Research Journal of Recent Sciences ____ ISSN 2277-2502
Vol. **3(2),** 26-30, February **(2014)** *Res.J.Recent Sci.*

Effect of Mixture Prior in Case of Poorly Specified Prior

Muneeb Javed[1], Muhammad Saleem[2] and Alamgir[1]
[1]Department of Statistics, University of Peshawar, Peshawar, PAKISTAN
[2]Department of Statistics, Government Post Graduate College of Science, Faisalabad, PAKISTAN

Available online at: www.isca.in, www.isca.me
Received 30th April 2013, revised 22nd May 2013, accepted 10th July 2013

Abstract

The Shifted exponential distribution is appropriate for modeling the distribution of the time to failure of systems under constant failure rate condition. In this regard, the parameter is related to the mean life plus shifted parameter. In this research paper we present shifted exponential as likelihood function and conjugate inverted gamma prior for making Bayesian inference comparatively robust against a prior density poorly specified. Making use of a mixture of conjugate Square root inverted gamma priors assists us to make robust inference against misspecified prior. In case of having a very different likelihood than what will be expected for the given prior density, a large posterior probability of misspecification is obtained, and our posterior distribution will lean heavily on the likelihood.

Keywords: Shifted exponential distribution, joint posterior, mixture posterior, mixture of two components inverted gamma prior.

Introduction

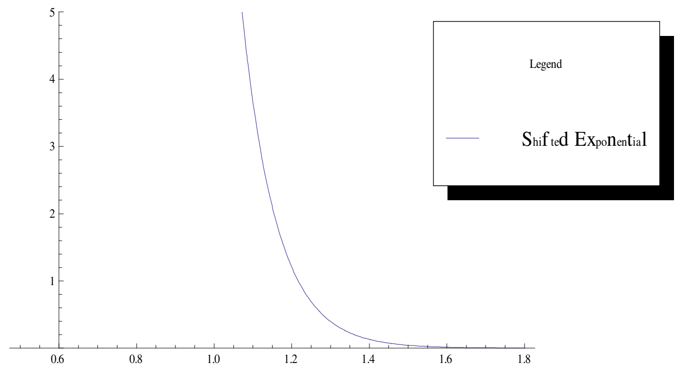
Shifted exponential distribution finds its application in failure data analysis and reliability analysis. The value of sifted parameter is some time taken as minimum life or guarantee time. When shifted parameter equals to zero, it simply becomes the exponential distribution. It was the first lifetime distribution for which researchers developed statistical methods. The exponential distribution is evaluated as a survival time model, particularly, used in industrial life testing[1,5]. Many researchers have added to the statistical methodology of the exponential model. Rich literature is available in this area[6,8].

In the current research paper, we consider the Shifted exponential distribution as sample information to update posterior distribution of the parameter (time-to-failure of the system) using the Inverted Gamma (IG) prior. Effect of misspecified prior is shown with the help of numerical example. Researchers have considered robust Bayesian inference by using the two components of mixture priors[9]. A mixture prior is also assumed that combine with likelihood to give mixture posterior distribution. The derivations of the mixture prior, the joint posterior and the mixture posterior is presented in order to eliminate the effect of misspecified prior and draw graphs to show that with the use of mixture prior, mixture posterior gives us better result in case of misspecified prior. A neighborhood class of mixture priors has been considered by researchers[10]. Some researchers have considered Bayesian analysis of the Rayleigh life time model incorporating Square root inverted gamma prior and mixture of two component Square root inverted gamma prior[11].

Methodology

Let $x_1, x_2, ..., x_n$ be a random sample drawn from Shift edexponential distribution with shift parameter $c = 1$ and scale parameter, that is $\lambda = 0.09$. The pdf and graph of the distribution are given below:

$$f(X \mid \lambda) = \frac{1}{\lambda} e^{-\frac{(x-c)}{\lambda}} \qquad x > c \qquad (1)$$



In pdf, λ is called a "survival parameter" as the random variable X is the amount of time that a given biological or mechanical system, R, manages to survive. Therandom variable $(\mathrm{X} - \mathrm{c}) \sim$ Exponential(λ) yields $E[X] = \lambda$. The joint pdf is called the likelihood function, which is a function of λ. Symbolically,

$$L(\mathbf{x}, \lambda) = f(x_1|\lambda)\ f(x_2|\lambda)\ \ldots\ f(x_n|\lambda)$$

