JP Journal of Biostatistics
Volume 9, Number 1, 2013, Pages 1-26
Published Online: June 2013
Available online at http://pphmj.com/journals/jpb.htm
Published by Pushpa Publishing House, Allahabad, INDIA

COMPARING RELATIVE EFFICIENCY FROM A RIGHT SKEWED MODEL

Hafiz M. R. Khan

Department of Biostatistics
Robert Stempel College of Public Health and Social Work
Florida International University
11200 S.W. 8th Street
Miami, FL 33199, U. S. A.
e-mail: hmkhan@fiu.edu

Abstract

Type II censored samples have been extensively collected from many fields including engineering and biomedical sciences. These samples frequently contain extreme values and the predictive inference from such type of samples may produce invalid results. In this paper, we compare predictive variability with respect to the location parameter mean as well as median. A simulation study and a real life study are considered to determine the predictive results. It is obtained that the predictive estimates with the location parameter median give more precise results compared with the location parameter mean when type II censored sample contains extreme values.

1. Introduction

Predictive inference has been used to solve many problems in the fields

Keywords and phrases: censored sample, half-normal model, statistical inference, predictive inference.

Communicated by K. K. Azad

Received December 9, 2012

of engineering and biomedical sciences. In order to make a predictive inference, the recorded sample may be viewed as a censored sample. A censored sample is a partially observed sample from an experiment. There are several published statistical research papers available based on censored samples, including Ahmed and Saleh [3], Ahmed and Rahbar [4], and Buhamra et al. [10].

In life testing experiments, the failure times of certain devices may be recorded, but in order to record the failures of the devices, the experimenter would have to wait until all the devices have failed. Sometimes, in order to save both time and money, the experimenter may choose to stop the experiment after observing a reasonable number of devices' failure times. In such a situation, failure times occur in an ordered fashion and form a type II censored sample. In the biomedical field, one may wish to determine the effectiveness of a certain drug by analyzing survival data in connection with lab experiments. The survival times for a given treatment are recorded as life data. These data usually occur in an ordered fashion; for example, in life tests the weakest 'unit' fails first, the second weakest fails next, and so on. These failure times data are modeled as life testing models.

Studies related to type II censored sample have been published in several referred articles, for example, Khan et al. [15], Buhamra et al. [10, 11]. Ahmed [2] described an asymptotic estimation of reliability in a life testing model. Bakilzi and Ahmed [7] used the Weibull lifetime model to discuss the estimation of the reliability function.

The mean is an appropriate measure of the central tendency when the distribution of a data set is symmetrical. In this case, the data points are usually concentrated at the middle, where the mean is located. When the distribution is skewed, then the mean is not in the middle. At this stage, the median is particularly useful due to a few extreme data points. For example, the mean of the ten numbers 1, 1, 1, 2, 2, 3, 5, 6, 15, 30 is 6.6. Eight of the ten numbers are less than the mean, with only two of the ten numbers greater than the mean. In this situation, a better measure of the center for this distribution would be the median, which is 2.5. In this example, the mean is

greater than the median. This is common for a distribution that is skewed to the right. One may choose to use the middle observations for statistical inference from this type of data set.

Another example, in calculating the mean salary obtained by a group of people in a community with different professions, it may happen that a few people are earning extraordinarily high salaries. To calculate the mean salary of the people, if one considers all values including very high values, then mean will be affected by a few extreme salaries. Statistical models from such a type of sample may be viewed as right skewed models. Predictive inferences from such a sample are not accurate and often lead to invalid results. In such situations, one may consider the predictive analysis based on the location parameter mean as well as median and compare their results. The median can be used as a measure of location when a distribution is skewed. When data contain extreme values or there may exist measurement error that may not be known, researchers may typically derive the predictive estimates for future responses by choosing an appropriate location parameter. The present work deals with the predictive inferences for future responses from the right skewed half-normal model based on the location parameter mean as well as median.

Bayesian predictive inference has been studied by many authors, including Thabane and Haq [18], Khan [17]. In order to derive the predictive model for a future response, the Bayesian approach may be considered. Berger [8] discussed a general Bayesian prediction problem. Geisser [13] described the inferential use of predictive distributions, and Geisser [14] considered various Bayesian predictive problems for future responses. Evans and Nigm [12] used the Bayesian approach to derive future responses from the Weibull distribution. Additional applications of the Bayesian approach to predictive inference have been discussed for instance, by Bernardo and Smith [9], and Thabane and Haq [18]. Ahsanullah and Ahmed [5] discussed in detail Bayes and empirical Bayes estimates of survival, and hazard functions of a class of distributions. Ahmed [1] used a priori information in the estimation of Poisson parameter.

