On Bayesian Estimation of an Exponential Distribution

Adegoke, T. M.[1], Adegoke, G. K.[2]; Yahaya, A. M.[3], Uthman, K. T.[4]; Odigie, A.D.[5]

[1]Department of Statistics, University of Ilorin, Nigeria.
[2]Department of Mathematical Science, Nasarawa State University, Keffi, Nigeria
[3]Department of Mathematics and Statistics, University of Maiduguri, Maiduguri, Nigeria
[4]National Center for Gentic Resources and Biotechnology, Ibadan, Nigeria.
[5]Center for Human Right and Climate Change Research, Edo State.
e-mail: adegoketaiwom@gmail.com[1]

Abstract **— The Bayesian estimation of unknown parameter of exponential distribution was examined under different priors using Markov Chain Monte Carlo simulation method in OpenBUGS by incorporating a module into OpenBUGS. The posterior distributions for the unknown parameter of the Exponential distribution were derived using the following priors: Uniform distribution, Jeffery's Prior, Gamma distribution, Gamma-Chi-square distribution, Gamma - Exponential distribution and Chi-square-Exponential distribution. The R functions were developed to study the various statistical properties of Exponential distribution and the output of Markov Chain Monte Carlo simulation samples were generated from OpenBUGS.**

Keywords- *Poisson distribution, prior distribution, posterior distribution, OpenBUGS.*

I. Introduction

Clima
Exponential distribution is a special case of Gamma distribution

$$G(a,\gamma) = \frac{\gamma^a}{\Gamma(a)} e^{-\gamma x} x^{a-1} \quad a > 0, \gamma > 0 \qquad (1)$$

When a = 1, it becomes an exponential distribution with probability mass function of the form

$$f(x,\gamma) = \gamma e^{-\gamma x}; \; x = 0,1,2,3 \ldots and \,, \gamma > 0 \qquad (2)$$

The exponential distribution may be viewed as a continuous counterpart of the geometric distribution, which describes the number of Bernoulli trials necessary for a "discrete" process to change state. In contrast, the exponential distribution describes the time for a continuous process to change state. In real-world scenarios, the assumption of a constant rate (or probability per unit time) is rarely satisfied. For example, the rate of incoming phone calls differs according to the time of day. In queuing theory, the service times of agents in a system (e.g. how long it takes for a bank teller etc. to serve a customer) are often modeled as exponentially distributed variables. (The interarrival of customers for instance in a system is typically modeled by the Poisson distribution in most management science textbooks.) The length of a process that can be thought of as a sequence of several independent tasks is better modeled by a variable following the Erlang distribution (which is the distribution of the sum of several independent exponentially distributed variables).

Reliability theory and reliability engineering also make extensive use of the exponential distribution. Because of the "memoryless" property of this distribution, it is well-suited to model the constant hazard rate portion of the bathtub curve used in reliability theory. It is also very convenient because it is so easy to add failure rates in a reliability model. The exponential distribution is however not appropriate to model the overall lifetime of organisms or technical devices, because the "failure rates" here are not constant: more failures occur for very young and for very old systems.

In physics, if you observe a gas at a fixed temperature and pressure in a uniform gravitational field, the heights of the various molecules also follow an approximate exponential distribution.

.

II. Reseach Methodology

A. *Posterior Distribution of Unknown Parameter θ Using Uniform Prior Distribution*

If $x_1, x_2, \ldots, x_n$ are iid observations from an (2), then the likelihood function is

$$L(\gamma) = \gamma^n e^{\gamma \Sigma x} \qquad (3)$$

[3,4] found that it worked exceptionally well to always select the prior for θ to be constant. Consider the uniform prior

$$p(\gamma) \propto 1 \,; 0 < \gamma < \infty \qquad (4)$$

The posterior distribution $p(\gamma|X)$ of the parameter γ is found using (3) and (4) as $p(\gamma|X) \; \alpha \; p(X|\gamma)$

$$p(\gamma|X) \propto \gamma^n e^{-\gamma \Sigma x}$$

which is the density function of a Gamma distribution of

$$p(\gamma|X) = \frac{(\sum x)^{n+1}}{\Gamma(n+1)}\gamma^{n}e^{-\gamma\sum x} \quad (5)$$

with parameters (n+1 , $\sum x$)

