

## Bayesian Survival Analysis of Generalized DUS Exponential Distribution

**Mohammed H AbuJarad**      **Athar Ali Khan**      **Eman S A AbuJarad**  
Aligarh Muslim University,      Aligarh Muslim University,      Aligarh Muslim University,  
Department of Statistics and      Department of Statistics and      Department of Mathematics  
Operations Research      Operations Research

---

### Abstract

Modeling and analysis of survival rate has proved a fruitful aspect of statistical work in many fields of science. This paper aims at using Bayesian approach to fit generalized DUS Exponential distribution (GDUSED). Kumar, Singh, and Singh (2015) proposed a renovation and called it DUS transformation. A Bayesian approach has been assumed to fit this model as a survival model. A real survival data set is utilized for illustration. Implementation is done using LaplaceApproximation and JAGS. Some graphical representations related to the probability density function and hazard function of the (GDUSED) are provided. LaplaceApproximation and JAGS codes have been provided to implement censoring mechanism using both optimization and simulation tools.

*Keywords:* GDUSED, LaplaceApproximation, LaplacesDemon, posterior, survival function, hazard function, simulation, Bayesian inference, JAGS, R.

---

### 1. Introduction

There is a variety of models, available in the literature, to examine the lifetime data. At first, exponential distribution was broadly used to investigate lifetime information because of its straightforwardness and systematic tractability. Albeit, a one-parameter exponential distribution, has a number of fascinating properties; however, its utilization is improper, all these things are taken into consideration in circumstances where related risk rate is not steady. To suit the circumstance of non steady danger rate, analysts tried relentlessly to grow new lifetime distribution with the aim that these turn out to be more adaptable than the ones used currently with regards to the state of their probability density function (*pdf*) and their related hazard rate. Kumar *et al.* (2015) proposed a change and called it a DUS change to get another distribution. In this case, the off chance that  $G(x)$  is the baseline cumulative distribution function (*cdf*), DUS transformation yields new *cdf*  $F(x)$  as given beneath:

$$F(x) = \frac{e^{G(x)} - 1}{e - 1}.$$

Maurya, Kaushik, Singh, and Singh (2017) proposed another class of distribution which comprises a wide range of hazard rates for fitting decision of shape parameter and proposed the

use of DUS transformation on the exponentiated *cdf*, henceforth referred to as generalized DUS (GDUS) transformation. This new distribution is of important interest to the researcher who aims to utilize it in applied operations by using Bayesian analysis. Survival analysis has numerous applications in various science fields, for example, in biological science, medicine, engineering, management, and public health. These statistical distributions are essentially utilized to show the life of an item with the end goal to think about its important features. In this way, suitable and proper distribution may yield valuable data that give good results and help settle on right choices. In this paper, the researcher goes for examining how Bayesian approach fits the (GDUSED) model by using `LaplaceApproximation` and `JAGS` (Just Another Gibbs Sampler). The tools and techniques which are used in this paper are implemented by using `LaplacesDemon` and `r2jags` packages of R, and hence they are in the Bayesian field. In light of the Bayes rule, Bayesian inference can give a balanced strategy to invigorate our belief in the light of new and upcoming data. Since researchers face difficulty in finding solutions for high dimension, in this article, the researcher uses a package to help in solving the problem, one package is `LaplacesDemon` ([Statisticat \(2015\)](#)) which helps and facilitates high dimensional Bayesian inference, mainly characterized by its own intellect and beneficial analysis and the other package is `R2jags` in which there is the function `JAGS` (Just Another Gibbs Sampler) can be run directly from R by using `R2jags` package. Also, `R2jags` package used for simulation from posterior density. The `JAGS` function takes data and starting values as an input. The `JAGS` function writes automatically a jags script, calls the model, and saves the simulations for easy access in R. A real survival data set is used to explain `LaplaceApproximation` and `JAGS`.

Therefore, Bayesian analysis of (GDUSED) model has been initiated to achieve the following aims:

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

## 2. The generalized DUS exponential distribution

[Maurya et al. \(2017\)](#) projected another class of distribution which comprises a wide range of hazard rates for proper choice of the shape parameter. Also, they proposed the use of DUS transformation on the exponentiated *cdf*, henceforth referred to as generalized DUS (GDUS) transformation. The distribution obtained by GDUS transformation is needed to have both monotone and bathtub-shaped hazard rates depending upon the decision of the values of the parameters. To illustrate the perspective, we consider exponential distribution as the base distribution because of its effortlessness and prominence in life testing problems, despite the fact that its use is limited to those phenomena where hazard rate is steady.