Although a number of studies related to the half-normal model have been conducted for the parameters by several authors based on non-censored samples but none of the studies used predictive inference for future responses based on type II censored samples. The goal of this paper is to derive predictive model for future responses and compare the variability by considering the location parameter mean as well as median on the basis of a type II censored sample from the two-parameter half-normal model whose density function is given by

$$p(x\,|\,\mu,\,\sigma) = \begin{cases} \dfrac{2}{\sigma\sqrt{2\pi}} \exp\left\{-\dfrac{(x-\mu)^2}{2\sigma^2}\right\}, & \text{for } x \geq \mu;\ \sigma > 0, \\ 0, & \text{elsewhere,} \end{cases} \tag{1}$$

where μ is the mean as a location parameter and σ is a scale parameter.

The rest of the paper is organized as follows: Section 2 presents the predictive density for a single future response under a type II censored sample. To illustrate the results, a simulated sample and a real life sample are considered in Section 3. Finally, a conclusion is added in Section 4.

2. The Predictive Approach

In a biomedical experiment, one may decide to terminate an experiment after recording the lifetimes of a preassigned number of mice. Let m mice be followed after receiving a concentration of dose of a drug and assume that the experiment is terminated after k mice's survival days are recorded. Let x_1 be the first mouse's survival days, x_2 be the next mouse's survival days, and so on. In total, $k(\leq m)$ lifetimes are being observed with no observation on the remaining $(m-k)$ mice. Thus, $x_1 \leq x_2 \leq \cdots \leq x_k$; and $\mathbf{x} = (x_1, ..., x_k)'$ forms an observed type II censored sample. Following Khan [16], let z be a future response, which is independent of the observed data, and the pdf of z may be defined from model (1). The predictive density for a single future response z given $\mathbf{x} = (x_1, ..., x_k)$ is

$$p(z \mid \mathbf{x}) = \begin{cases} \Psi_1(\mathbf{x}) \int_{\mu=0}^{\min(x_1, z)} \\ \times \sum_{n=0}^{+\infty} \dfrac{(-1)^{n(m-k)} ((2n-1)!!)^{(m-k)} \Gamma\left(\dfrac{k-(2n+1)(m-k)+2b+1}{2} \right)}{2^{n(m-k)-k+(2n+1)(m-k)-2b} (x_k - \mu)^{(2n+1)(m-k)}} \\ \times \left[a + (z-\mu)^2 + \sum_{i=1}^{k} (x_i - \mu)^2 + (m-k)(x_k - \mu)^2 \right]^{-\frac{2k-m(2n+1)+2nk+2b+1}{2}} d\mu; \\ \qquad \text{for } z \geq 0, \\ 0, \qquad \text{elsewhere,} \end{cases} \tag{2}$$

where $\Psi_1(\mathbf{x})$ is a normalizing constant. The above model can be used to obtain the predictive inference for a single future response on the basis of the location parameter mean (μ).

One may replace the location parameter μ with $\tilde{\mu}$ in equation (2), where $\tilde{\mu}$ is the median. Thus the predictive density for a single future response z given $\mathbf{x} = (x_1, ..., x_k)$ is

$$p(z \mid \mathbf{x}) = \begin{cases} \Phi_1(\mathbf{x}) \int_{\tilde{\mu}=0}^{\min(x_1, \tilde{z}_{(\cdot)})} \\ \times \sum_{n=0}^{+\infty} \dfrac{(-1)^{n(m-k)} ((2n-1)!!)^{(m-k)} \Gamma\left(\dfrac{k-(2n+1)(m-k)+2b+1}{2} \right)}{2^{n(m-k)-k+(2n+1)(m-k)-2b} (x_k - \tilde{\mu})^{(2n+1)(m-k)}} \\ \times \left[a + (z-\tilde{\mu})^2 + \sum_{i=1}^{k} (x_i - \tilde{\mu})^2 + (m-k)(x_k - \tilde{\mu})^2 \right]^{-\frac{2k-m(2n+1)+2nk+2b+1}{2}} d\tilde{\mu}; \\ \qquad \text{for } z \geq 0, \\ 0, \qquad \text{elsewhere,} \end{cases} \tag{3}$$

where $\Phi_1(\mathbf{x})$ is a normalizing constant. Equation (3) can be used to obtain

predictive inferences for a future response on the basis of the location parameter median $(\tilde{\mu})$.

The hyperparameters are usually unknown and may be estimated from the data set by applying the maximum likelihood method; see Geisser [14] and Berger [8]. When the maximum likelihood equations (MLEs) do not yield closed form representations for *a* and *b*, one may then consider some arbitrary values for *a* and *b* to perform statistical analysis, see Khan et al. [15]. An advanced statistical package may then be used to display the predictive density graphically with several choices of *a* and *b* and a scientific conclusion can be drawn.