$$Mean = E(X) = \frac{\sum x}{n+1}$$

$$Var(X) = \frac{\sum x}{(n+1)^2}$$

B. *Posterior Distribution of Unknown Parameter γ Using Jefferey's Prior*

According to [1] the Jeffrey's prior for the parameter γ having distribution (2) is

$$p(\gamma) \propto \frac{1}{\gamma} \quad (6)$$

And the posterior distribution p($\gamma|X$) of the parameter γ is derived using (3) and (6) as

$$p(\gamma|y) \ \propto \ \gamma^{n-1}e^{-\gamma\sum x}$$

which is the density function of a Gamma distribution of

$$p(\gamma|X) = \frac{(\sum x)^{n}}{\Gamma(n)}\gamma^{n-1}e^{-\gamma\sum x} \quad (7)$$

with parameters (n , $\sum x$).

$$Mean = E(X) = \frac{\sum x}{n}$$

$$Var(X) = \frac{\sum x}{n^2}$$

C. *Posterior Distribution of Unknown Parameter γ Using Gamma Prior*

The single prior distribution of γ is a Gamma distribution with hyper parameters α_1 and β_1 is

$$\boldsymbol{p(\gamma)} = \frac{{\beta_1}^{\alpha_1}}{\Gamma(\alpha_1)}\gamma^{\alpha_1-1}e^{-\beta_1\gamma};\ \alpha_1 > 0\,,\beta_1 > 0 \quad (8)$$

And the posterior distribution p($\gamma|X$) of the parameter γ is derived using (3) and (8) as

$$p(\gamma|X) \ \propto \ \gamma^{n+\alpha_1-1}e^{-\gamma(\beta_1+\sum x)}$$

which is the density function of a Gamma distribution of

$$p(\gamma|X) = \frac{(\beta_1+\sum x)^{n+\alpha_1}}{\Gamma(n+\alpha_1)}\gamma^{n+\alpha_1-1}e^{-\gamma(\sum x+\beta_1)} \quad (9)$$

With parameters $(n+\alpha_1, \sum x+\beta_1)$

$$Mean = E(X) = \frac{\beta_1+\sum x}{n+\alpha_1}$$

$$Var(X) = \frac{\beta_1+\sum x}{(n+\alpha_1)^2}$$

D. *Posterior Distribution of Unknown Parameter γ Using Gamma-Chi-Square Prior*

Assume that the prior distribution of γ is a Gamma distribution with hyper parameters α_2 and β_2 as

$$p_1(\gamma) = \frac{{\beta_2}^{\alpha_2}}{\Gamma(\alpha_2)}\gamma^{\alpha_2-1}e^{-\beta_2\gamma};\ \alpha_2 > 0\,,\beta_2 > 0 \quad (10)$$

The second prior assumed is a Chi-square distribution with hyper parameter ψ_1 given by

$$p_2(\gamma) = \frac{1}{2^{\frac{\psi_1}{2}}\Gamma\left(\frac{\psi_1}{2}\right)}\gamma^{\frac{\psi_1}{2}-1}e^{-\frac{\gamma}{2}};\psi_1 > 0, \gamma > 0 \quad (11)$$

A double prior is defined for γ by combining the two priors in (10) and (11) as:

$$p(\gamma) \propto \ p_1(\gamma)p_2(\gamma)$$

$$p(\gamma) \propto \ \gamma^{\alpha_2+\frac{\psi_1}{2}-2}e^{-\gamma\left(\frac{1}{2}+\beta_2\right)} \quad (12)$$

And the posterior distribution p($\gamma|X$) of the parameter γ is derived using (3) and (12) as

$$p(\gamma|X) \ \propto \ \gamma^{\left(\alpha_2+\frac{\psi_1}{2}+n-1\right)-1}e^{-\gamma(\beta_1+\sum x+\frac{1}{2})}$$

which is the density function of a Gamma distribution of

$$p(\gamma|X) = \frac{\left(\beta_1+\sum x+\frac{1}{2}\right)^{\alpha_2+\frac{\psi_1}{2}+n-1}}{\Gamma(\alpha_2+\frac{\psi_1}{2}+n-1)}\gamma^{\left(\alpha_2+\frac{\psi_1}{2}+n-1\right)-1}e^{-\gamma(\beta_1+\sum x+\frac{1}{2})} \quad (13)$$