Where $\mathbf{x} = (X_1, X_2, ..., X_n)$

The likelihood function is used to find the set of parameter values that gives the highest possible likelihood, but in Bayesian statistics, the purpose is to obtain a complete probability distribution over all possible parameter values, the likelihood function of the shifted exponential distribution is given as $\mathrm{L}(\mathbf{x},\lambda)=\frac{1}{\lambda^{n}}e^{-\frac{\Sigma(x-c)}{\lambda}}$

Bayesian Analysis of Shifted-Exponential Model Using IG Prior: Bayesian approach has several advantages over the more commonly used classical approach that it makes use of not only the sample information but also the prior information. Now here we take prior distribution for unknown parameter λ as Inverted Gamma and mixture of two component of Inverted Gamma distribution which reflects our beliefs about the population being studied. The inverted gamma is a two parameter probability distributions. The inverted gamma distribution's pdf is defined over the range $\lambda > 0$, with shape parameter a and scale parameter b.

$$P(\lambda)=\frac{b^{a}}{\Gamma(a)}\lambda^{-(a+1)}e^{-\frac{b}{\lambda}},\ \lambda>0 \tag{2}$$

The posterior information is proportional to the product of the prior information and the sample information.

$$P(\lambda\mid\mathbf{x})=\frac{(b+\Sigma x-nc)^{a+n}}{\Gamma(a+n)}\lambda^{-((a+n)+1)}e^{-\frac{b+\Sigma x-nc}{\lambda}},\ \lambda>0 \tag{3}$$

The above posterior is Inverted Gamma distribution. In the above density, the constants a and b are called the hyper-parameters.

Specification of Prior: Due to uncertainty about the true value of the parameter, it is considered as random variate. Hence the probability rules are used in order to draw inferences about the underlying parameters. The prior information is the probability statement about parameter which is interpreted as degree of belief or the relative weights that expert gives to every possible value of the parameter. These are the information the expert has before the data is observed. The prior should have relatively high probabilities over the whole range where the likelihood is considerably important.

But sometimes we have relatively high prior probabilities over the range that does not support the likelihood. In case of incorrect prior it affects posterior distribution[12].

Bayes' Theorem with Mixture Priors: Based on personal knowledge the researcher specifies the prior density for the unknown parameters of a statistical distribution. Let our prior distribution be $P_0(\lambda)$. Now it is quite precise that if our prior density is incorrectly specified, we would not have very good idea of what values λ should take. In that case we use uniform prior for λ that is $P_1(\lambda)$. Let $P_0(\lambda\mid\mathbf{x})$ be the posterior density of λ given the observations when we start with $P_0(\lambda)$ as the prior. Likewise, let $P_1(\lambda\mid\mathbf{x})$ be the posterior distribution of λ given the observations when we start with $P_1(\lambda)$ as the prior:

$$P_i(\lambda\mid\mathbf{x})\propto P_i(\lambda)\mathrm{L}(\mathbf{x},\lambda)\quad where\ i=0,1$$

By applying the updating rules we find them using conjugate prior family or flat prior.

The Mixture Prior: Let us define a new parameter, I ,taking two possible values. If $i=0$, then λ comes from the density $P_0(\lambda)$. On the other hand, if $i=1$, then λ comes from $P_1(\lambda)$. The conditional prior density of λ given I is given by

$$P_i(\lambda)=\begin{cases}P_0(\lambda) & if\quad i=0\\ P_1(\lambda) & if\quad i=1\end{cases} \tag{4}$$

Suppose the prior probability of I be $P(I=0)=p_0$, where p_0 is chosen some high value like .9,.95or .99, as we think of our prior $P_0(\lambda)$ to be correct. Let $p_1(I=0)=p_0$ be the prior probability that our prior density is misspecified. The joint prior density of λ and I is obtained as

$$P(\lambda,i)=p_i\times P_i(\lambda)\qquad\text{for}\qquad i=0,1. \tag{5}$$

It is worthwhile to be noted that the above joint density is continuous in the parameter λ and is discrete in the parameter I . One can easily obtain the marginal prior density of λ by marginalizing the joint densityover the values of I. It has mixture prior distribution as its distribution

$$P(\lambda)=\sum_{0}^{1}p_iP_i(\lambda) \tag{6}$$

is a mixture of the two prior densities.