The new distribution through GDUS transformation introduced by [Maurya et al. \(2017\)](#) as follows :

$$f(t) = \frac{\alpha g(t) G^{\alpha-1}(t) \exp(G^\alpha(t))}{e - 1}. \quad (1)$$

$$F(t) = \frac{\exp(G^\alpha(t)) - 1}{e - 1}. \quad (2)$$

Now, we consider that the distribution in Equations (1&2) is an exponential distribution with parameter  $\lambda$ , also, we recall the probability density function (*pdf*), cumulative distribution function (*cdf*), survival function  $S(t)$  and hazard function  $h(t)$  generalized DUS Exponential

distribution  $GDUSED(\theta, \lambda)$  which can be in (3), (4), (5) and (6), respectively, as in Figure 1.

$$f(t; \alpha, \lambda) = \frac{\frac{\alpha}{\lambda} e^{-\frac{t}{\lambda}} (1 - e^{-\frac{t}{\lambda}})^{\alpha-1} \exp(1 - e^{-\frac{t}{\lambda}})^{\alpha}}{e - 1}; \quad \lambda, \alpha, t > 0. \quad (3)$$

$$F(t; \alpha, \lambda) = \frac{\exp(1 - e^{-\frac{t}{\lambda}})^{\alpha} - 1}{e - 1}; \quad \lambda, \alpha, t > 0. \quad (4)$$

$$S(t; \alpha, \lambda) = 1 - F(t) = 1 - \frac{\exp(1 - e^{-\frac{t}{\lambda}})^{\alpha} - 1}{e - 1}; \quad \lambda, \alpha, t > 0. \quad (5)$$

$$h(t; \alpha, \lambda) = \frac{f(t)}{S(t)} = \frac{\frac{\alpha}{\lambda} e^{-\frac{t}{\lambda}} (1 - e^{-\frac{t}{\lambda}})^{\alpha-1} \exp(1 - e^{-\frac{t}{\lambda}})^{\alpha}}{e - \exp(1 - e^{-\frac{t}{\lambda}})^{\alpha}} \quad \lambda, \alpha, t > 0. \quad (6)$$

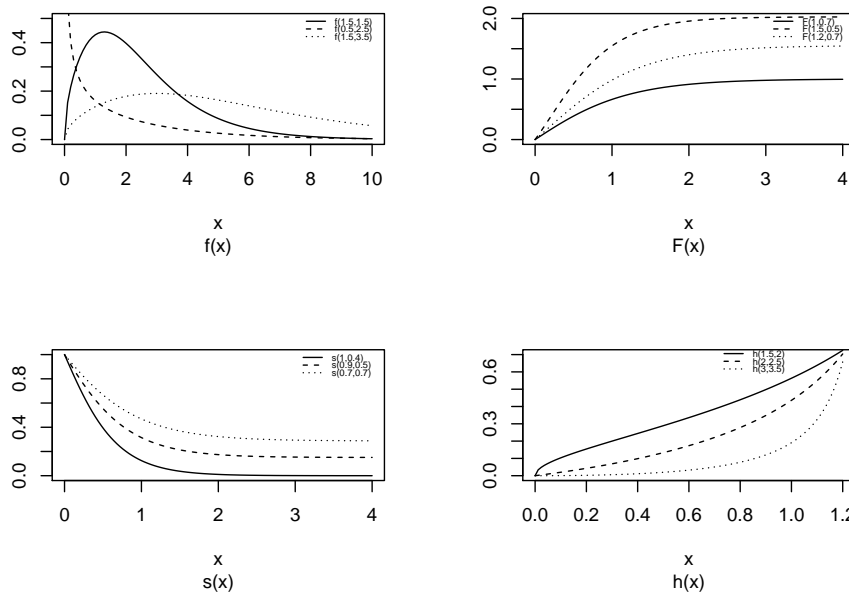


Figure 1: Probability density plots  $f(x)$ , cdf  $F(x)$ , survival  $s(x)$  and hazard curves  $h(x)$  of the Generalized DUS Exponential Distribution for different values of,  $\alpha$  and  $\lambda$ .

### 3. Bayesian inference

Gelman, Stern, Carlin, Dunson, Vehtari, and Rubin (2013) applied Bayesian modeling into the following three steps:

1. Establishing 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. Computing the posterior probability of unknown quantities conditioned on observed quantities. The unknown quantities may admit unobservable quantities such as parameters and potentially observable quantities such as predictions for outlook observations.
3. Estimating the model fit to the data. This includes evaluating the implications of the posterior.

At this point, we pinpoint the very important three points in Bayesian inferences which are as follows:

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

$$P(\theta|y) = p(y|\theta) \times p(\theta). \quad (7)$$

#### 4. The prior distributions

Having the prior distribution, the Bayesian inference can provide the information concerning an uncertain parameter  $\theta$  connected through the probability distribution of data. This uncertain parameter can help to obtain the posterior distribution  $p(\theta|y)$ . In the case of the Bayesian inference, it is very important for prior information to be identified through the value of the specified parameter. The information which is gathered before analyzing the experimental data by using a probability distribution function is referred to as the prior probability distribution (or the prior). In this article, we use two types of priors, the half-Cauchy prior and the Normal prior. The simplest type 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. (AbuJarad and Khan (2018a)) applied the half-Cauchy distribution by setting scale parameter  $\alpha = 25$ .

