Qatar Univ. Sci. Bull. (1986), 6 : 23-32

ON SOME SHRINKAGE TECHNIQUES FOR ESTIMATING THE PARAMETERS OF EXPONENTIAL DISTRIBUTION

M. A. Q. YOUSEF, M. S. ABU-SALIH and M. A. ALI

Yarmouk University, Irbid, Jordan

Key words : Shrinkage techniques, exponential distribution

ABSTRACT

A variety of shrinkage methods have been proposed for estimation of some unknown parameter by considering estimators based on a prior guess of the value of the parameter. We compare some of the shrunken estimators for the parameters μ and θ of the exponential distribution through simulation.

INTRODUCTION

In the estimation of an unknown parameter there often exists some form of prior knowledge about the parameter which one would like to utilize in order to get a better estimate. Thompson (1968) described a shrinkage technique for estimating the mean of a population. Mehta and Srinivasan (1971) proposed another class of shrunken estimator for the mean of a population and have shown that this class has better performance than that of Thompson (1968) in terms of mean squared error. Pandey and Singh (1977) and Pandey (1979) described shrinkage techniques for estimating the variance of a normal population. Lemmer (1981) considered a shrunken estimator for the parameter of the binomial distribution. His estimator is similar to the Pandey (1979) estimator for the variance of a normal distribution.

We consider a variety of shrinkage methods for estimating the parameters u and 0 of the exponential distribution. These estimators are compared through simulation.

ESTIMATORS CONSIDERED

Let the length of life X of a certain system be distributed as

$$f(x, \theta, \mu) = \frac{1}{\theta} \exp\left[-(x-\mu)/\theta\right], \qquad 0 \leqslant \mu \leqslant x, \qquad \theta > 0$$

A random sample of n such systems is subjected to test and the test terminated as soon as the first $r(\leqslant n)$ items fail. Let $\underset{\sim}{x} = \{x_{(1)} <..., < x_{(r)}\}$ be the first r ordered failure times. It is well known from Epstein and Sobel (1954) that

$$\hat{\theta} = [\sum_{i=1}^{r} x_{(i)} + (n-r)x_{(r)} - nx_{(1)}]/(r-1), r > 1$$

and

$$\hat{\mu} = x_{(1)} - \hat{\theta}/n,$$

are the minimum variance unbiased estimators of θ and u respectively. The variances of these estimators age given by

$$\text{var}(\hat{\theta}) = \theta^2/(r-1)$$

$$\text{var}(\hat{\mu}) = r\theta^2/n^2(r-1)$$

(see Bain (1978), p – 163).

The first estimator considered is :

$$\check{\mu}_T = \mu_o + C(\hat{\mu} - \mu_o) \quad 0 \leqslant c \leqslant 1 \tag{2.1}$$

where μ_o is the guessed value of μ. $\check{\mu}_T$ is the actual Thompson–type estimator. Thompson suggested to determine C from

$$\frac{\partial \, \text{MSE}(\check{u}_T)}{\partial C} = 0,$$

with $\text{MSE}(\check{\mu}_T) = E(\mu_T - \mu)^2$, the mean squared error of $\check{\mu}_T$. It follows that

$$C = (\mu - \mu_o)^2/[(\mu - \mu_o)^2 + \text{var}(\hat{\mu})] \tag{2.2}$$

In practice C in (2.2) is estimated by replacing the unknown parameters by their sample estimates. Substituting the estimated value of C in (2.1) we have

$$\hat{\mu}_T = \mu_o + (\hat{\mu} - \mu)^3/[(\hat{\mu} - \mu_o)^2 + r\hat{\theta}^2/n^2(r-1)] \tag{2.3}$$

Secondly, we consider the Mehta and Srinivasan–type estimator (cf. Mehta and Srinivasan (1971) for μ :

$$\hat{\mu}_M = \hat{\mu} - a(\hat{\mu} - \mu_o) \exp[-bn^2(r-1)(\hat{\mu} - \mu_o) / r\hat{\theta}^2], \qquad (2.4)$$

where a and b are suitably chosen positive constants a < b. No general guidance has been given on how a and b should be chosen.

Now we consider the Pandey-type estimator of μ :

$$\check{\mu}_P = a[K\mu + (1-K)\mu_o], \qquad 0 \leqslant K \leqslant 1 \qquad (2.5)$$

with K a constant specified by the experimentor according to his belief in μ_o and a is determined from $\partial MSE(\check{\mu}_p)/\partial a = 0$. It follows that $a = d_1\mu^2 / [K^2 \text{var}(\hat{\mu}) + d_1\mu^2]$ where $d_1 = K + (1-K)\mu_o/\mu$. Usually a is estimated by replacing the unknown parameters by their sample estimates.

