Bayesian Analysis of Exponentiated Pareto Distribution

Mohammed H AbuJarad[1*], Eman S AbuJarad[2], Athar Ali Khan[3]

[1,3]Department of Statistics and Operations Research, Aligarh Muslim University, Aligarh, 202002, India.
[2]Department of Mathematics, Aligarh Muslim University, Aligarh, 202002, India.
[*]Corresponding Author: m.jarad@gu.edu.ps

ABSTRACT

Modeling and analysis of Survival has been proved a fruitful aspect of statistical work in many fields of science. This paper aims at using Bayesian approach to fitting exponentiated Pareto distribution first introduced by Gupta et al in 1998 . A Bayesian approach has been adopted to fit this model as survival model. A real survival data set is used for illustration. Implementation is done using LaplaceApproximation, some graphical representations related to the probability density function and hazard function of the exponentiated Pareto are provided. LaplaceApproximation codes have been provided to implement censoring mechanism using both optimization and simulation tools.

Keywords: Exponentiated Pareto, LaplaceApproximation, Laplaces Demon, Posterior, Survival function, Hazard function, Simulation, Bayesian Inference, R.

1. Introduction

The Pareto distribution is one of the most famous models known after its founder the professor of economics: Vilfredo Pareto. one of its main objectives was basically proposed to model the unequal distribution of wealth Pareto noticed that a larger portion of the wealth of any society is in the hands of a very small percentage of people. moreover, the Pareto distribution has an important role not only in the field of economics, but also in analyzing a wide range of authentic and real-world situations. It was until Gupta et al. (1998) that a new distribution, called the Exponentiated Pareto distribution was introduced. This new distribution is of a main interest to the researcher who aims to employ it in applied operations using Bayesian analysis. Survival analysis has its many important applications such as in medicine, engineering, biological science, management, and public health. these statistical

distributions are basically employed to model the life of an item in order to study its important features. Therefore, suitable and proper distribution may give useful information that provides sound conclusions and help make right decisions. In this paper,the researcher aims at examining how Bayesian approach fits the Exponentiated Pareto distribution using LaplaceApproximation. The tools and techniques used in this paper are in implemented using LaplacesDemon package and thus they are in the Bayesian environment. Based on the Bayes rule, Bayesian inference can provide a rational method to update our beliefs in the light of new and upcoming information. Since researchers face difficulty to find solutions for high dimension, the researcher, in this paper, uses package to help alleviate the problem. The package is LaplacesDemon (Statisticat LLC 2015) which helps and facilitates high dimensional Bayeisan inference, mainly characterized by its own intellect and beneficial analysis. This package normally contains two functions: The function LaplaceApproximation presents the posterior results analytically and the function LaplacesDemon uses one of the several Metropolis algorithms Markov Chain Monte Carlo (MCMC) to simulate the results from the posterior density. A real survival data set is used to explain LaplaceApproximation .

Therefore, Bayesian analysis of Exponentiated Pareto distribution (EP) has been initiated to meet the following aims:

- To define a Bayesian model, that is, specification of likelihood and prior distribution.
- To write down the R code for approximating posterior densities with, LaplaceApproximation and simulation tools.
- To illustrate numeric as well as graphic summaries of the posterior densities.

2. Exponentiated Pareto Distribution

The Exponentiated Pareto distribution $EP(\theta, \lambda)$ have two-parameter, has been introduced recently by Gupta et al. (1998). We recall the probability density function (pdf), cumulative distribution function (cdf), survival function $S(t)$ and hazard function $h(t)$ Exponentiated Pareto distribution $EP(\theta, \lambda)$ are given by (1), (2), (3) and (4), respectively, as in Figure(1)

$$f(t; \theta, \lambda) = \theta \lambda [1 - (1+t)^{-\lambda}]^{\theta-1}(1+t)^{-(\lambda+1)}; \quad \lambda, \theta, t > 0. \qquad (1)$$

$$F(t; \theta, \lambda) = [1 - (1+t)^{-\lambda}]^{\theta}; \quad \lambda, \theta, t > 0. \qquad (2)$$

$$S(t; \theta, \lambda) = 1 - F(t) = 1 - [1 - (1+t)^{-\lambda}]^{\theta}; \quad \lambda, \theta, t > 0. \qquad (3)$$

$$h(t;\theta,\lambda) = \frac{f(t)}{S(t)} \quad \lambda,\theta,t > 0. \tag{4}$$

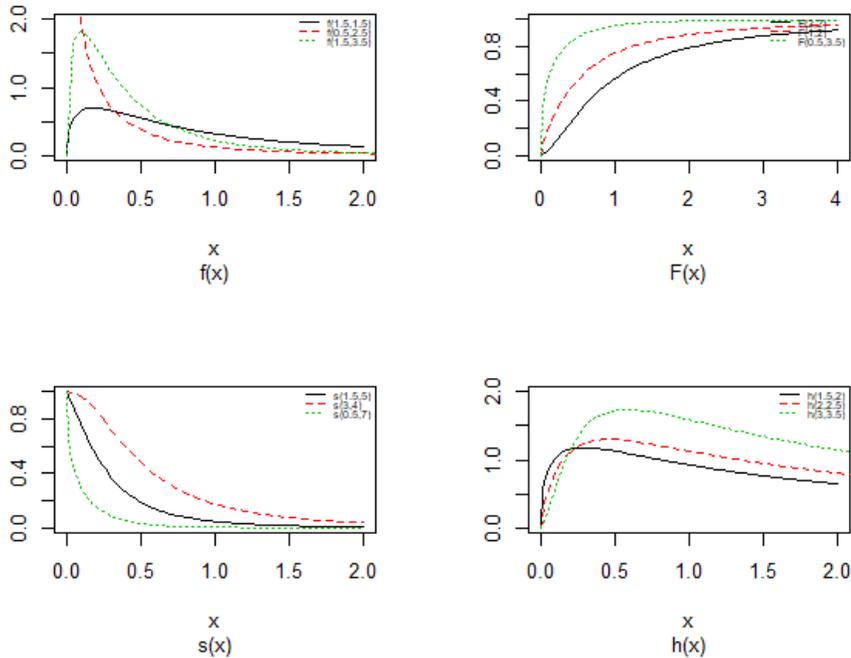Figure 1: Probability density plots, cdf, survival and hazard curves of Exponentiated Pareto distribution for different values of scale and shape

