

On The Estimation of Survival Function and Parameter Exponential Life Time Distribution

Hadeel Salim Al-Kutubi and Noor Akma Ibrahim
Institute for Mathematical Research, University Putra Malaysia,
43400, UPM, Serdang, Selangor, Malaysia

Abstract: Problem statement: The study and research of survival or reliability or life time belong to the same area of study but they may belong to a different area of application. In survival analysis one can use several life time distribution, exponential distribution with mean life time θ is one of them. To estimate this parameter and survival function we must be used estimation procedures with less MSE and MPE. **Approach:** The only statistical theory that combined modeling inherent uncertainty and statistical uncertainty is Bayesian statistics. The theorem of Bayes provided a solution to how learn from data. Bayes theorem was depending on prior and posterior distribution and standard Bayes estimator depends on Jeffery prior information. In this study we annexed Jeffery prior information to get the modify Bayes estimator and then compared it with standard Bayes estimator and maximum likelihood estimator to find the best (less MSE and MPE). **Results:** when we derived Bayesian and Maximum likelihood of the scale parameter and survival functions. Simulation study was used to compare between estimators and Mean Square Error (MSE) and Mean Percentage Error (MPE) of estimators are computed. **Conclusion:** The new proposed estimator of modify Bayes estimator in parameter and survival function was the best estimator (less MSE and MPE) when we compared it with standard Bayes and maximum likelihood estimator.

Key words: Extension of Jeffery prior, survival function, exponential distribution, maximum likelihood estimation

INTRODUCTION

Survival analysis refers to the techniques used to study the time to occurrence of some event in a population and is often called time to event analysis. The survival function $S(t)$ is defined as the probability that human being will surviving at time period t . The two parameter of exponential distribution was used when failure will never occur prior to some specified time t_0 . The parameter t_0 is a location parameter that shifts the distribution an amount equal to t_0 the right on the time (horizontal axis). Now $t \geq t_0$, the probability density function of exponential distribution becomes:

$$f(t; \theta) = \frac{1}{\theta} \exp\left(-\frac{t-t_0}{\theta}\right)$$

and the survival function is given by:

$$S(t) = \int_t^{\infty} f(u; \theta) du$$

Some studies deal with this aspects in detail and provide the conceptual framework and methodology for such analysis. In^[1] show how the classical estimators can be obtained from various choices made within a Bayesian framework. A Bayesian approach to hypothesis testing is presented by Fernandez^[2]. He studied the problem of estimating the mean lifetime, hazard rate and survival function of exponential lifetime model. Hahn^[3] show that it is not always Jeffrey' prior applied to panel models with fixed effects yields posterior inference free from the incidental parameter problem. In^[4] derived Bayesian and non Bayesian estimators of the shape parameter and reliability in case of complete and censored samples and the MSE of estimators are computed.

The main aim of this study is to obtain maximum likelihood estimator and Bayesian estimators of the parameters and survival functions under Jeffrey's prior and extension of Jeffrey's prior and then compared between them to find the best estimator (less MSE and MPE). Based on simulation study, the Mean Square Error (MSE) and Mean Percentage Error (MPE) are

Corresponding Author: Hadeel Salim Al-Kutubi, Institute for Mathematical Research, University Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia

compared with those of their Maximum likelihood and Bayes estimator under prior distribution and loss function.

MATERIALS AND METHODS

Maximum likelihood estimation: We have the set of random lifetime t_1, t_1, \dots, t_n and a vector of unknown parameters $\theta = (\theta_1, \dots, \theta_n)$, the i th element of the score vector is:

$$U_i(\theta) = \frac{\partial \ln L(t, \theta)}{\partial \theta_i}, i = 1, 2, \dots, p$$

Now we can find the maximum likelihood estimator by using exponential population with parameter θ . The probability density function of two parameters exponential distributions is given by:

$$f(t; \theta) = \frac{1}{\theta} \exp\left(-\frac{t-t_0}{\theta}\right), \theta > 0$$