With parameters $(\alpha_2+\frac{\psi_1}{2}+n-1, \beta_1+\sum x+\frac{1}{2})$

$$Mean = E(X) = \frac{\beta_2+\sum x+\frac{1}{2}}{n+\alpha_2+\frac{\psi_1}{2}-1}$$

$$Var(X) = \frac{\beta_2+\sum x+\frac{1}{2}}{\left(n+\alpha_2+\frac{\psi_1}{2}-1\right)^2}$$

E. *Posterior Distribution of Unknown Parameter γ Using Gamma-Exponential Prior*

The double prior for γ is defined to be a Gamma distribution with hyper parameters (α_3, β_3) and an exponential distribution with hyper parameter δ_1 as

$$p_3(\gamma) = \frac{{\beta_3}^{\alpha_3}}{\Gamma(\alpha_3)}\gamma^{\alpha_3-1}e^{-\beta_3\gamma};\ \alpha_3 > 0\,,\beta_3 > 0 \quad (14)$$

$$p_4(\gamma) = \ \delta_1 e^{\,\delta_1\gamma};\ \delta_1 > 0 \quad (15)$$

A double prior is defined for γ by combining the two priors in (14) and (15) as:

$$p(\gamma) \propto \ p_3(\gamma)p_4(\gamma)$$

$$p(\gamma) \propto \ \gamma^{\alpha_3-1}e^{-\gamma(\delta_1+\beta_3)} \quad (16)$$

and the posterior distribution p($\gamma|X$) of the parameter γ is derived using (3) and (16) as

$$p(\gamma|X) \propto \ \gamma^{n+\alpha_3-1}e^{-\gamma(\delta_1+\beta_3+\sum x)}$$

which is the density function of a Gamma distribution of

$$p(\gamma|X) = \frac{(\delta_1+\beta_3+\sum x)^{n+\alpha_3}}{\Gamma(n+\alpha_3)}\gamma^{n+\alpha_3-1}e^{-\gamma(\delta_1+\beta_3+\sum x)} \quad (17)$$

With parameters $(n+\alpha_3, \delta_1+\beta_3+\sum x)$

$$Mean = \ E(X) = \frac{\delta_1+\beta_3+\sum x}{n+\alpha_3}$$

$$Var(X) = \frac{\delta_1 + \beta_3 + \sum x}{(n + \alpha_3)^2}$$

F. Posterior Distribution of Unknown Parameter γ Using Chi-Square-Exponential Prior

The double prior for γ is defined to be a Gamma distribution with hyper parameters ψ_2 and an exponential distribution with hyper parameter δ_2 as

$$p_5(\gamma) = \frac{1}{2^{\frac{\psi_2}{2}}\Gamma\left(\frac{\psi_2}{2}\right)}\gamma^{\frac{\psi_2}{2}-1}e^{-\frac{\gamma}{2}}; \psi_2 > 0, \gamma > 0 \qquad (18)$$

$$p_6(\gamma) = \delta_2 e^{\delta_2 \gamma}; \; \delta_2 > 0 \qquad (19)$$

A double prior is defined for γ by combining the two priors in (18) and (19) as:

$$p(\gamma) \propto \gamma^{\frac{\psi_2}{2}-1}e^{-\gamma\left(\frac{1}{2}+\delta_2\right)} \qquad (20)$$

and the posterior distribution p($\gamma|X$) of the parameter γ is derived using (3) and (20) as

$$p(\gamma|X) \propto \gamma^{\left(n+\frac{\psi_2}{2}-1\right)}e^{-\gamma\left(\delta_2+\sum x+\frac{1}{2}\right)}$$

which is the density function of a Gamma distribution of

$$p(\gamma|X) = \frac{\left(\delta_2+\sum x+\frac{1}{2}\right)^{n+\frac{\psi_2}{2}}}{\Gamma\left(n+\frac{\psi_2}{2}\right)}\gamma^{\left(\alpha_2+\frac{\psi_1}{2}+n-1\right)}e^{-\gamma\left(\delta_2+\sum x+\frac{1}{2}\right)} \qquad (21)$$

With parameters $\left(n+\frac{\psi_2}{2}, \delta_2+\sum x+\frac{1}{2}\right)$

$$Mean = E(X) = \frac{\delta_2 + \sum x + \frac{1}{2}}{n + \frac{\psi_2}{2}}$$

$$Var(X) = \frac{\delta_2 + \sum x + \frac{1}{2}}{\left(n + \frac{\psi_2}{2}\right)^2}$$

G. Simulation Design

A simulation study was carried out in order to be able to estimate the parameter on interest. Using R programming language was used to simulate a sample size of 800 from an exponential distribution with a mean 5.