3. Functions for Exponentiated Pareto distribution in R

3.1 The Probability Density Function Using R

```
dEP<-function(x,shape,scale){
 a1<-(shape * scale)
a2<-(1-(1+x)^(-scale))^(shape-1)
 a3<-(1+x)^-(scale+1) a<- (a1*a2*a3)
 return(a)
 }
```

3.2 The Cumulative Distribution Function Using R

```
pEP<-function(x,shape,scale){
 p<-(1 - (1+x)^(-scale))^(shape)
 return(p)
 }
```

3.3 The Survival Function Using R

survEP<-function(x,shape,scale){
surv<-1-pEP(x,shape,scale)
return(surv)
}

3.4 The Hazard Function Using R

hexp<-function(x,shape,scale){
haz<-dexp(x,shape,scale)/survexp(x,shape,scale)
return(haz)
}

4. Bayesian Inference

Gelman et al.,(2013) split applied Bayesian modeling into the subsequent three steps:

1. put up a full probability model for all observable and unobservable quantities. This model should be consistent with obtainable knowledge of the data being modeled and how it was collected.
2. compute the posterior probability of unknown quantities conditioned on observed quantities. The unknowns may take in unobservable quantities such as parameters and potentially observable quantities such as predictions for outlook observations.
3. estimate the model fit to the data. This includes evaluating the implications of the posterior.

In general, this cycle will be repeated until a sufficient fit is achieved in the third step. Stan automates the calculations involved in the second and third steps (Carpenter et al., 2017).

We have to identify at this point the very important in Bayesian inference which are as for each the following :

- prior distribution: $p(\theta)$: The parameter θ can set a prior distribution elements that using probability as a means of quantifying uncertainty about θ before taking the data into acount.
- Likelihood $p(y \mid \theta)$: likelihood function for variables are related in full probability model.
- Posterior distribution $p(\theta \mid y)$: is the joint posterior distribution that expresses uncertainty about parameter θ after considering about the prior and the data, as in equation.

$$P(\theta \mid y) = p(y \mid \theta) \times p(\theta) \qquad\qquad (5)$$

5. The Prior Distributions

The Bayesian inference, having the prior distribution, can provide the information concerning an uncertain parameter θ connected through the probability distribution of data. This uncertain parameter is able to help obtain the posterior distribution $p(\theta \mid y)$. In the case of the Bayesian paradigm, it is very important for prior information to be identified through the value of the specified parameter. The information which are gathered before analyzing the experimental data with using a probability distribution function is referred to as the prior probability distribution (or the prior). In the remain of this paper, The researchers make use of two types of priors: half-Cauchy prior and Normal prior. The simplest types of priors is a conjugate prior which facilitates posterior calculations. In addition, a conjugate prior distribution is intended for an unknown parameter which leads to a posterior distribution for which there is a simple formula for posterior means and variances.

(Akhtar and Khan, 2014a) apply the half-Cauchy distribution by scale parameter α = 25 while a prior distribution for scale parameter.

Hereinafter we determination talk about the types of prior distribution:
- Half-Cauchy prior.
- Normal prior.

First, the probability density function of half-Cauchy distribution by scale parameter α is specified as a result

$$f(x) = \frac{2\alpha}{\pi(x^2 + \alpha^2)} \qquad x > 0, \alpha > 0.$$

Half-Cauchy distribution does not exist for mean and variance, although its mode is equal to 0. The half-Cauchy distribution by scale $\alpha = 25$ is a suggested, default, weakly informative prior distribution used for a scale parameter. On this scale $\alpha = 25$, the density of half-Cauchy is almost flat however not completely (see Figure 4), prior distributions that are not completely flat afford adequate information for the numerical approximation algorithm to continue to look at the target density; the posterior distribution. The inverse-gamma is often used as a non-informative prior distribution for scale parameter, but; this model creates a trouble for scale parameters close to zero; (Gelman and Hill, 2007) suggest that, the uniform, otherwise if more information is needed, the half-Cauchy is a better option. Consequently, in this paper, the half-Cauchy distribution with scale parameter

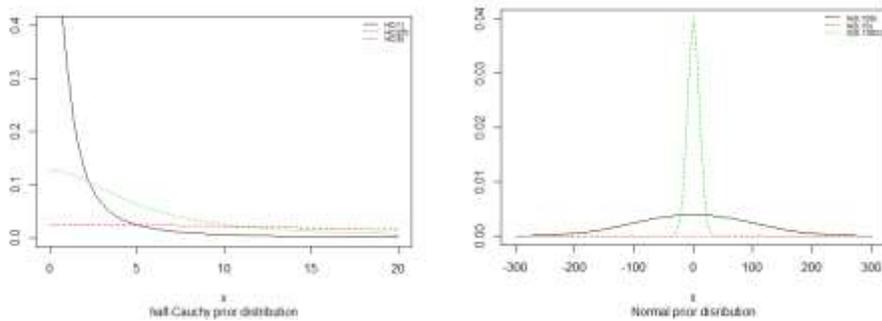$\alpha = 25$ is used as a weakly informative prior distribution.