The Joint Posterior: The joint posterior distribution of λ, I given the observation $\mathbf{x}$ is given by

$$P(\lambda,i\mid\mathbf{x})=a\times P(\lambda,i)\times f(\mathbf{x}\mid\lambda,i)\ \text{for}\ i=0,1 \tag{7}$$

Where a is the constant of proportionality. Since the sample data depends onlyon λ, not on I, therefore, the joint posterior reduces to

$$P(\lambda,i\mid\mathbf{x})=a\times p_iP_i(\lambda)f(\mathbf{x}\mid\lambda)\ \text{for}\quad i=0,1$$

$$=a\times p_ih_if(\lambda,\mathbf{x})\qquad\text{for}\qquad i=0,1 \tag{8}$$

where $h_i(\lambda,\mathbf{x}) = P_i(\lambda) f(\mathbf{x} \mid \lambda)$ is the joint distribution of the parameter and the data, when $P_i(\lambda)$ is the correct prior. To obtain the marginal posterior probability $P(I = i \mid \mathbf{x})$ we integrate λ out of the joint posterior:

$$\begin{aligned} P(I = i \mid \mathbf{x}) &= \int P(\lambda, i \mid \mathbf{x})\, d\lambda \\ &= a \times p_i \int h_i(\lambda,\mathbf{x})\, d\lambda \\ &= a \times p_i f_i(\mathbf{x}) \end{aligned}$$

for $i = 0,1$, where $f_i(\mathbf{x})$ is the marginal probability (or probability density) of the data. The posterior probabilities sum to 1 and the constant "*a*" is the normalizing constant, so

$$P(I = i \mid \mathbf{x}) = \frac{p_i f_i(\mathbf{x})}{\sum_{i=0}^{1} p_i f_i(\mathbf{x})} \quad (9)$$

These can be easily evaluated.

The Mixture Posterior: We find the marginal posterior of λ as

$$P(\lambda \mid \mathbf{x}) = \sum_{i=0}^{1} P(\lambda, i \mid \mathbf{x}). \quad (10)$$

Yet there is alternate way of arranging the joint posterior from conditional probabilities, that is,

$$P(\lambda, i \mid \mathbf{x}) = P(\lambda \mid i, \mathbf{x}) \times P(I = i \mid \mathbf{x}). \quad (11)$$

where $P(\lambda \mid i, \mathbf{x}) = P_i(\lambda \mid \mathbf{x})$ is the posterior distribution when we started with $P_i(\lambda)$ as the prior. Thus the marginal posterior of λ is

$$P(\lambda \mid \mathbf{x}) = \sum_{i=0}^{1} P_i(\lambda \mid \mathbf{x}) \times P(I = i \mid \mathbf{x}). \quad (12)$$

The expression in (12) is the mixture of the two posterior, where the weights are the posterior probabilities of the two values of*i* given the data.

Results and Discussion

Let we draw a random sample of size *n*=10, so our prior Inverted Gamma distribution is $G^{-1}(a,b)$. Here we take *a* = 8 and *b* = 2. From equation (3) our posterior distribution is $G^{-1}(18, 12.30)$.

From figure 1 we see that the conjugate inverted gamma prior and the likelihood are very far from each other and the posterior is in between. It gives high posterior probability to values that are not supported by the data (likelihood) which is shifted exponential distribution and are not strongly supported by prior either. This is not satisfactory. This shows how an incorrect prior can arise.

Now we reanalyze the data with a mixture prior. We let $P_0(\lambda)$ be the same $G^{-1}(8,2)$ prior that we used. We let $P_1(\lambda)$ be the uniform prior. We let the prior probability $p_0 = 0.95$.