First, the probability density function of half-Cauchy distribution by scale parameter  $\alpha$  is defined as

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

The mean and variance of the Half-Cauchy distribution dose not exist but its mode is equal to 0. The half-Cauchy distribution with 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 but not completely see Figure 2. Thus, the prior distributions which are not completely flat provide sufficient information for the numerical approximation algorithm to continue to look at the target density (posterior distribution). The inverse-gamma distribution is often used as a non-informative prior distribution for scale parameter, but generates a trouble in the model fitting process (Gelman and Hill (2006)). Otherwise if more information is needed, the half-Cauchy is a favorable option. Consequently, in this article, the half-Cauchy distribution with scale parameter  $\alpha = 25$  is used as a weakly informative prior distribution.

Next, in the normal (or Gaussian), every parameters are assigned a weak information Gaussian prior probability distribution. Our aims in this work is to use the parameters  $\beta_i$  independently in the normal distribution with mean=0 and standard deviation=1000, i.e.,  $\beta_j \sim N(0, 1000)$ , therefore, we get a flat prior, as we see in Figure 2. We see that the large variance indicates a lot of uncertainty about each parameter and hence a weak informative distribution.

#### 5. Laplace approximation

Laplace Approximation is a technique first known in De Laplace (1774). For this technique, it is important to specify and decide information prior to data analysis. For example, many

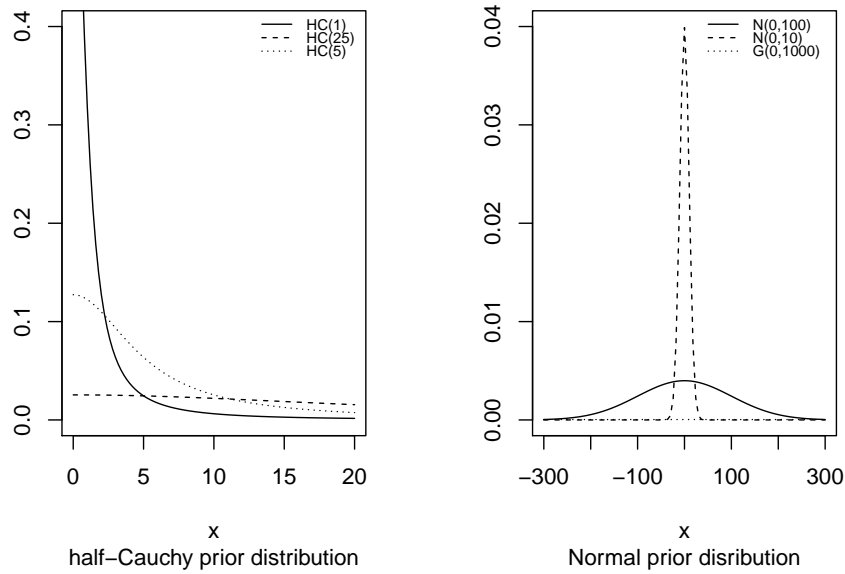


Figure 2: If we set the scale 25, then the half-Cauchy prior distribution will be a uniform distribution. Also, the Normal prior distribution will be flat with (0,1000).

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 validation of a non-informative prior distribution depends on the amount of information existing in the data, in the simple cases as the sample size  $n$  increases, the influence of the prior distribution on posterior inference decreases. These thoughts are sometimes 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 concise, informal description of Laplace Approximation method.

Suppose  $-h(\theta)$  is a smooth, bounded unimodal function with a maximum at  $\hat{\theta}$  and  $\theta$  is a scalar. By using Laplace Approximation (e.g., Tierney and Kadane (1986)) the following integral

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

can be approximated by

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

As presented in Mosteller and Wallace (1964), Laplace Approximation is used to expand  $\hat{\theta}$  and to find the following integral:

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

by setting  $h'(\hat{\theta}) = 0$ , we get

$$\begin{aligned} I &= \int f(\hat{\theta}) \exp\left(-n\left[h(\hat{\theta}) + \frac{(\theta - \hat{\theta})^2}{2}h''(\hat{\theta})\right]\right) d\theta \\ &= f(\hat{\theta}) \exp[-nh(\hat{\theta})] \int \exp\left(\frac{-n(\theta - \hat{\theta})^2}{2\sigma^2}\right) d\theta \end{aligned}$$

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

Spontaneously, if  $\exp[-nh(\theta)]$  is exceptionally topped about  $\hat{\theta}$ , then the integral is able to be well approximated by the performance of the integrand near  $\theta$ , which can be expressed in the following equation:

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

To evaluate moments of posterior distributions, we have to estimate expressions as follows (Tanner (2012)):

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

where  $\exp[-nh(\theta)] = L(\theta|y)p(\theta)$ .

## 6. Bayesian analysis of the generalized DUS exponential

Bayesian analysis is the best method to get the marginal posterior distribution of the specific parameters of interest. Basically, the best approach to meet this point is self-evident. First, we calculate and obtain the joint posterior distribution of all unknown parameters, at that point, we integrate this distribution over the unknown parameters that are not of immediate interest to obtain the needed marginal distribution. Also, at the same level, we draw samples from the joint posterior distribution by using simulation, this will lead to the ultimatum goal of treating the parameters of interest and ignoring the values of the other unknown parameters (see, e.g., AbuJarad and Khan (2018b)).