Figure 2: half-Cauchy and normal distribution

Next, the normal (or Gaussian), every parameters is assigned a weak information Gaussian prior probability distribution. The researcher aims to use the parameters β_i independently in the normal distribution with mean=0 and standard deviation=1000, i.e., $\beta_j : N(0,1000)$, for this, we get a flat prior. As of (Figure 4), we see that the large variance indicates a lot of uncertainty about each parameter and hence, a weak informative distribution.

6. Laplace approximation

Laplace approximation is a technique first known in (Laplace, 1774(reprinted Stigler)).For this technique, it is of great importance to specify and determine information prior to data analysis. For example, Many simple Bayesian analyses give the similar results to standard non-Bayesian approaches when built on non informative prior distribution. for example, the posterior t-interval for the normal mean with unknown variance. An objective justification of a noninformative prior distribution depends on the amount of information available in the data; in the simple cases as the sample size n increases, the influence of the prior distribution on posterior inference decreases.These ideas, sometime referred to as asymptotic approximation theory because they refer to properties that hold in the limit as n becomes large. Thus, a special method of asymptotic approximation is the Laplace approximation which accurately approximates the unimodal posterior moments and marginal posterior densities in many cases. In this section we introduce a brief, informal description of Laplace approximation method. Suppose $-h(\theta)$ is a smooth, bounded unimodal function, with a maximum at $\hat{\theta}$, and θ is a scalar. By Laplace's method (e.g., Tierney and Kadane 1986), the integral

$$I = \int f(\theta)exp[-nh(\theta)]d\theta ,$$

can be approximated by

$$\hat{I} = f(\hat{\theta})\sqrt{\frac{2\pi}{n}}\sigma exp[-nh(\hat{\theta})]$$

while offered in (Mosteller andWallace, 1964), Laplace's method is to develop concerning $\hat{\theta}$ to find

$$I = \int f(\hat{\theta})exp(-n[h(\hat{\theta}) + (\theta - \hat{\theta})h'(\hat{\theta}) + \frac{(\theta - \hat{\theta})^2}{2}h''(\hat{\theta})])d\theta$$

Recall to facilitate that $h'(\hat{\theta}) = 0$, we include

$$I = \int f(\hat{\theta})exp(-n[h(\hat{\theta}) + \frac{(\theta - \hat{\theta})^2}{2}h''(\hat{\theta})])d\theta$$

$$= f(\hat{\theta})exp[-nh(\hat{\theta})]\int exp(\frac{-n(\theta - \hat{\theta})^2}{2\sigma^2})d\theta$$

$$= f(\hat{\theta})\sqrt{\frac{2\pi}{n}}\sigma exp[-nh(\hat{\theta})]$$

Spontaneously, if $exp[-nh(\theta)]$ is exceptionally topped about $\hat{\theta}$, then the integral be able to be well approximated by the performance of the integrand near θ. additional officially, it be able to be exposed that

$$I = \hat{I}[1 + O(\frac{1}{n})]$$

On the way to evaluate moments of posterior distributions, we have to estimate expressions for example:

$$E[g(\theta)] = \frac{\int g(\theta)exp[-nh(\theta)]d\theta}{\int exp[-nh(\theta)]d\theta},$$

where $exp[-nh(\theta)] = L(\theta | y)p(\theta)$ (e.g., Tanner 1996).

7. Bayesian Analysis of Exponentiated Pareto Model

To obtain the marginal posterior distribution of the particular parameters of interest, Bayesian analysis is the best method. Actually, the way to fulfill this aim is obvious; first, we calculate and obtain the joint posterior distribution of all unknown parameters, then, we integrate this distribution over the unknowns parameters that are

not of immediate interest to obtain the desired marginal distribution. Or in the same level, using simulation, so that the researchers draw samples from the joint posterior distribution. this will lead us to the ultimatum goal of looking at the parameters of interest and ignoring the values of the other unknown parameters.

7.1 The Model

Currently, the probability density function (pdf) is specified through this equation;

$$f(t;\theta,\lambda) = \theta\lambda[1-(1+t)^{-\lambda}]^{\theta-1}(1+t)^{-(\lambda+1)}.$$

Moreover, the survival function is set through;

$$S(t;\theta,\lambda) = 1 - F(t) = 1 - [1-(1+t)^{-\lambda}]^{\theta}.$$

We be able to condition the likelihood function for right censored (at the same time as is our case the data are right censored) as

$$L = \prod_{i=0}^{n} Pr(t_i,\delta_i)$$
$$= \prod_{i=0}^{n}[f(t_i)]^{\delta_i}[S(t_i)]^{1-\delta_i}$$

where δ_i is an indicator variable which takes value 0 if observation is censored and 1 if observation is uncensored. Consequently, the likelihood function is specified via

$$L = \prod_{i=0}^{n}[\theta\lambda(1-(1+t)^{-\lambda})^{\theta-1}(1+t)^{-(\lambda+1)}]^{\delta_i}[1-(1-(1+t)^{-\lambda})^{\theta}]^{1-\delta_i}.$$

As a result, the posterior distribution of our beliefs about the underlying bias is derived in the usual way by applying Bayes' rule (Statisticat LLC 2015), the joint posterior density is given by(AbuJarad and Khan 2018). Here the functions involve two parameters:

$$p(\beta\theta\,|\,t,X) \propto L(t\,|\,X,\beta\theta) \times p(\beta) \times p(\theta)$$

$$\propto \prod_{i=0}^{n}[\theta e^{X\beta}(1-(1+t)^{-e^{X\beta}})^{\theta-1}(1+t)^{-(e^{X\beta}+1)}]^{\delta_i}[1-(1-(1+t)^{-e^{X\beta}})^{\theta}]^{1-\delta_i}$$

$$\times \prod_{i=0}^{J}\frac{1}{\sqrt{2\pi\times10^3}}exp(-\frac{1}{2}\frac{\beta_j^2}{10^3})\times\frac{2\times25}{\pi(\theta^2+25^2)}..$$