The likelihood function is:

$$\begin{aligned} L(t, \theta) &= \prod_{i=1}^n f(t_i, \theta) = \prod_{i=1}^n \frac{1}{\theta} \exp\left(-\frac{t_i - t_0}{\theta}\right) \\ &= \theta^{-n} \exp\left(-\frac{\sum_{i=1}^n t_i - t_0}{\theta}\right) \end{aligned}$$

The score vector is:

$$U(\theta) = \frac{\partial \ln L(t, \theta)}{\partial \theta} = \frac{-n}{\hat{\theta}} + \frac{\sum_{i=1}^n t_i - t_0}{\hat{\theta}^2}$$

Let $U(\theta) = 0$, then the maximum likelihood estimator is:

$$\hat{\theta}_M = \frac{\sum_{i=1}^n t_i - t_0}{n} \tag{1}$$

Since the maximum likelihood estimator is invariant and one to one mapping. The maximum likelihood estimator of survival function is:

$$\hat{s}_M(t) = \exp\left(-\frac{t - t_0}{\hat{\theta}}\right) \tag{2}$$

Bayes estimation:

Estimator of parametric: We can find Bayes estimator of parametric distribution by using posterior distribution, depend on joint probability density function and marginal probability density function, so the posterior distribution is given by:

$$\Pi(\theta | t_1, \dots, t_n) = \frac{H(t_1, \dots, t_n, \theta)}{p(t_1, \dots, t_n)}$$

Now the joint probability density function $f(t_1, \dots, t_n, \theta)$ is given by:

$$H(t_1, \dots, t_n, \theta) = L(t_1, \dots, t_n | \theta) g(\theta) = \frac{k\sqrt{n}}{\theta^{n+1}} \exp\left(-\frac{\sum_{i=1}^n t_i - t_0}{\theta}\right)$$

And the marginal probability density function of t_1, \dots, t_n is given by:

$$p(t_1, \dots, t_n) = \int H(t_1, \dots, t_n, \theta) d\theta = \frac{(k\sqrt{n})(n-1)!}{\left(\sum_{i=1}^n t_i - t_0\right)^n}$$

So the posterior distribution (conditional probability density function of θ) is given:

$$\begin{aligned} \Pi(\theta | t_1, \dots, t_n) &= \frac{H(t_1, \dots, t_n, \theta)}{p(t_1, \dots, t_n)} \\ &= \frac{\exp\left(-\frac{\sum_{i=1}^n t_i - t_0}{\theta}\right)}{\theta^{n+1}} \frac{\left(\sum_{i=1}^n t_i - t_0\right)^n}{(n-1)!} \end{aligned}$$

By using Risk function, we can obtain the Bayes estimator, such that:

$$\begin{aligned} R(\hat{\theta}, \theta) &= E\ell(\hat{\theta}, \theta)^2 \\ &= \int_0^\infty \ell(\hat{\theta}, \theta) \Pi(\theta | t_1, \dots, t_n) d\theta \\ &= c\hat{\theta}^2 - 2c\hat{\theta} \int_0^\infty \left(\frac{\sum_{i=1}^n t_i - t_0}{(n-1)!}\right) \theta^{-n} \exp\left(-\frac{\sum_{i=1}^n t_i - t_0}{\theta}\right) d\theta + \zeta(\theta) \end{aligned}$$