### 6.1. The generalized DUS exponential

The probability density function (*pdf*) of the generalized DUS exponential model is defined as:

$$f(t; \alpha, \lambda) = \frac{\frac{\alpha}{\lambda}e^{-\frac{t}{\lambda}}(1 - e^{-\frac{t}{\lambda}})^{\alpha-1}\exp(1 - e^{-\frac{t}{\lambda}})^{\alpha}}{e - 1}.$$

Also, the survival function  $S(t)$  of the generalized DUS exponential model is defined as:

$$S(t; \alpha, \lambda) = 1 - F(t) = 1 - \frac{\exp(1 - e^{-\frac{t}{\lambda}})^{\alpha} - 1}{e - 1}.$$

Now, we have the capacity to condition the likelihood function for right censored (in the meantime similar to our case the data are right censored) as follows:

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

where  $\delta_i$  is an indicator variable which takes the value 0 if observation is censored and the value 1 if observation is uncensored.

The likelihood function of the generalized DUS exponential model in terms of the  $f(t; \alpha, \lambda)$  and  $S(t; \alpha, \lambda)$ , is defined as:

$$L = \prod_{i=0}^n \left[ \frac{\frac{\alpha}{\lambda}e^{-\frac{t_i}{\lambda}}(1 - e^{-\frac{t_i}{\lambda}})^{\alpha-1}\exp(1 - e^{-\frac{t_i}{\lambda}})^{\alpha}}{e - 1} \right]^{\delta_i} \left[ 1 - \frac{\exp(1 - e^{-\frac{t_i}{\lambda}})^{\alpha} - 1}{e - 1} \right]^{1-\delta_i}. \quad (8)$$

The posterior distribution of our convictions about the basic inclination is determined in the standard way by applying Bayes' rule (Statisticat (2015)), the joint posterior density is given by AbuJarad and Khan (2018a) and defined as follows:

$$\begin{aligned}
 p(\beta, \alpha | t, X) &\propto L(t|X, \beta, \alpha) \times p(\beta) \times p(\alpha) \\
 &\propto \prod_{i=0}^n \left[ \frac{\frac{\alpha}{e^{\frac{t}{X\beta}}} e^{-\frac{t}{X\beta}} (1 - e^{-\frac{t}{X\beta}})^{\alpha-1} \exp(1 - e^{-\frac{t}{X\beta}})^{\alpha}}{e-1} \right]^{\delta_i} \left[ 1 - \frac{\exp(1 - e^{-\frac{t}{X\beta}})^{\alpha} - 1}{e-1} \right]^{1-\delta_i} \\
 &\times \prod_{i=0}^J \frac{1}{\sqrt{2\pi} \times 10^3} \exp\left(-\frac{1}{2} \frac{\beta_j^2}{10^3}\right) \times \frac{2 \times 25}{\pi(\alpha^2 + 25^2)}. \tag{9}
 \end{aligned}$$

To complete Bayesian inference in the **generalized DUS Exponential model**, we choose a prior distribution for  $\alpha$  and  $\beta$ 's. Also, we discuss the issue related to deciding prior distributions in section 4. Now, we expect that the prior distribution for  $\alpha$  is half-Cauchy on the interval  $[0, 25]$  and for  $\beta$  is Normal with  $[0, 1000]$ . Basic application of Bayes rule as showed in equation (7) which is applied toward the condition (8) to gives the back thickness for  $\alpha$  and  $\beta$  as illustrated in condition (9). The result proposed for this negligible back appropriation gets high-dimensional fundamental over every single model parameter  $\beta_j$ . To solve this integral, we use the approximated by means of Markov Chain Monte Carlo methods. In any case, due to the availability of software package similar to `LaplacesDemon`, this model can be effectively fitted in Bayesian paradigm through `Laplaceapproximation` in addition to MCMC techniques as well as JAGS.

## 7. IUD data set

In an attempt to investigate and analyze menstrual bleeding data from women using contraceptives trials to prevent pregnancy, the World Health Organization (WHO 1987) has made available clinical data of 18 women aged between 18 and 35 years by using intrauterine device (IUD), known as the Multi-load 250 until discontinuation because of menstrual bleeding problems. The time origin starts with the first day in which women use the IUD, and it is finished with discontinuation because of bleeding problems. It was also mentioned that some women in the study stopped using the IUD because of the desire for pregnancy, or because of no need for a contraceptive, or just simply losing follow up. The study aimed at documenting those women for two years from the time origin. For some practical reasons those women could not be examined for two years to see if they were still using the IUD. This explains why there are three times of discontinuation of more than 104 weeks that are right-censored. The table below illustrates the number of weeks of using IUD (Collett (2015)).

Table 1: Times in weeks to discontinuation of the use of an IUD (Collett (2015)).

10	13*	18*	19	23*	30	36	38*	54*
56*	59	75	93	97	104*	107	107*	107*

## 8. Implementation using with Laplace approximation

Bayesian fitting of **generalized DUS Exponential model** for this data can be done in R by utilizing capacity `LaplaceApproximation` for analytic approximation and then with `LaplacesDemon` for MCMC simulations. In this manner, usage has been made by using `LaplacesDemon` package.