To complete Bayesian inference in the Exponentiated Pareto model,we ought to decide an prior distribution for θ and $\beta's$. We talked about the issue related with determining prior distributions in segment 5, in any case, for straightforwardness now, we expect that the prior distribution for θ is half-Cauchy on the interval [0, 5]

and for β is Normal with [0, 5]. Basic application of Bayes rule as showed in (4.5), Applied in the direction of the Equation (14), after that gives the posterior density for b and β as Equation (15). consequence intended for this marginal posterior distribution get high-dimensional integral over all model parameters β_j. To solve this integral, we utilize the approximated via Markov Chain Monte Carlo methods. Notwithstanding, due to the availability of computer software package similar to LaplacesDemon, this compulsory model be capable of effortlessly be fitted in Bayesian paradigm via Laplaceapproximation in addition to MCMC techniques.

7.2   The Data: Breast Cancer

Breast cancer is one of the most common forms of cancer occurring in women living in the Western World. The data given in Table refers to the survival times (in months) of women who had received a simple or radical mastectomy to treat a tumour. The data is carried out at the Middlesex Hospital, and documented in Leathem and Brook (1987) and is also discussed by Collet (2003). In the table, the survival times of each woman are classied according to whether their tumour was positively or negatively stained. Censored survival times are labeled with an asterisk:

Negatively Stained: 23, 47, 69, 70*, 71*, 100*, 101*, 148, 181, 198*, 208*, 212*, 224*

Positively Stained: 5, 8, 10, 13, 18, 24, 26, 26, 31, 35, 40, 41, 48, 50, 59, 61, 68, 71, 76*, 105*, 107*, 109*, 113, 116*, 118, 143*, 154*, 162*, 188*, 212*, 217*, 225*

8.   Implementation Using with Laplace approximation

Bayesian fitting of Exponentiated Pareto(EP) model for this data can be done in R by using function LaplaceApproximation for analytic approximation and then with LaplacesDemon for MCMC simulations. Thus, implementation has been made by using LaplacesDemon package.

8.1 Creation of breast cancer data

Though most R functions use data in the form of a data frame, Laplace's Demon uses one or more numeric matrices in a list. It is much faster to process a numeric matrix than a data frame in iterative estimation. For the above data of 45 patients of prognosis of women withbreast cancer has given the survival times of women with tumour that were negatively or positively stained with HPA.

library(LaplacesDemon
*y<-c(23,47,69,70,71,100,101,148,181,198,208,212,224,5,8,10,13,18,24,26,
26,31,35,40,41,48,50,59,61,68,71,76,105,107,109,113,116,118,143,154,162,188,21
2,217,225)*

x1<-c(rep(0,13),rep(1,32))
censor<-c(rep(1,3),rep(0,4),rep(1,2),rep(0,4),rep(1,18),rep(0,4),1,0,1,1,rep
(0,6))
N<-length(y)
J<-2 #number of predictorsJ
X<-cbind(1,x1) #model matrix X
M1<-lm(log(y)~x1)
mon.names<-c("LP")
parm.names<-as.parm.names(list(beta=rep(0,J),alpha=0))
MyData<-list(J=J,X=X,mon.names=mon.names,censor=censor,parm.names=parm
.names ,y=y)

y is the vector of survival time containing both groups in it, $x1$ is the indicator vector (0 : negatively, 1 : positively), censor is a binary vector of censoring using 1 for uncensored and 0 for censored observation. The matrix X is created by the function cbind termed as model matrix. Its first column is of one whereas second column $x1$ is a column of indicator of staining. (0 : negatively stained, 1 : positively stained).

8.2 Model Specification

At the moment we motivation look at the posterior estimates of the parameters as soon as the Exponentiated Pareto Model is fitted to the above mentioned information (data). Therefore the meaning of the probability (likelihood) becomes the topmost necessity for the Bayesian fitting. at this juncture, we contain the likelihood like:

$$L(\theta \mid t) = \prod_{i=1}^{n} f(t_i)^{\delta_i} S(t_i)^{1-\delta_i}$$

$$= \prod_{i=1}^{n} \left(\frac{f(t_i)}{S(t_i)}^{\delta_i} S(t_i) \right)$$

$$= \prod_{i=1}^{n} h(t_i)^{\delta_i} S(t_i),$$

this approach, our log-likelihood progresses toward becoming

$$logL = \sum_{i=i}^{n} (log[h(t_i)]^{\delta_i} + log(S_t)).$$

$$t : EP(\theta, \lambda),$$

wherever $\lambda = exp(X\beta)$ is a linear combination of explanatory variables, log is the natural log used for the time to failure event. The Bayesian system requires the determination and specification of prior distributions for the parameters. At this point, we stick to subjectivity and thus introduce weakly informative priors for the parameters. Priors for the β, as well as α are in use to be normal and half-Cauchy at the same time as follows:

$$\beta_j : N(0,1000); \quad j = 1,2,3,...J$$

$$\theta : HC(0,25).$$

```
Model<-function(parm,Data){
 ### Parameters
 beta<-parm[1:Data$J]
 alpha<-exp(parm[Data$J+1])
 ### Log-Prior
beta.prior<-sum(dnorm(beta,0,1000,log=TRUE))
alpha.prior<-dhalfcauchy(alpha,25,log=TRUE)
### Log-Likelihood
mu<-exp(tcrossprod(Data$X,t(beta)))
lamda<-mu
lf<-function(y,alpha,lamda)(alpha*lamda)*(1-(1+y)^-lamda)^(alpha-1)*(1+y)^-(lamda+1)
ls<-function(y,alpha,lamda)1-(1-(1+y)^-lamda)^alpha
LL<-censor*log(lf(y,alpha,lamda))+(1-censor)*log(ls(y,alpha,lamda))
LL<-sum(LL)
### log-Posterior
LP<-LL+beta.prior+alpha.prior
Modelout<-list(LP=LP,Dev=-2*LL,Monitor=c(LP),yhat=mu,parm=parm)
return(Modelout)
 }
```

Into this approach, we obtain the log posterior of the Exponentiated Pareto Model.