Let $\frac{\partial R(\hat{\theta}, \theta)}{\partial \hat{\theta}} = 0$, then the Bayes estimator is:

$$\hat{\theta}_B = \frac{\left(\sum_{i=1}^n t_i - t_0\right)^n}{(n-1)!} \int_0^\infty \theta^{-n} \exp\left(-\frac{\sum_{i=1}^n t_i - t_0}{\theta}\right) d\theta \tag{3}$$

$$\hat{\theta}_B = \frac{\sum_{i=1}^n t_i - t_0}{n-1}$$

Estimator of survival function: We can find the estimator of survival function; also by depend on posterior distribution, such that:

$$\begin{aligned} \Pi(\theta|t_1, \dots, t_n) &= \frac{H(t_1, \dots, t_n, \theta)}{p(t_1, \dots, t_n)} \\ &= \frac{\exp\left(-\frac{\sum_{i=1}^n t_i - t_0}{\theta}\right) \left(\sum_{i=1}^n t_i - t_0\right)^n}{\theta^{n+1} (n-1)!} \end{aligned}$$

So, the estimator of survival function is given by:

$$\begin{aligned} \hat{s}_B(t) &= \int_0^\infty \exp\left(\frac{-t_i + t_0}{\theta}\right) \Pi(\theta|t_1, \dots, t_n) d\theta \\ &= \int_0^\infty \exp\left(\frac{-t_i + t_0}{\theta}\right) \frac{\exp\left(-\frac{\sum_{i=1}^n t_i - t_0}{\theta}\right) \left(\sum_{i=1}^n t_i - t_0\right)^n}{\theta^{n+1} (n-1)!} d\theta \\ \hat{S}_B(t) &= \left(\frac{\sum_{i=1}^n t_i - t_0}{t_i - t_0 + \sum_{i=1}^n t_i - t_0}\right)^n \end{aligned} \tag{4}$$

Modify Bayes estimation:

Estimator of parametric: The extension of Jeffery prior is given by:

$$g(\theta) \propto [I(\theta)]^{c_1}, \quad c_1 \in \mathbb{R}^+$$

So:

$$g(\theta) = k \frac{n^{c_1}}{\theta^{2c_1}}$$

The same way above, we can find the modify Bayes estimator by depend on posterior distribution as following: The joint probability density function is given by:

$$\begin{aligned} H(t_1, \dots, t_n, \theta) &= \prod_{i=1}^n f(t_i|\theta) g(\theta) \\ &= \frac{1}{\theta^n} \exp\left(-\frac{\sum_{i=1}^n t_i - t_0}{\theta}\right) \frac{kn^{c_1}}{\theta^{2c_1}} \\ &= \frac{kn^{c_1}}{\theta^{n+2c_1}} \exp\left(-\frac{\sum_{i=1}^n t_i - t_0}{\theta}\right) \end{aligned}$$

And the marginal probability density function is given by:

$$\begin{aligned} p(t_1, \dots, t_n) &= \int_0^\infty H(t_1, \dots, t_n, \theta) d\theta \\ &= \frac{kn^{c_1} (n + 2c_1 - 2)!}{\left(\sum_{i=1}^n t_i - t_0\right)^{n+2c_1-1}} \end{aligned}$$

Then the posterior distribution given the data t_1, \dots, t_n is given by:

$$\begin{aligned} \Pi(\theta|t_1, \dots, t_n) &= \frac{\frac{kn^{c_1}}{\theta^{n+2c_1}} \exp\left(-\frac{\sum_{i=1}^n t_i - t_0}{\theta}\right)}{\frac{kn^{c_1} (n + 2c_1 - 2)!}{\left(\sum_{i=1}^n t_i - t_0\right)^{n+2c_1-1}}} \\ &= \frac{\theta^{-n-2c_1} \exp\left(-\frac{\sum_{i=1}^n t_i - t_0}{\theta}\right)}{\left(\sum_{i=1}^n t_i - t_0\right)^{1-n-2c_1} (n + 2c_1 - 2)!} \end{aligned}$$

By risk function, we can obtain modify Bayes estimator, such that:

$$R(\hat{\theta}, \theta) = \int_0^\infty c(\hat{\theta} - \theta)^2 \frac{\theta^{-n-2c_1} \exp\left(-\frac{\sum_{i=1}^n t_i}{\theta}\right)}{\left(\sum_{i=1}^n t_i\right)^{1-n-2c_1} (n + 2c_1 - 2)!} d\theta$$