### 8.1. Creation of IUD data

Despite the fact that most R functions utilize data as a frame, Laplace's Demon utilizes at least one numeric matrices in a list. It is much faster to process a numeric matrix than to process a data frame in iterative estimation. The above data of 18 menstrual bleeding from women (IUD), have given the survival times. The data and related codes are given in Appendix A2.

### 8.2. Specification of the generalized DUS exponential

At that moment, we are motivated to look at the posterior estimates of the parameters as soon as the generalized DUS Exponential model (GDUSED) 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 point, the likelihood function can be written as:

$$\begin{aligned} L(\theta|t) &= \prod_{i=1}^n f(t_i)^{\delta_i} S(t_i)^{1-\delta_i} \\ &= \prod_{i=1}^n \left( \frac{f(t_i)^{\delta_i}}{S(t_i)} S(t_i) \right) \\ &= \prod_{i=1}^n h(t_i)^{\delta_i} S(t_i), \end{aligned}$$

by taking the logarithm of the above likelihood function, we obtain

$$\log L = \sum_{i=1}^n \left( \log \left[ h(t_i) \right]^{\delta_i} + \log(S_t) \right).$$

$$t \sim GDUSED(\alpha, \lambda),$$

where  $\lambda = \exp(X\beta)$  is a linear combination of explanatory variables and  $\log$  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 we introduce weakly informative priors for the parameters. Priors for the  $\beta$  and  $\alpha$  are in use to be normal and half-Cauchy at the same time as follows (see, e.g., [AbuJarad and Khan \(2018a\)](#)):

$$\begin{aligned} \beta_j &\sim N(0, 1000); \quad j = 1, 2, 3, \dots, J \\ \alpha &\sim HC(0, 25). \end{aligned}$$

In this methodology, we acquire the log posterior of the (GDUSED). The model specification and related codes are given in Appendix A3.

### 8.3. Fitting of generalized DUS exponential with Laplace approximation

Now, we use Laplace's method with the function `LaplaceApproximation` to facilitate approximating the joint posterior density. 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 to 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 (GDUSE) model for the data using `Laplaceapproximation`. This method is implemented in `LaplaceApproximation` function with `TR` as a method. An object `GDUSEDLA` has been created as a result of using `LaplaceApproximation` function. The fitting of model with Laplace Approximation and related codes are given in Appendix A4.



Table 2: Summary of the Trust region TR algorithm with `Mode` and `SD` for standard deviation. `LB`, `UB` are 2.5%, 97.5% quantiles, respectively.

	Mode	SD	LB	UB
beta	4.1697286	0.4006627	3.368403	4.971054
log.alpha	-0.6592323	0.3724867	-1.404206	0.085741

Table 3: Simulated results using sampling importance resampling (SIR) method, `Mode`, `SD` for standard deviation and `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	4.3236620	0.4132184	0.013067114	1000	3.605601	4.3064117	5.1240133
log.alpha	-0.8052070	0.3717694	0.011756380	1000	-1.512526	-0.7700861	-0.1521542
Deviance	100.6451801	1.5835276	0.050075538	1000	98.829493	100.2342456	104.6335808
LP	-61.8201652	0.7917613	0.025037690	1000	-63.814008	-61.6148299	-60.9124145
alpha	0.4778718	0.1734020	0.005483452	1000	0.220353	0.4629733	0.8588574

#### 8.4. Summarizing output

The analytic results by using `LaplaceApproximation` function are shown in Table 2. It may be noted that the posterior mode of parameters `beta` and `log.alpha` are 4.160.40,  $-0.650.37$ , respectively. According to 95% credible intervals, `beta` is found to be statistically significant and `log.alpha` is found to be statistically non-significant. Hence, they are appropriate variables for modeling survival data. The simulated results using sampling importance resampling (SIR) method are shown in Table 3, which represents the posterior mode (`Mode`), posterior standard deviation (`SD`), Monte Carlo standard error (`MCSC`), effective sample size (`ESS`) and respective credible intervals `LB` (2.5%), `Median` (50%) and `UB` (97.5%).

#### 8.5. Simulation study of IUD data of GDUSED

In this section, the simulation will be performed by utilizing random walk Metropolis algorithm. With the end goal of outline IUD data has been utilized. The R commands for the usage of IM are given underneath with protest name `GDUSEDLD` by utilizing function `LaplacesDemon` and the outcomes are outlined in Table 4, alongside the Posterior density plot is presented in Figure 3. The simulation study of IUD data of GDUSED and related codes is given in Appendix A5.

Table 4: Simulated posterior summary obtained by independent Metropolis (IM) algorithm.