8.3 Fitting of model with Laplace approximation

To facilitate approximate the joint posterior density, Laplace's method is utilized with the function LaplaceApproximation. For the idea of optimization, several algorithms have been implemented in this function. Among all the algorithms, we have discovered that the BFGS, LBFGS and TR perform well in a large portion of the cases. However, for this specific case Trust region TR algorithm of Nocedal and Wright (1999) is protected due its proficiency of convergence in the least number of

iteration. To start the optimization, some initial values for the parameters must be defined and in this way zero is set to the regression coefficients. currently, the time has come to call the function LaplaceApproximation to approximate the posterior densities of (EP) model for the data using Laplaceapproximation.This method is implemented in LaplaceApproximation function with TR as a method. An object Fit1 has been created as a result of using LapalceApproximation function.

Initial.Values<-c(coef(M1),log(1))

Fit1<-LaplaceApproximation(Model=Model, Initial.Values, Data=MyData, Iterations=500, Method="TR",Samples=1000)

Table1: Summary of the Trust region TR algorithm with, Mode, SD for standard deviation, LB, UB are 2.5%, 97.5% quantiles, respectively

	Mode	SD	LB	UB
beta[0]	-0.6520609	0.2035186	-1.05909803	-0.2450238
beta[1]	0.3043317	0.1413387	0.02165431	0.5870091
alpha	2.6330080	0.4084416	1.81612480	3.4498912

Table 2: Simulated results using sampling importance resampling (SIR) method, Mode, SD for standard deviation, ESS for number effective sample size, LB, Median, UB are 2.5%, 50%, 97.5% quantiles, respectively

	Mode	SD	MCSE	ESS	LB	Median	UB
beta[0]	-0.6950589	0.1918306	0.006066216	1000	-1.0634839	-0.690319	-0.333636
beta[1]	0.3083100	0.1496222	0.004731469	1000	0.02120982	0.3105953	0.6382411
alpha	2.5463563	0.3841378	0.012147504	1000	1.86579138	2.5196950	3.2848838
Deviance	311.2194476	2.1478647	0.067921446	1000	308.533006	310.77113	316.43061
LP	-175.2145248	1.0816780	0.034205661	1000	-177.978831	-174.94854	-173.90605

The second method is Nelder-Mead algorithm. This method is used and implemented in the LaplaceApproximation function with object name Fit2. The posterior summaries obtained by Nelder-Mead algorithm is accounted for Table 3. *Fit2<-LaplaceApproximation(Model=Model,Initial.Values,Data=MyData,Iterations=500, Method="NM",Samples=1000)*

Table 3: Summary of Nelder-Mead NM algorithm with, Mode, SD for standard deviation, LB, UB are 2.5%, 97.5% quantiles, respectively

	Mode	SD	LB	UB
beta[0]	-0.6523431	0.2036053	-1.05955368	-0.2451324
beta[1]	0.3046496	0.1413657	0.02191823	0.5873809
alpha	2.6329333	0.4084817	1.81596982	3.4498967

Table 4: simulated results using sampling importance resampling (SIR) method, Mode, SD for standard deviation, Monte Carlo standard error (MCSE), effective sample size ESS for number effective sample size, LB, Median, UB are 2.5%, 50%, 97.5% quantiles, respectively

	Mode	SD	MCSE	ESS	LB	Median	UB
beta[0]	-0.7239468	0.2096642	0.006630164	1000	-1.14870063	-0.7096739	-0.3414673
beta[1]	0.3097145	0.1497282	0.004734820	1000	0.01797051	0.3083062	0.6363872
alpha	2.5001636	0.4113588	0.013008308	1000	1.73490762	2.4749350	3.2912642
Deviance	311.530886	2.6254440	0.083023828	1000	308.47472870	310.7720762	318.37445
LP	-175.35663	1.3077761	0.041355510	1000	-178.69255055	-175.020489	-173.83753