3. Numerical Studies

We consider two numerical studies. The first study describes a simulated sample from the half-normal model, and the second study presents a real life survival sample of hepatocellular carcinoma of liver patients. Equations (2) and (3) are utilized for both studies.

3.1. Simulation study

We have simulated a sample of size $m = 25$ with some higher extremes from the half-normal model. To make a type II censored sample, data are arranged in order, and the last five observations are discarded from the ordered sample. Therefore, the value for the x_k was the 20th position. This sample was utilized to evaluate its corresponding normalizing constants and to plot the predictive graphs. It is also used to calculate the predictive variance and average absolute deviation from mean as well as median.

We used the numerical integration command 'NIntegrate' in conjunction with the symbolic computational software Mathematica version 7.0, Wolfram Research. This package was utilized to carry out all predictive related calculations. For the estimation of the hyperparameters, this software failed to estimate the hyperparameters for *a* and *b* from the likelihood equations and the Newton-Raphson iterative algorithm also fails to achieve solutions for *a* and *b*, simultaneously. Thus, we considered some arbitrary values of *a*

and b, and these values were substituted in equations (2) and (3) to determine the predictive density.

Graphical representations of the predictive densities are given with certain values of the hyperparameters with the location parameter mean (Figure 1). The figure at the top $(a = 1, b = 1; a = 1, b = 5)$ indicates that the predictive density with hyperparameters $a = 1,\ b = 5$ has a lower variance compared with the variance of the predictive density with $a = 1,\ b = 1$. Similarly, the predictive density with $a = 16,\ b = 5$ has a higher variance compared with the variance of the predictive density with the hyperparameters $a = 2,\ b = 5$.

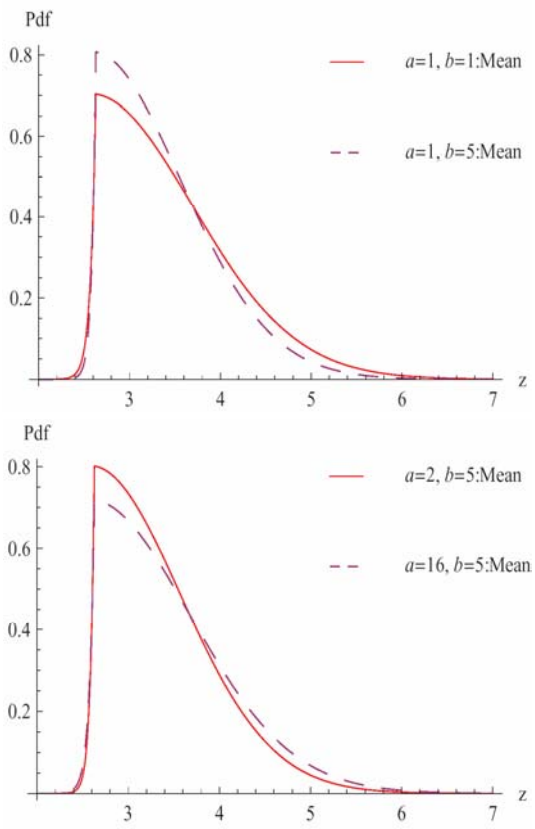

Figure 1. Comparison of variability of the predictive densities for a single future response with the location parameter mean under certain values of the hyperparameters.

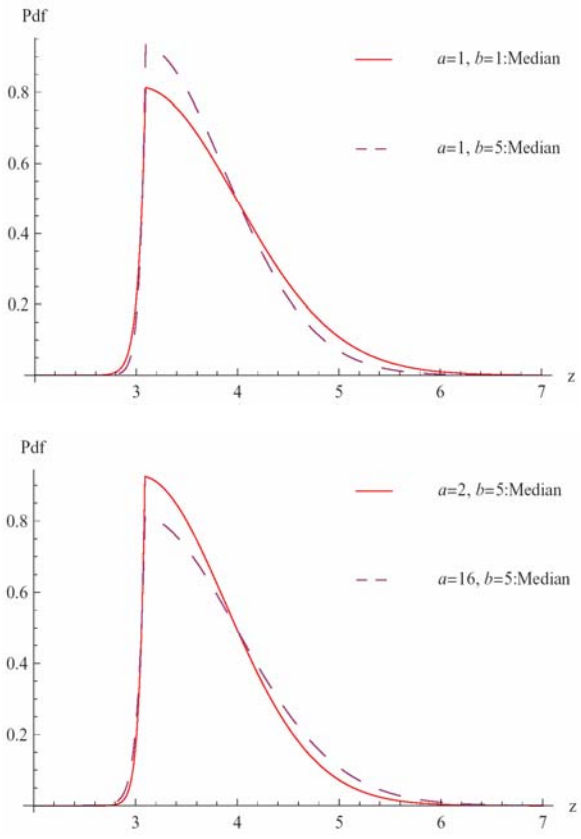

Figure 2. Comparison of variability of the predictive densities for a single future response with the location parameter median under certain values of the hyperparameters.