	Mean	SD	MCSE	ESS	LB	Median	UB
beta	4.1859020	0.2344156	0.003308514	4765.90	3.7279120	4.1841940	4.6573480
log.alpha	-0.6738969	0.2203885	0.003108755	5000.00	-1.1109668	-0.6721282	-0.2402241
Deviance	99.4684210	0.7273085	0.010679895	4946.49	98.7853611	99.2393235	101.4355342
LP	-61.2318293	0.3637238	0.005340583	4946.98	-62.2161453	-61.1173321	-60.8902876
alpha	0.5221997	0.1158907	0.001643958	5000.00	0.3292405	0.5106207	0.7864516

## 9. Fitting Bayesian generalized DUS exponential in JAGS

JAGS (Martyn (2011)) is Just Another Gibbs Sampler that was mainly written by Martyn Plummer in order to provide a BUGS. It is a program for the analysis of Bayesian models using Markov Chain Monte Carlo (MCMC). `R2jags` (Su and Yajima (2012)) is an R package

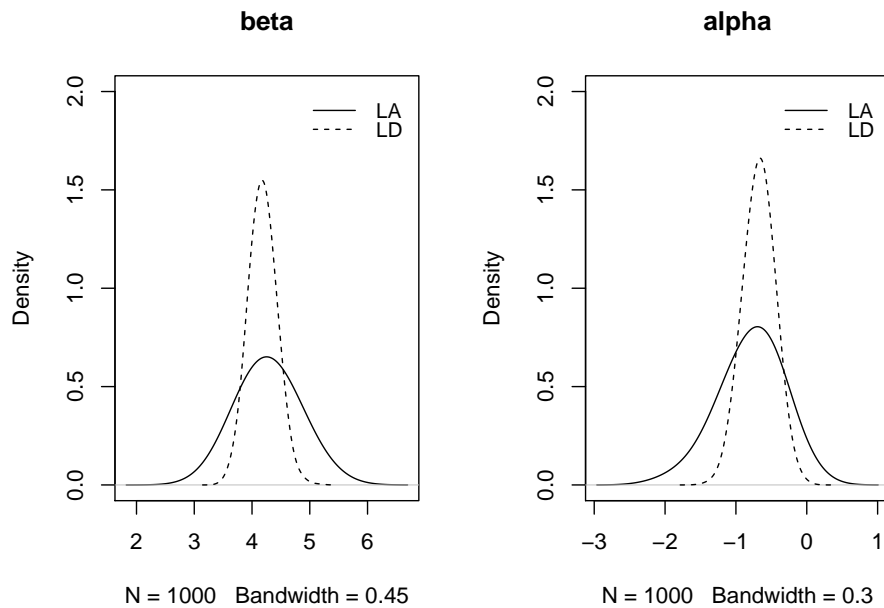


Figure 3: Posterior density plots for  $GDUSED(\beta, \alpha)$  model, LA stands for LaplaceApproximation and LD for LaplacesDemon.

that allows fitting JAGS models from within R (Gelman *et al.* (2013)). This program is based on a version of numerical library (Rmath) used for R. Many of the functions in base R for mathematical and statistical calculations are also available in the JAGS (Lunn, Jackson, Best, Spiegelhalter, and Thomas (2012)). Let us consider the Bayesian analysis of IUD data with JAGS using its interface of R that is, R2jags package of R. R2jags which is designed for inference on Bayesian models using Markov chain Monte Carlo (MCMC) simulation, which is also used for simulation from the posterior density. The JAGS function takes data and starting values as input. It automatically writes a jags script, calls the model, and saves the simulations for easy access in R.

### 9.1. Creation of data

In order to fit the model with JAGS, the data is required to be listed containing the name of each vector. This can be done by R related codes that are given in Appendix A6.

### 9.2. Definition of the generalized DUS exponential

The GDUSED is used for modeling the IUD data, as presented in the following formula:

$$t_i \sim GDUSED(\alpha, \lambda)$$

with log-link function

$$\log \lambda = X\beta,$$

where  $X$  is model matrix and  $\beta$  is the vector of regression coefficients.

Moreover, the prior probabilities for the parameters  $\beta$  and  $\alpha$  are specified, respectively, as follows:

$$\begin{aligned} \beta &\sim N(0, 0.001) \\ \alpha &\sim U(0, 100). \end{aligned}$$

Thus, the JAGS codes of the this model are given in Appendix A7.

### 9.3. Summarizing output

The summary of JAGS simulations after being fitted to the  $GDUSED(\alpha, \lambda)$  model for the IUD data JAGS simulates the data from posterior density using Metropolis within Gibbs algorithm and approximates the results which are reported in Table 5. Rhats are very close to 1.0, which is indicated to a good convergence. A plot of the posterior densities can be seen in Figure 4.

Table 5: Posterior mean, posterior SD, quantiles, Rhat and effective sample size (n.eff)

	Mean	SD	2.5%	50%	97.5%	Rhat	n.eff
beta	4.353	0.432	3.632	4.304	5.372	1.003	2500
alpha	0.483	0.154	0.213	0.474	0.797	1.001	1800
deviance	100.001	2.054	98.193	100.101	104.381	1.000	1