III. Results And Discussion

A. Results

The Exponential distribution provides a realistic model for many random phenomena. Because the values of an Exponential random variable are non-negative integers, any random phenomena for which a count is of interest is a candidate for modelling by assuming an Exponential distribution. The numerical and graphical illustration of posterior densities of the parameters of interest conveys a convincing and comprehensive picture of Bayesian data analysis. Several programs were developed to calculate posterior densities of the Exponential distribution under various priors in R Software and OpenBUGS for this study using the method described in [5] and [6]. These programs illustrate the strength of Bayesian methods in practical situations.

In order to estimate the unknown parameter γ a simulation of 800 samples were generated using R programming with mean 5. The QQ plot, PP Plot and empirical distribution function, the fitted distribution function for both probability density function and the cumulative probability density function are shown in Fig 1. The plots show that the dataset simulated fit in perfectly to the distribution in which they are simulated from (exponential distribution).

The posterior mean and variance under all the assumed priors is calculated assuming the values of all hyper parameters are 2 and are displayed in Table 1.

Bayesian Analysis

We ran the model to generate two Markov Chains at the length of 20,000 with different starting points of the parameters. The convergence is monitored using trace and ergodic mean plots, we find that the Markov Chain converge
together after approximately 2000 observations. Therefore, burnin of 4000 samples is more than enough to erase the effect of starting points (initial values). Finally, samples of size 8000 are formed from the posterior by picking up equally spaced every two outcome, i.e. thin = 2, starting from 4001. Therefore, we have the posterior sample $\{\gamma_{1i}\}$, i = 1,…, 8000 from chain 1 and $\{\gamma_{2i}\}$, i = 1,…,8000 from chain 2.

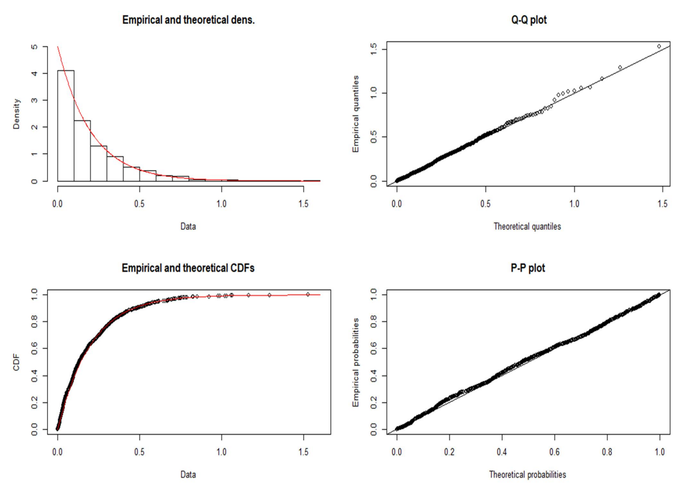

Fig 1: The graph of empirical distribution and fitted distribution function, QQ plots, PP plots.

Table 1: Posterior Mean and Posterior Variance of an Exponential Distribution with Different Priors for Chain 1

Prior	Chain 1		
	Mean	Variance	MC_error
Uniform	4.98	0.1783	0.001599
Jeffrey's	4.976	0.1771	0.001994
Gamma	4.93	0.1735	0.001349
Gamma Chi-square	4.974	0.1766	0.001214
Gamma Exponential	4.96	0.1765	0.001159
Chi-Square Exponential	4.97	0.1773	0.001758

Table 2: Posterior Mean and Posterior Variance of an Exponential Distribution with Different Priors for Chain 2

Prior	Chain 2		
	Mean	Variance	MC_error
Uniform	4.99	0.1785	0.001600
Jeffrey's	4.977	0.1755	0.002774
Gamma	4.927	0.1725	0.001949
Gamma Chi-square	4.972	0.1759	0.001801
Gamma Exponential	4.958	0.1756	0.001732
Chi-Square Exponential	4.969	0.1776	0.002437