Substituting the estimated value of a in (2.5) we obtain

$$\hat{\mu}_P = \hat{d}_1^2 \hat{\mu}^3 / [\hat{d}_1^2 \hat{\mu}^2 + K^2 r\hat{\theta}^2 / n^2(r-1)] \qquad (2.6)$$

with

$$\hat{d}_1 = [K + (1-K)\mu_o / \hat{\mu}].$$

Finally, we consider Lemmer-type estimator (cf. Lemmer (1981)) for μ :

$$\hat{\mu}_L = K\hat{\mu} + (1-K)\mu_o \qquad (2.7)$$

which follows from (2.5) if a = 1. Of all estimators considered, $\hat{\mu}_L$ is the simplest. As $\hat{\mu}_P$ and $\hat{\mu}_L$ depend on K, different values of K have been considered.

All the above approaches can be used to define variety of shrunken estimators for the parameter θ. We present all the estimators considered in the following table.

COMPARISON OF ESTIMATORS

Simulation experiments are used to estimate the mean squared errors for the five estimators of μ and θ. The procedure is described below :

Table 1

Shrunken Estimators For u and θ

Parameter	Type of Estimator	Estimator
Location Parameter μ	Thompson	$\hat{\mu}_T = \mu_o + (\hat{\mu} - \mu_o)^3 / [(\hat{\mu} - \mu_o)^2 + r\hat{\theta}^2 / \{n^2 (r-1)\}]$
	Mehta–Srinivasan	$\hat{\mu}_M = \hat{\mu} - a(\hat{\mu} - \mu_o) \exp[-\{bn^2 (r-1)(\hat{\mu} - \mu_o) / r\hat{\theta}^2\}]$
	Pandey	$\hat{\mu}_P = \hat{d}_1^2 \hat{\mu}^3 / [\hat{d}^2 \hat{\mu}^2 + K^2 r \hat{\theta}^2 / n^2 (r-1)]$
	Lemmer	$\hat{\mu}_L = K\hat{\mu} + (1-K)\mu_o$
Scale parameter θ	Thompson	$\hat{\theta}_T = \theta_o + (\hat{\theta} - \theta_o)^3 / [(\hat{\theta} - \theta_o)^2 + \hat{\theta}^2 / (r-1)]$
	Mehta–Sirivasan	$\hat{\theta}_M = \hat{\theta} - a(\hat{\theta} - \theta_o) \exp[-\{b(r-1)(\hat{\theta} - \theta_o) / \hat{\theta}^2\}]$
	Pandey	$\hat{\theta}_P = \hat{d}_2^2 \mu^3 / [\hat{d}_2^2 \hat{\theta}^2 + K^2 \hat{\theta}^2 / (r-1)]$
	Lemmer	$\hat{\theta}_L = K\hat{\theta} + (1-K)\theta_o$

$\hat{d}_1 = K + (1-K)\mu_o / \hat{\mu}$, $\hat{d}_2 = K + (1-K)\theta_o / \hat{\theta}$, K is a known constant between zero and one, a and b are positive constants $a < b$, μ_o and θ_o are the guessed values for μ and θ respectively.

We generate a random sample of size n from a two-patameter exponential distribution,

$$f(x,\theta,\mu)=\frac{1}{\theta}\exp[-(x-\mu)/\theta], \qquad 0 \leqslant \mu \leqslant x, \qquad \theta > 0$$

with $\mu = 80$, and $\theta = 7.0$. The vector

$$\underset{\sim}{x} = \{x_{(1)} < x_{(2)} < \dots < x_{(r)}\}$$

of the first r-ordered observation is recorded. Then the minimum variance unbiased estimators $\hat{\mu}$ and $\hat{\theta}$ of μ and θ respectively are computed using the following formulas.

$$\hat{\mu} = x_{(1)} - \hat{\theta}/n$$

and

$$\hat{\theta} = [\sum_{i=1}^{r} x_{(i)} + (n-r)x_{(r)} - nx_{(1)}]/(r-1).$$

For a known constant K between zero and one and for specific values of μ_o and θ_o, the quantities

$$\hat{d}_1 = K+(1-k)\mu_o/\hat{\mu} \quad \text{and} \quad \hat{d}_2 = K+(1-K)\theta_o/\hat{\theta}$$

are obtained. Then the estimators $\hat{\mu}_T, \hat{\mu}_M, \hat{\mu}_P$ and $\hat{\mu}_L$ of u are computed using the relations (2.3), (2.4), (2.6) and (2.7). Similarly, the estimators $\hat{\theta}_T$, $\hat{\theta}_M$, $\hat{\theta}_P$ and $\hat{\theta}_L$ of θ are obtained using the formulas shown in Table – 1.