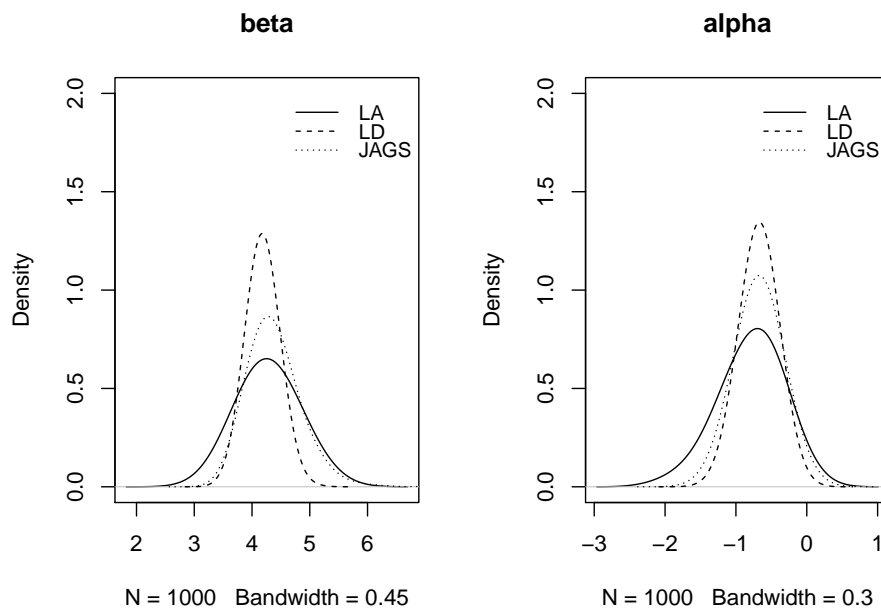


Figure 4: Posterior density plots for  $GDUSED(\beta, \alpha)$  model, LA stands for LaplaceApproximation, LD for LaplacesDemon and JAGS.

## 10. Conclusion

In this paper, GDUSED has been used for the analysis of IUD data in Bayesian paradigm. For illustration of data, a real survival data set is used. The analytic approximation and simulation methods are implemented using LaplacesDemon and R2jags packages of R. As indicated by the findings, it is clear that simulation tools provide better results when compared with those obtained by asymptotic approximation.

## Appendix

### A1

Functions Generalized DUS Exponential Distribution in R  
The Probability Density Function Using R

```
dGDUSED<-function(x,shape,scale){
a1<-(shape/scale)
a2<-(exp(-(x/scale)))
a3<-(1-exp(-(t/scale)))^(shape-1)
a4<-exp(1-exp(-(t/scale)))^(shape)
a5<-exp(1)-1
a<- ((a1*a2*a3*a4)/a5)
return(a)
}
```

The Cumulative Distribution Function Using R

```
pGDUSED<-function(x,shape,scale){
p<- exp(1-exp(-(t/scale)))^(shape) -1
return(p)
}
```

The Survival Function Using R

```
survGDUSED<-function(x,shape,scale){
surv<-1-pGDUSED(x,shape,scale)
return(surv)
}
```

The Hazard Function Using R

```
hexp<-function(x,shape,scale){
haz<-dGDUSED(x,shape,scale)/survGDUSED(x,shape,scale)
return(haz)
}
```

### A2

Creation of IUD Data

```
library(LaplacesDemon)
IUD <- c(10,13,18,19,23,30,36,38,54,56,59,75,93,97,104,107,107,107)
N <- 18
J <- 1
X <- matrix(1, nrow=N, ncol=J)
y <- IUD
censor <- c(1,0,0,1,0,1,1,rep(0,3),1,1,1,1,0,1,0,0)
mon.names <-c ("LP", "alpha")
parm.names <- as.parm.names(list(beta=rep(0,J), log.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, censor is a binary vector of censoring using 1 for uncensored and 0 for censored observation.

### A3

Model Specification

```
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 <- tcrossprod(beta, Data$X)
lambda <- exp(mu)
lf <- log(alpha)-log(lambda)-(y/lambda)+(alpha-1)*
log(1-exp(-(y/lambda)))+alpha*log(1-exp(-(y/lambda)))-log(exp(1)-1)
ls <- (1-(((exp(1-exp(-(y/lambda)))^(alpha)) -1)/(exp(1)-1)))
LL <- censor * lf + (1-censor) * ls
LL <- sum(LL)
## Log-Posterior
LP <- LL + beta.prior + alpha.prior
Modelout <- list(LP=LP, Dev=-2*LL, Monitor=c(LP,alpha),
yhat=lambda, parm=parm)
return(Modelout)
}
```

### A4

Fitting of Model with Laplace Approximation

```
Initial.Values <- c(coef(lm(log(y)~1)), log(1))
GDUSEDLA <- LaplaceApproximation(Model, Initial.Values, Data=MyData,
Iterations=5000, Method="TR")
print(GDUSEDLA)
```

### A5

Simulation Study of IUD Data of GDUSED

```
## Fitting with LaplacesDemon
Initial.Value<-as.initial.values(GDUSEDLA)
GDUSEDLD<-LaplacesDemon(Model, Data=MyData,Initial.Values, Covar=GDUSEDLA$Covar,
Iterations=50000,Status=100,Thinning=10,Algorithm="IM",
Specs=list(mu=GDUSEDLA$Summary1[1:length(Initial.Values),1]))
print(GDUSEDLD)
```

### A6

Creation of Data in JAGS