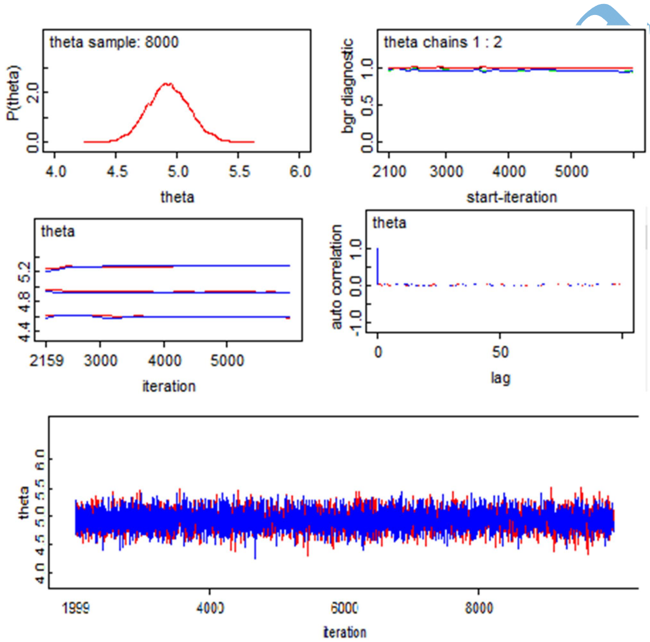

Fig. 2: Convergence Plots for parameter γ

The chain 1 is considered for convergence diagnostics plots. The visual summary is based on posterior sample obtained from chain 2 whereas the numerical summary is presented for both the chains.

B. Discussion

From Fig 2, the history plot shows that there is ample evidence of convergence of chain as the plots show no long upward or downward trends, but look like a horizontal band, then we has evidence that the chain has converged. Also, the BGR diagnostic plot clear show that convergence is achieved and the autocorrelation graph shows that the correlation is almost negligible. According to [2] if both quantiles are approximately 1.0, effective convergence may be diagnosed. We may conclude that the samples are independent. Thus we can obtain the posterior summary statistics.

In Table 1 and 2, we have considered various quantities of interest (posterior means, posterior variances and mc_error) and their numerical values based on MCMC sample of posterior characteristics for Exponential model with different prior distributions.

It was observed that the posterior mean and posterior variance under Gamma distribution is less compared to the posterior mean and posterior variance for other assumed prior. Also, the MC_error for posterior mean and variance for Gamma distribution is less compared to the MC_error for other assumed prior distributions.

IV. Conclusion

In this research work, the researcher considered different priors distribution in estimating the rate of an exponential distribution. One of the most important motivations for using a Bayesian approach instead of a frequentist one is the fact that the Bayesian estimation of uncertainty (like variances and confidence intervals) is not based on asymptotic sampling arguments that requires the availability of larges samples.

Based on the results obtained, it was observed that the results obtained are very close to the true value of $\gamma = 5$ of an exponential distribution. We concluded that Gamma prior distribution is more efficient compared to other priors and the lower variation in posterior distribution assists in more precise Bayesian estimates of the true unknown parameter γ of an Exponential distribution.

Acknowledgment

The authors are grateful to anonymous reviewers for their valuable comments on the original draft of this manuscript

References

[1] Berger, J. O. (1985). Statistical decision theory and Bayesian analysis, 2nd Edition. New York, NY: Springer-Verlag.

[2] Gelman, A., Rubin, D. B. (1992), Inference from iterative simulation using multiple sequences (with discussion). Stat. Sci., Hayward, v.7, p.457-511.

[3] Laplace, P. S. (1774, translated 1986). Memoire sur la probabilities des causes par les evenemens. Translation by S. Stigler. Memoir on the probability of causes of events. Statistical Science, 1, 364-378.

[4] Laplace, P. (1812). Theorie Analytique des Probabilite's. Paris: Courcier.

[5] Thomas, A. (2004). OpenBUGS, *Developer Manual* http://mathstat.helsinki.fi/openbugs/.

[6] Thomas, A. (2007). *OpenBUGS Developer Manual*, Version 3.0.2. URLhttp://mathstat.helsinki.fi/openbugs/