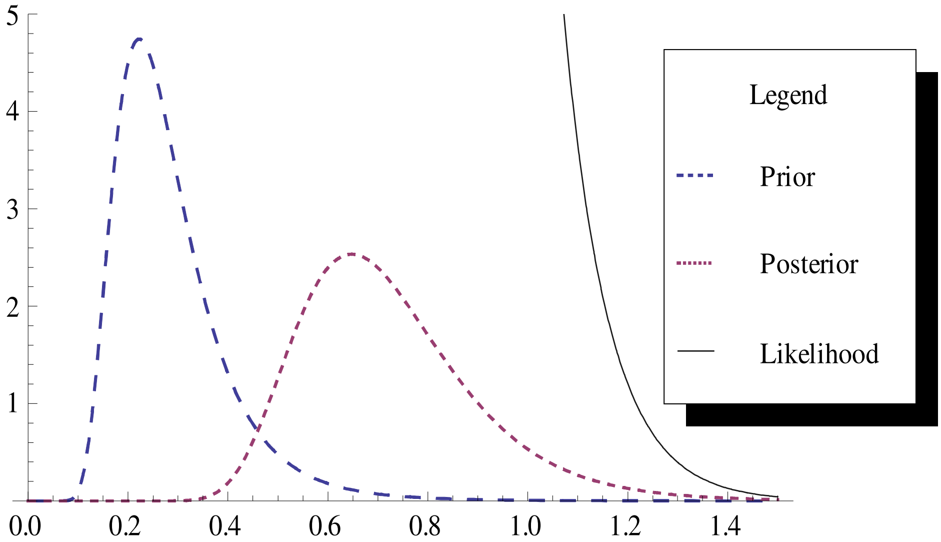


Figure-1
Likelihood, Prior and Posterior

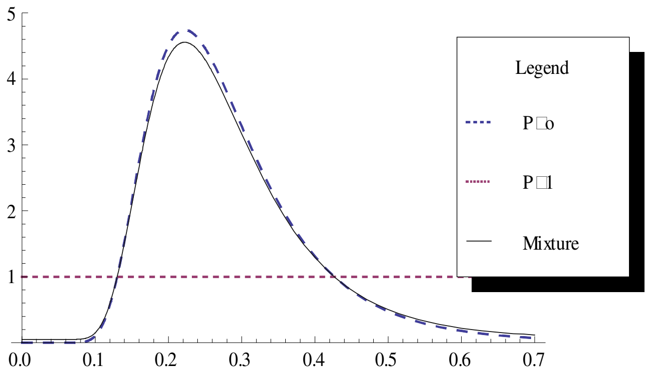


Figure-2
Mixture Prior and its Component

Mixture Prior is quite similar to single prior in figure 2. However, mixture prior has heavier weight in the tails. This gives makes our prior robust against prior specification. In this case, $h_i(\lambda, x)$ is:

$$h_0(\lambda,\mathbf{x}) = \frac{(2)^8}{\Gamma(8)}\lambda^{-(18+1)}Exp[-\tfrac{12.30}{\lambda}]$$

$$h_1(\lambda,\mathbf{x}) = \lambda^{-(9+1)}Exp[-\tfrac{10.30}{\lambda}]$$

Now integrating them with respect to λ gives

$$\int_0^\infty h_0(\lambda,\mathbf{x})d\lambda$$

$$= \frac{(2)^8}{\Gamma(8)}\int_0^\infty \lambda^{-(18+1)}Exp[-\tfrac{12.30}{\lambda}]d\lambda$$

$$= \frac{(2)^8}{\Gamma(8)}\frac{\Gamma(18)}{(12.30)^{18}}$$

We can evaluate the integral numerically

$$f_0(\mathbf{x}) = \int_0^\infty h_0(\lambda,\mathbf{x})d\lambda$$

$= 4.35097\times10^{-7}$ and

$$\int_0^\infty h_1(\lambda,\mathbf{x})d\lambda$$

$$= \int_0^\infty \lambda^{-(9+1)}Exp[-\tfrac{10.30}{\lambda}]d\lambda$$

$$= \frac{\Gamma(9)}{(10.30)^9}$$

To evaluate the integral numerically

$$f_1(\mathbf{x}) = \int_0^\infty h_1(\lambda, y_1,\ldots,y_n)d\lambda$$

$= 0.0000309019$.

So the posterior probabilities using the equation (9) are $P(I=0\mid\mathbf{x}) = 0.211057$ and $P(I=1\mid\mathbf{x}) = 0.788943$.

From equation (12) posterior distribution is given by
$P(\lambda\mid\mathbf{x}) = 0.211057\times P_0(\lambda\mid\mathbf{x}) + 0.788943\times P_1(\lambda\mid\mathbf{x})$,

Where $P_0(\lambda\mid\mathbf{x})$ and $P_1(\lambda\mid\mathbf{x})$ are the conjugate posterior densities obtained by using P_0 and P_1 as the respective priors.