```
IUD <- data.frame(time = c(10,13,18,19,23,30,36,38,54,56,59,75,93,97,
104,107,107,107), censor = c(1,0,0,1,0,1,1,rep(0,3),1,1,1,1,0,1,0,0))
censor <- IUD$censor
y <- IUD$time
n <- length(y)
J <- 1
X <- matrix(1,nrow=n,ncol=J)
zeros <- rep(0,n)
C <- 10000
data <- list(n=n, J=J, y=y, X=X, zeros=zeros, censor=censor, C=C)
```

where  $n$  is number of observations,  $J$  number of predictors,  $X$  is the model matrix,  $zeros$  is a vector of zero values equal to the number of observations,  $censor$  indicates the censoring status, and  $C$  is sufficiently large positive value. Poisson zeros trick used for modelling Generalized DUS Exponential distribution GDUSED in JAGS.

## A7

### Model Definition

```
library(R2jags)
cat("model{
for (i in 1:n){
zeros[i] ~ dpois(phi[i])
phi[i] <- - l[i] + C
l[i] <- censor[i] * (log(alpha)-log(lambda[i])-(y[i]/lambda[i])+(alpha-1)
*log(1-exp(-(y[i]/lambda[i]))) + alpha*log(1-exp(-(y[i]/lambda[i]))) - log(exp(1)-1))
+ (1-censor[i]) * log(1-(((exp(1-exp(-(y[i]/lambda[i])))^(alpha)) - 1))/(exp(1)-1)))
log(lambda[i]) <- inprod(X[i,], beta[])
}
## Priors
for (j in 1:J){
beta[j] ~ dnorm(0, 0.001)
}
alpha ~ dunif(0, 100)
}", file="GDUSED.txt")
```

To Start the MCMC simulation, the initial values for the parameters are

```
inits <- list(list(alpha=0.78, beta=4.7), list(alpha=1.17, beta=1.5*4.7))
```

The above defined model is fitted with JAGS function

```
GDUSED.jags <- jags(data=data, inits=inits, param=c("beta","alpha"),
n.chains=2, n.iter=5000, model.file="GDUSED.txt")
print(GDUSED.jags)
```

## References

- AbuJarad MH, Khan AA (2018a). “Bayesian Survival Analysis of Topp-Leone Generalized Family with Stan.” *International Journal of Statistics and Applications*, **8**(5), 274–290.
- AbuJarad MH, Khan AA (2018b). “Exponential Model: A Bayesian Study with Stan.” *International Journal of Recent Scientific Research*, **9**(8), 28495–28506.
- Collett D (2015). *Modelling Survival Data in Medical Research*. Chapman and Hall/CRC.
- De Laplace PS (1774). “Mémoire sur les Suites Récurro-récurrentes et sur Leurs Usages dans la Théorie des Hasards.” *Mémoires de l’Académie royale des sciences (Paris)*, **6**, 353–371.
- Gelman A, Hill J (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge university press.
- Gelman A, Stern HS, Carlin JB, Dunson DB, Vehtari A, Rubin DB (2013). *Bayesian Data Analysis*. Chapman and Hall/CRC.
- Kumar D, Singh U, Singh SK (2015). “A Method of Proposing New Distribution and Its Application to Bladder Cancer Patients Data.” *Journal of Statistics Applications & Probability Letters*, **2**(3), 235–245.
- Lunn D, Jackson C, Best N, Spiegelhalter D, Thomas A (2012). *The BUGS Book: A Practical Introduction to Bayesian Analysis*. Chapman and Hall/CRC.
- Martyn P (2011). “JAGS Version 3.1. 0 User Mannual.” *International Agency for Research on Cancer*.
- Maurya SK, Kaushik A, Singh SK, Singh U (2017). “A New Class of Distribution Having Decreasing, Increasing, and Bathtub-shaped Failure Rate.” *Communications in Statistics-Theory and Methods*, **46**(20), 10359–10372.
- Mosteller F, Wallace D (1964). *Inference and Disputed Authorship: The Federalist*. Addison-Wesley.
- Nocedal J, Wright SJ (1999). “Numerical Optimization.” *New York*.
- Statisticat L (2015). “LaplacesDemon: Complete Environment for Bayesian Inference.” *R package version*, **15**(01). URL <http://www.bayesian-inference.com/software>.
- Su YS, Yajima M (2012). “R2jags: A Package for Running jags from R.” *R package version 0.03-08*, URL <http://CRAN.R-project.org/package=R2jags>.
- Tanner MA (2012). *Tools for Statistical Inference*. Springer.
- Tierney L, Kadane JB (1986). “Accurate Approximations for Posterior Moments and Marginal Densities.” *Journal of the American Statistical Association*, **81**(393), 82–86.

### Affiliation:

Mohammed H AbuJarad  
 Department of Statistics and Operations Research  
 Aligarh Muslim University  
 202002 Aligarh, India  
 E-mail: [m.jarad@gu.edu.ps](mailto:m.jarad@gu.edu.ps)



Athar Ali Khan  
Department of Statistics and Operations Research  
Aligarh Muslim University  
202002 Aligarh, India  
E-mail: [atharkhan1962@gmail.com](mailto:atharkhan1962@gmail.com)

Eman S A AbuJarad  
Department of Mathematics  
Aligarh Muslim University  
202002 Aligarh, India  
E-mail: [emanjarad2@gmail.com](mailto:emanjarad2@gmail.com)