8.4 Comparison of Optimization Techniques

Breast cancer data has been analyzed by two optimization methods that are, trust region and Nelder-Mead optimization algorithms. It is to be found from Table 1&2 and 3&4 that they are very close in terms of numerical posterior summaries. However, the calculations vary as far as convergence. Posterior density plots obtained by trust region and Nelder-Mead algorithms are reported in Figure 3 and 4 respectively.
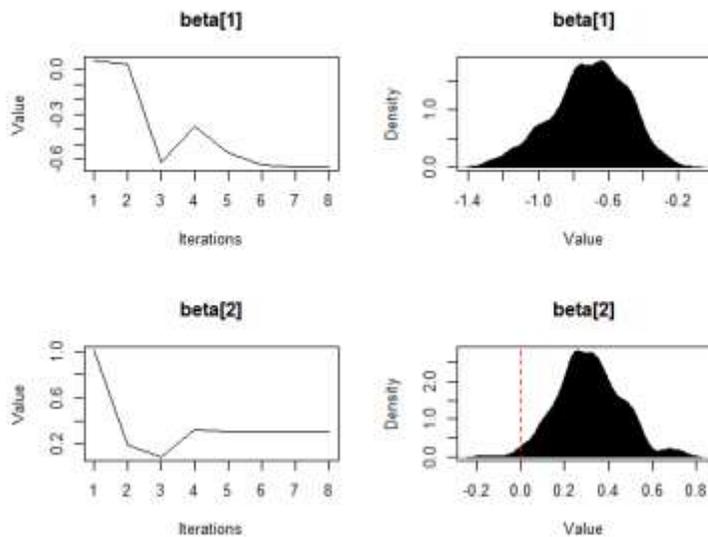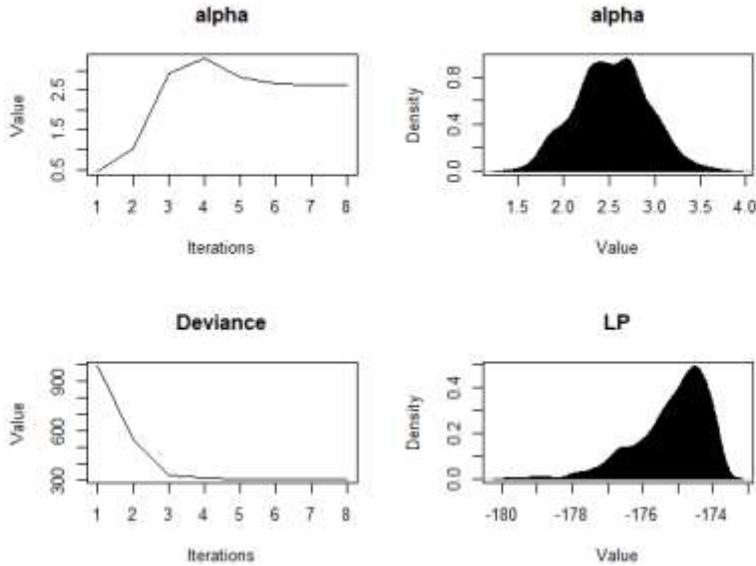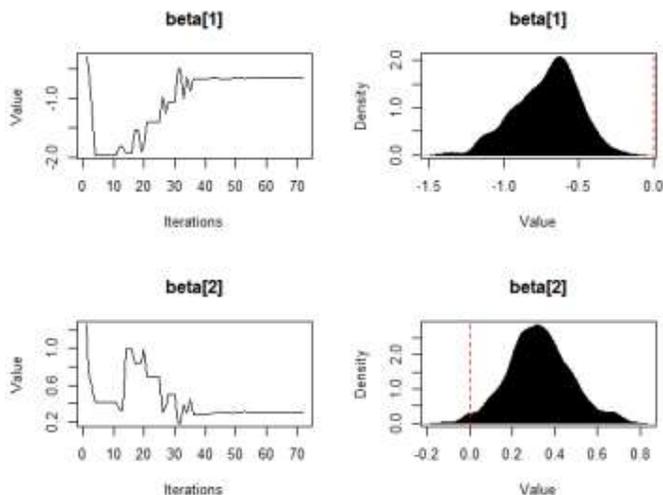
Figure 3: Posterior density plots of regressor variables obtained by TR

From Figure 3, we can see that the convergence of algorithm begin from 6th iteration. After this period the chain is balanced out inside a zone. subsequently the quantity of iterations we need to dispose of is 8 to screen the sampled values which exhibit much better conduct with little periodicities of course, from Figure 4 the calculation begins combines at 40th iteration. now, the iterations to be disposed of is around 40. The convergence speed of calculation by N-M is much slower than from trust region calculation. Be that as it may, N-M is a simplex base optimization algorithm and does not require any derivatives, while, trust district strategy requires derivatives of the goal work. Strikingly, in the execution of TR in LaplaceApproximation, supply of derivatives isn't required.
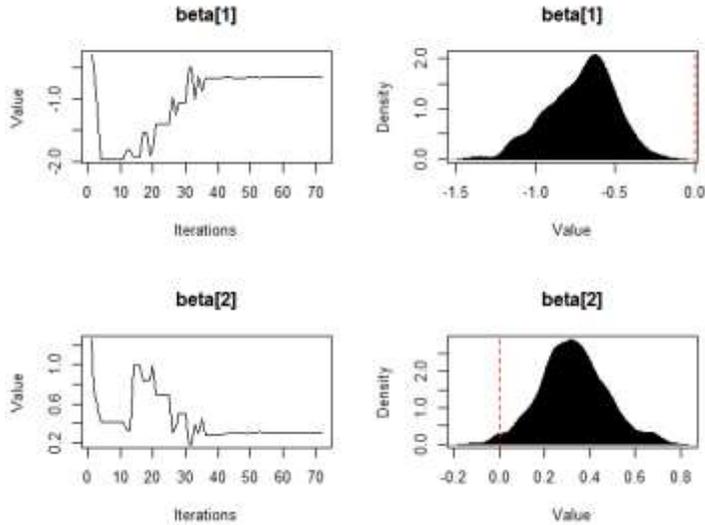
Figure 4: Posterior density plots of regressor variables obtained by NM

8.4.1 Simulation Study of Breast Cancer Data of EP model

In this area simulation will be performed by utilizing two algorithms in particular, random walk Metropolis algorithm and independent Metropolis algorithm. With the end goal of outline breast cancer data has been utilized. The R commands for the usage of RWM is given underneath with protest name Fit3 by utilizing function LaplacesDemon and the outcomes are outlined in Table 5 alongside the histograms of generated values, and their comparing portion evaluations of the posterior densities are portrayed in Figure 5.