Monte Carlo experiments are repeated 500 times. The average of the 500 sample values of each squared error, e.g. $(\hat{\mu}-\mu)^2$, is taken as an estimate of the corresponding mean squared error which is denoted by MSE (.).

The estimates of the mean squared errors of the various estimators of μ and θ and the relative efficiencies, e.g.

$$R(\hat{\mu}_T/\hat{\mu}) = MSE(\hat{\mu}_T)/MSE(\hat{\mu})$$

are calculated for n = 30, r = 10, 20, 30, K = 0.20, 0.70, a = 1, 5, b = 20, 50, $\mu = 80, \theta = 7.0, \mu_o = 70$, and $\theta_o = 5.0$.

Results of the simulation experiments are given in Table 2 – 3.

CONCLUSION

Although the results derived above apply strictly to only very limited cases, they are suggestive of some general conclusions regarding the relative efficiencies of the various methods. Note from Tables 2 that the MSE of $\hat{\mu}_T$ are always smaller than that of other estimators. It is obvious that $\hat{\mu}_T$, $\hat{\mu}_M$, and $\hat{\mu}_L$ have smaller mean squared error than the minimum variance unbiased estimator $\hat{\mu}$.

The mean squared error of $\hat{\mu}_P$ is always higher than the MSE of $\hat{\mu}$. The advantages of $\hat{\mu}_T$ and $\hat{\mu}_L$ are most marked when r is small.

Further, the comparison statistics in Table – 3 show that the MSE of $\hat{\theta}_T$, $\hat{\theta}_M$, and $\hat{\theta}_L$ are always smaller than the MSE of $\hat{\theta}$, the minimum variance unbiased estimator. The mean squared error of $\hat{\theta}_P$ is always greater than the MSE of $\hat{\theta}$. The MSE of Thompson–type estimator is smaller than those of the remaining estimators.

Table 2

Relative Efficiencies of Various Shrunken Estimators of μ

Sample size $n = 30$, $\mu = 80$, $\theta = 7.0$, $\mu_o = 70$, $\theta_o = 5.0$

No. of failures r	**Average of M.V.U.E. of μ**	**$R(\hat{\mu}_T/\hat{\mu})$**	**a=1, b=20**	**a=5, b=50**	**K = 0.20**	**K = 0.50**	**K = 0.20**	**K = 0.50**
			$R(\hat{\mu}_M/\hat{\mu})$		**$R(\hat{\mu}_P/\hat{\mu})$**		**$R(\hat{\mu}_L/\hat{\mu})$**	
10	79.871	4.43×10^{-4}	$6.54 \times 1-^{-2}$	0.544	3.707	3.258	2.49×10^{-2}	0.249
20	80.349	2.52×10^{-3}	0.226	0.998	3.788	2.076	2.50×10^{-2}	0.255
30	79.841	1.18×10^{-2}	0.016	0.012	4.004	2.736	2.48×10^{-2}	0.245

Table 3

Relative Efficiencies of Various Shrunken Estimators of θ

Sample size $n = 30$, $\mu = 80$, $\theta = 7.0$, $\theta_o = 5.0$

No. of failures	**Average of M.V.U.E.**	**$R(\hat{\theta}_T/\hat{\theta})$**	**a=1, b=20**	**a=5, b=50**	**K = 0.20**	**K = 0.70**	**K = 0.20**	**K = 0.70**
			$R(\hat{\theta}_M/\hat{\theta})$		**$R(\hat{\theta}_P/\hat{\theta})$**		**$R(\hat{\theta}_L/\hat{\theta})$**	
10	7.358	9.34×10^{-3}	0.066	0.769	1.150	1.270	2.33×10^{-3}	0.095
20	7.338	1.48×10^{-2}	0.645	1.0	3.965	1.435	2.53×10^{-2}	0.223
30	7.207	4.13×10^{-2}	0.876	1.0	3.978	1.753	2.50×10^{-2}	0.256

REFERENCES

Lemmer, H. H. 1981. Note on shrinkage estimators for the Binomial distribution. Commum. Statist. Theory. Meth. A 10 (10), 1017 – 1027.

Mehta, J. S.; and **Srinivasan, R. 1971.** Estimation of the mean by shrinkage to a point Jr. Amer. Statist. Assoc. 66, 86 – 90.

Pandey, B. N. 1979. On shrinkage estimation of normal population variance. Commun. Statist. Theory. Meth. A 8 (4), 359 – 365.

Pandey, B. N. and **Singh, J. 1977.** Estimation of variance of normal population prior information. Jr. Indian Statist. Assoc. 15, 141 – 50.