Let:

$$\frac{\partial R(\hat{\theta}, \theta)}{\partial \hat{\theta}} = 0$$

then

$$\hat{\theta}_{B_M} = \frac{\sum_{i=1}^n t_i}{n + 2c_1 - 2} \tag{5}$$

Estimator of survival function: By using posterior distribution, we can found the survival function such that:

$$\begin{aligned} \hat{s}_{BE}(t) &= \int_0^\infty \exp\left(-\frac{t_i - t_0}{\theta}\right) \Pi(\theta | t_1, \dots, t_n) d\theta \\ &= \int_0^\infty \exp\left(-\frac{t_i - t_0}{\theta}\right) \frac{\theta^{-n-2c_1} \exp\left(-\frac{\sum_{i=1}^n t_i - t_0}{\theta}\right)}{\left(\sum_{i=1}^n t_i - t_0\right)^{1-n-2c_1} (n + 2c_1 - 2)!} d\theta \tag{6} \\ &= \left(\frac{t_i - t_0 + \sum_{i=1}^n t_i - t_0}{\sum_{i=1}^n t_i - t_0}\right)^{1-n-2c_1} \end{aligned}$$

Simulation: Two experiments were designed to compare the methods of estimation, Maximum likelihood and Bayesian approach, using the ordinary and proposed tool based on varieties of parameters and sample sizes. Using Mean Square Error (MSE) and Mean Percentage Error (MPE) as criteria for comparisons were some new results presented.

Where:

$$MSE(\hat{\theta}) = \frac{\sum_{i=1}^{1000} (\hat{\theta}_i - \theta)^2}{R}$$

$$MPE(\hat{\theta}) = \frac{\left[\sum_{i=1}^R \frac{|\hat{\theta}_i - \theta|}{\theta} \right]}{R}$$

RESULTS

We chooses the samples with size $n = 25, 50, 100$ with varieties of parameter value namely 0.5, 1, 1.5, 2. The size of replication is $R = 1000$. After the parameters and survival functions were estimated, using formulas Mean Square Error (MSE) and Mean

Percentage Error (MPE) to compare between the methods of estimator.

DISCUSSION

In the simulation 1, when we compared parametric estimators of exponential distribution in maximum likelihood and Bayes method, we found the Maximum Likelihood Estimator (MLE) is the best with 100% percentage valued in Mean Square Error (MSE) in Table 1 and Mean Percentage Error (MPE) in Table 2.

Table 1: Parameter Estimation using MLE and Bayes with respect to MSE

Size	θ	$\hat{\theta}_M$	$\hat{\theta}_M$	Best
25	0.5	0.0101	0.0115	$\hat{\theta}_M$
	1.0	0.0395	0.0441	$\hat{\theta}_M$
	1.5	0.0898	0.1029	$\hat{\theta}_M$
	2.0	0.1571	0.1765	$\hat{\theta}_M$
50	0.5	0.0049	0.0052	$\hat{\theta}_M$
	1.0	0.0189	0.0201	$\hat{\theta}_M$
	1.5	0.0445	0.0469	$\hat{\theta}_M$
	2.0	0.0778	0.0819	$\hat{\theta}_M$
100	0.5	0.0025	0.0026	$\hat{\theta}_M$
	1.0	0.0101	0.0104	$\hat{\theta}_M$
	1.5	0.0212	0.0218	$\hat{\theta}_M$
	2.0	0.0406	0.0417	$\hat{\theta}_M$