  Initial.Value<-as.initial.values(Fit1)

  Fit3<-LaplacesDemon(Model,Data=MyData,Initial.Values,Covar=Fit3$C ovar,Iterations=50000,Status=100,Thinning=10,Algorithm="IM",Specs=list(mu=F it3$Summary1[1:length(Initial.Values),

Table 5: Simulated posterior summary obtained by independent Metropolis (IM) algorithm

	Mean	SD	MCSE	ESS	LB	Median	UB
beta[0]	-0.6605593	0.12250919	0.001757011	5000.000	-0.9068956	-0.6576426	-0.4316411
beta[1]	0.3057407	0.08579998	0.001239852	5000.000	0.1374309	0.3061102	0.4706322
alpha	2.6157104	0.24038677	0.003441005	5000.000	2.1300264	2.6193497	3.0889460
Deviance	309.3807653	0.86354245	0.013070103	4576.711	308.358467	309.1783927	311.614727
LP	-174.2963489	0.43318594	0.006524472	4726.311	-175.418631	-174.1921747	-173.79473

In Figure 5 there are three plots for every parameter, the deviance, and each observed variable (which in this illustration are LP and alpha). The furthest left plot is a trace-plot, demonstrating the historical backdrop of the estimation of the parameter as per the iteration. The middlemost plot is a density plot. The furthest right plot is an ACF or autocorrelation plot, demonstrating the autocorrelation at various lags. The chains look stationary (don't display a trend), and the ACF's show low autocorrelation. in this plot two things in these follow plots; stationarity and great mixing. Stationarity alludes to the way remaining inside the posterior distribution. We locate that all the mean estimation of the chain is very stable from start to finish. A great mixing chain implies that each progressive example inside every parameter isn't exceptionally correlated with the example before it.
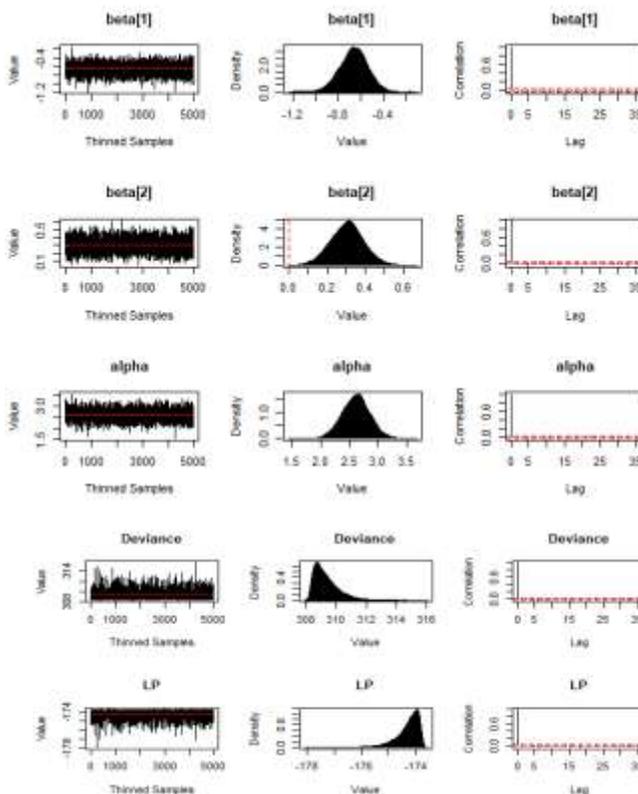

Figure 5: Posterior density plots of regressor variables obtained by IM

Presently the second simulation algorithm used to get simulated posterior summary is independent Metropolis algorithm. The contention of LaplacesDemon work for the usage of Random-walk Metropolis (RWM) algorithm with object name Fit3 is given as underneath:

Initial.Value<-as.initial.values(Fit1)

Fit4<-LaplacesDemon(Model,Data=MyData,Initial.Values,Covar=Fit3$C

ovar,Iterations=50000, Status=100, Thinning=1, Algorithm="RWM",Specs=NULL)

Table 6: Simulated posterior summary obtained by Random-walk Metropolis (RWM) algorithm

|	Mean	SD	MCSE	ESS	LB	Median	UB
beta[0]	-0.7241608	0.2259394	0.005405748	3438.7213	-1.22879524	-0.706542	-0.3343053
beta[1]	0.3172566	0.1558729	0.003654185	3684.6085	0.02610937	0.312076	0.6409900
alpha	2.4937777	0.4463295	0.012331563	2432.2910	1.65069149	2.500179	3.3100319
Deviance	311.947463	13.1560840	0.621644595	871.7042	308.53112331	310.662611	318.4655669
LP	-175.564152	6.5647157	0.310146791	872.6094	-178.75424383	-174.999932	-173.8785283
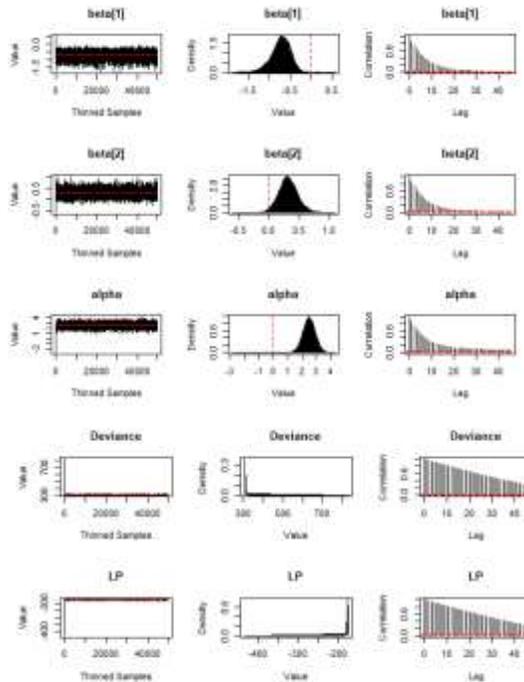
Figure 6: Posterior density plots of regressor variables obtained by RWM