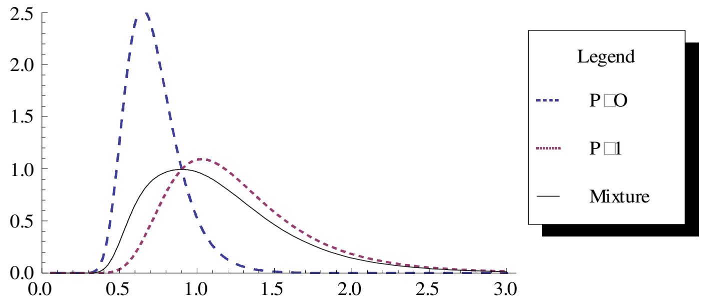


Figure-3
Mixture Posterior and its Two Components

Posterior distribution which is based on our prior experience denoted by P_O in figure 3 is very peaked if we compare with it by posterior distribution which is based on flat prior and mixture distribution. Mixture posterior distribution and posterior distribution based on uniform prior is quite similar.

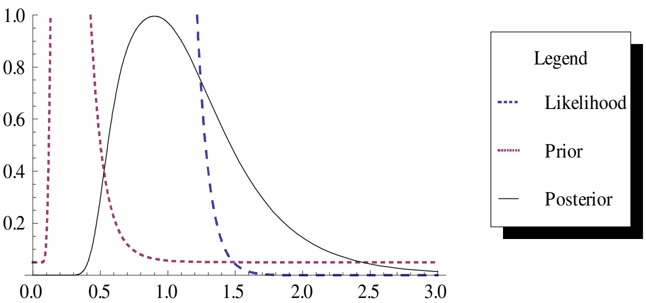


Figure-4
Mixture Prior, Likelihood and Mixture Posterior

When the prior and likelihood are very far from each other, we need to follow the likelihood function as it is determined from the observed data. Mixture prior and single prior look very similar to each other. However it has a heavier tail allowed by the mixture, and this has allowed its posterior to be very close to the likelihood as shown in figure 4. We see that this is much more satisfactory than the analysis that shown in figure 1.

Conclusion

In this study we presented a Bayesian analysis against a misspecified prior. By assuming the mixture of conjugate inverted gamma prior enabled us to do the robust analysis. We see that the in case of misspecified prior the posterior will be in between prior and likelihood, and will give high posterior probability to values neither supported by the likelihood or the prior. We give a small prior probability by using indicator random variable indicating that our original prior is misspecified.

The mixture posterior is very close to the original posterior when the original prior is correct. In case when the original prior is very far from the likelihood function, the posterior probability $P(I = 0 \mid \mathbf{x})$ is to be very small, and the mixture posterior will be close to the likelihood. This has set the conflict between the original prior and the likelihood by giving much more weight to the likelihood.

References

1. Epstein B. and Sobel M., Life testing, *J. Am. Statist. Ass,* **48**, 486-502 **(1953)**
2. Epstein B. and Sobel M., Some theorems relevant to life testing from an exponential distribution, *Ann. Math. Statist,* **25**, 373-381 **(1954)**
3. Epstein B. and Sobel M., Sequential life testing in the exponential case, *Ann.Math. Statist*, **26**, 82-93 **(1955)**
4. Epstein, B. Truncated life tests in the exponential case, *Ann. Math. Statist,* **25**, 555-564 **(1954)**
5. Epstein B., Testing for the validity of the assumption that the underlying distribution of life is exponential, *Technometrics*, **2,** 83-101 **(1960)**
6. Mendenhall, W. and Hader, R.J. Estimation of parameters of mixed exponentially distributed failure time distributions from censored life test data, *Biometrika,* **45**: No. ¾ **(1958)**
7. Johnson N.L., Kotz S. and Balakrishnan N., Continuous Univariate Distribution, Volume 1, Second Edition, Wiley-Interscience, New York, NY **(1994)**
8. Johnson N.L., Kotz S. and Balakrishnan N., Continuous Univariate Distribution. 2, Second Edition, Wiley-Interscience, New York, NY **(1995)**
9. Bolstad, W.M. Introduction to Bayesian Statistics, John Wiley & Sons Inc. New Jersey **(2004)**
10. Bose, S. Bayesian robustness with mixture classes of priors, *The Annals of Statistics*, **22(2),** 652-667 **(1994)**
11. Javed M. and Saleem M., A Bayesian Analysis of the Rayleigh Life Time Model, *World Applied Science Journal,* **16(2),** 220-226 **(2012)**
12. Javed M. and Saleem M. and Alamgir, Robust Bayesian Inference Using the Mixture Prior, *World Applied Science Journal*, **21(7)**, 1070-1074 **(2013)**