Table 2: Parameter Estimation using MLE and Bayes with respect to MPE

Size	θ	$\hat{\theta}_M$	$\hat{\theta}_M$	Best
25	0.5	0.1600	0.1679	$\hat{\theta}_M$
	1.0	0.1595	0.1659	$\hat{\theta}_M$
	1.5	0.1579	0.1661	$\hat{\theta}_M$
	2.0	0.1544	0.1625	$\hat{\theta}_M$
50	0.5	0.1124	0.1144	$\hat{\theta}_M$
	1.0	0.1108	0.1134	$\hat{\theta}_M$
	1.5	0.1116	0.1139	$\hat{\theta}_M$
	2.0	0.1095	0.1121	$\hat{\theta}_M$
100	0.5	0.0800	0.0812	$\hat{\theta}_M$
	1.0	0.0808	0.0816	$\hat{\theta}_M$
	1.5	0.0782	0.0790	$\hat{\theta}_M$
	2.0	0.0803	0.0808	$\hat{\theta}_M$

Table 3: Survival Estimation using MLE and Bayes with respect to MSE

Size	S	$\hat{S}_M(t)$	$\hat{S}_B(t)$	Best
25	0.5	0.0031000	0.0035000	$\hat{S}_M(t)$
	1.0	0.0032000	0.0034000	$\hat{S}_M(t)$
	1.5	0.0030000	0.0032000	$\hat{S}_M(t)$
	2.0	0.0031000	0.0030000	$\hat{S}_M(t)$
50	0.5	0.0015000	0.0017000	$\hat{S}_M(t)$
	1.0	0.0012000	0.0014000	$\hat{S}_M(t)$
	1.5	0.0013000	0.0015000	$\hat{S}_M(t)$
	2.0	0.0012000	0.0015000	$\hat{S}_M(t)$
100	0.5	0.0000748	0.0000741	$\hat{S}_M(t)$
	1.0	0.0000763	0.0000760	$\hat{S}_M(t)$
	1.5	0.0000714	0.0000711	$\hat{S}_B(t)$
	2.0	0.0000760	0.0000750	$\hat{S}_B(t)$

Table 4: Parameter Estimation using MLE and Bayes with respect to MPE

Size	S	$\hat{S}_M(t)$	$\hat{S}_B(t)$	Best
25	0.5	0.1865	0.2152	$\hat{S}_M(t)$
	1.0	0.1825	0.2101	$\hat{S}_M(t)$
	1.5	0.1908	0.2209	$\hat{S}_M(t)$
	2.0	0.0802	0.2082	$\hat{S}_M(t)$
50	0.5	0.1211	0.1291	$\hat{S}_M(t)$
	1.0	0.1186	0.1272	$\hat{S}_M(t)$
	1.5	0.1203	0.1289	$\hat{S}_M(t)$
	2.0	0.1176	0.1257	$\hat{S}_M(t)$
100	0.5	0.0836	0.0868	$\hat{S}_M(t)$
	1.0	0.0839	0.0866	$\hat{S}_M(t)$
	1.5	0.8060	0.0833	$\hat{S}_M(t)$
	2.0	0.0833	0.0858	$\hat{S}_M(t)$