8.5 Comparison of Simulation Techniques

The synopsis of the simulated posterior yield gotten by utilizing random walk Metropolis algorithm is accounted for in Table 6. This table comprises of seven columns which contains posterior mean and also posterior middle. Third and fourth columns of Table 6 is Monte Carlo standard error and effective sample size, individually. Column fifth and seventh speak to the 25% quantile meant by LB (Lower bound) and 97.5% quantile meant as UB (Upper bound), individually. Here,

it could seen that the value of MCSE is little, which demonstrates the convergence of algorithm. In addition, the reproduced posterior yield gotten by utilizing Independent Metropolis algorithm is accounted for in Table 5. which comprises same as the Table 5, yet value of MCSE is huge, which demonstrates the convergence of algorithm. In this two Tables(5 and 6) we see the essential lines which is Deviance. From Table 7 we can see that the Independent Metropolis (IM) is the best for simulation.

Table 7: Comparison of simulation techniques using Deviance and DIC

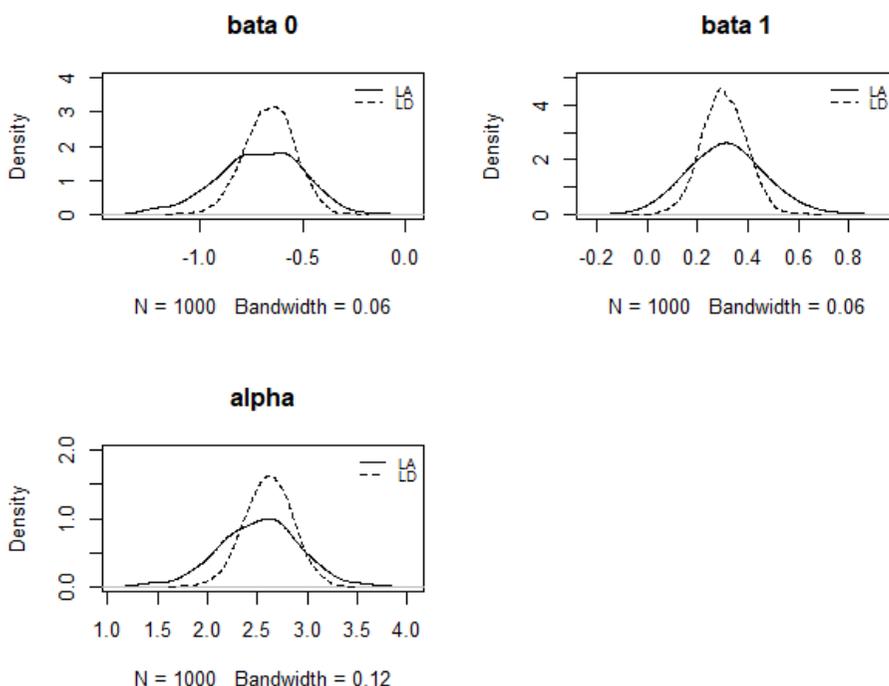|	Deviance	DIC
Independent Metropolis (IM)	309.38	309.735
Random-Walk (RWM)	311.94	398.48

Figure 7: Posterior density plots for $EP(\theta, \lambda)$ model, LA stands for LaplaceApproximation and LD for LaplacesDemon

9. Conclusion

In this paper, EP model is utilized to investigate the survival in Bayesian paradigm. A real survival data set is utilized for illustrative purposes. The analytic approximation and simulation methods are implemented using LaplacesDemon package of R. From the tables, the random walk Metropolis algorithm and independent Metropolis

algorithm both perform quite effectively for such a survival data under (EP) distribution. obviously simulation tools give better outcomes as we see in Figure 7, when contrasted with that acquired by asymptotic approximati

References

1. Akhtar, M. T. and Khan, A. A. (2014a). Bayesian analysis of generalized log-Burr family with R. *Springer Plus*, 3-185.

2. Collet, D. (2003). *Modelling Survival Data in Medical Research*, second edition. Chapman & Hall, London.

3. Gelman, A. and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, New York.

4. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. third edition. Chapman & Hall/CRC Press, London.

5. Gupta, R.C., Gupta, R.D. and Gupta, P.L. (1998). Modeling failure time data by Lehman alternatives. *Communications in Statistics: Theory and Methods*, 27, 887–904.

6. Laplace, P. (1774). Memoire sur la Probabilite des Causes par les Evenements. Academic Royale des Sciences, 621-656. English translation by Stigler, S.M. (1986). Memoir on the Probability of the causes of events in Statistical Science, 1, 364-378.

7. Leathem, A. and Brooks, S. (1987). Predictive Value of Lectin Binding on Breast-cancer Recurrence and Survival. *The Lancet*, 329(8541), 1054–1056.

8. Mohammed, H., Abu, J. and Khan, A.A. (2018). Exponential Model: A Bayesian Study With Stan. *Int J Recent Sci Res*, 9(8), pp. 28495-28506. DOI: 10.24327/ijrsr.2018.0908.2470.

9. Mosteller, F. and Wallace, D. (1964). *Inference and Disputed Authorship: The Federalist*. Reading: Addison- Wesley.

10. Nocedal, J., and Wright, S. J. (1999). *Numerical Optimization*. Springer-Verlag.

11. Statisticat, L.L.C. (2015). Laplaces Demon: Complete Environment for Bayesian Inference. R package version 3.4.0, http://www.bayesian-inference.com/software. URL http://www. bayesian-inference.com/software.

12. Tanner, M. A. (1996). *Tools for Statistical Inference*. Springer-Verlag, New York.

13. Tierney, L. and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, 81, 82-86.