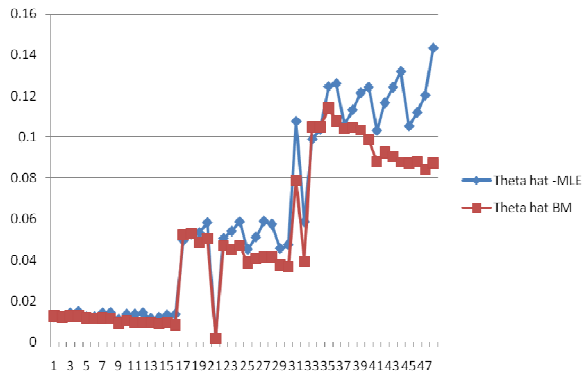


Fig. 1: Comparison parametric estimators in different sample size using MSE

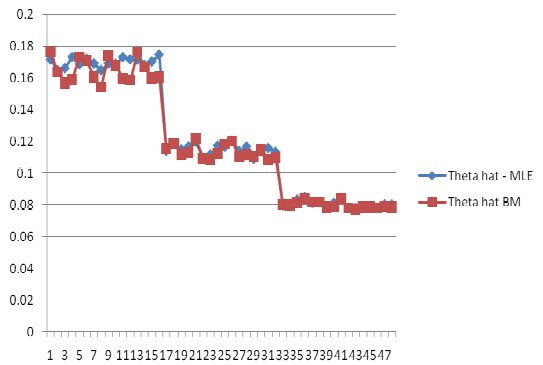


Fig. 2: Comparison parametric estimators in different sample size using MPE

Also, when compared the estimators of survival function for exponential distribution, we found the best estimator is MLE with 66.66% percentage valued with MSE in Table 3 and MPE with 100 % in Table 4.

In the simulation 2, when we compared the best estimator (MLE) in simulation 1 with proposed parameter (modify Bayes), we found the best estimator is proposed estimator, that is explain in MSE (Fig. 1) and MPE (Fig. 2),because this proposed gives the value of MSE and MPE less than the valued in other estimators. We have the same thing if we compared survival estimators (proposed and MLE), we found the proposed survival estimator with MSE in Fig. 3 and MPE in Fig. 4 is the best estimator.

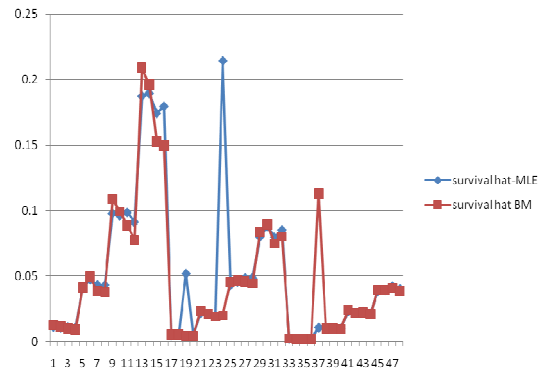


Fig. 3: Comparison survival estimators in different sample size using MSE

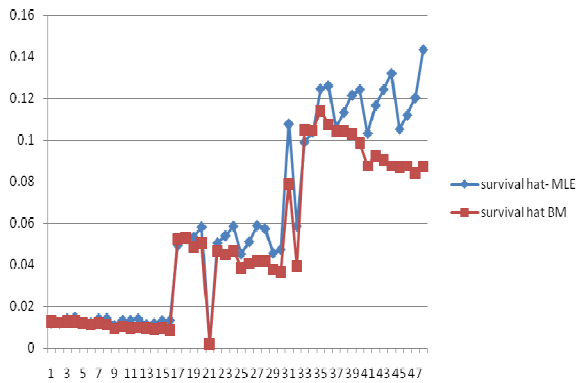


Fig. 4: Comparison survival estimators in different sample size using MPE

CONCLUSION

The new proposed estimator of modify Bayes estimator in parameter and survival function is the best estimator (less MSE and MPE) when we compared it with standard Bayes and Maximum likelihood estimator.

REFERENCE

1. Elfessi, A. and D.M. Reineke, 2001. Bayesian look at classical estimation: The exponential distribution. *J. Stat. Educ.*, 9: 1-7. <http://www.amstat.org/publications/jse/v9n1/elfessi.html>
2. Fernandez, A.J., 2000. Estimation and hypothesis testing for exponential lifetime models with double censoring and prior information. *J. Econ. Soc. Res.*, 2: 1-17. <http://www.fatih.edu.tr/~jesr/TestingforExponentialLifetimeModels.pdf>
3. Hahn, J., 2004. Dose Jeffrey’s prior alleviate the incidental parameter problem? *Econ. Lett.*, 82: 135-138. DOI: 10.1016/j.econlet.2003.04.001.
4. Shawky, A. and R. Bakoban 2008. Bayesian and non Bayesian estimations on the exponentiated Gamma distribution. *Applied Math. Sci.*, 2: 2521-2530. <http://www.m-hikari.com/ams/ams-password-2008/ams-password49-52-2008/bakobanAMS49-52-2008.pdf>