Based on the location parameter median, the graphical representations of the predictive densities are given with respect to certain values of the hyperparameters (Figure 2). The figure at the top $(a = 1, b = 1; a = 1, b = 5)$ indicates that the predictive density with hyperparameters $a = 1,\ b = 5$ has a lower variance compared with the predictive variance with $a = 1,\ b = 1$. Similarly, the predictive density with $a = 16,\ b = 5$ has a higher variance compared with the variance of the predictive density with the hyperparameters $a = 2,\ b = 5$.

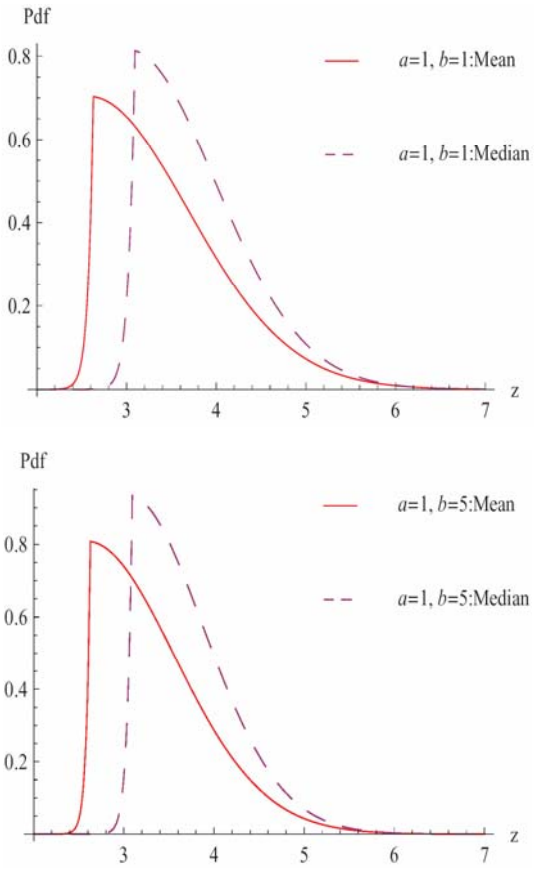

Figure 3. Comparison of variability of the predictive densities for a single future response with the location parameter mean and median under certain values of the hyperparameters.

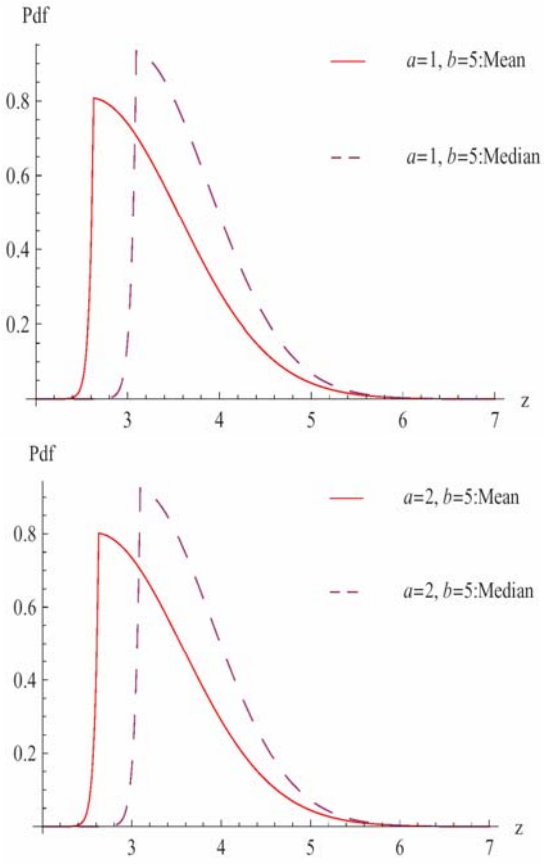

Figure 4. Comparison of variability of the predictive densities for a single future response with the location parameter mean and median under certain values of the hyperparameters.

Figures 3 and 4 are the comparisons of the predictive variability based on the location parameter mean and median with respect to some combination of values of the hyperparameters. In Figure 3, two graphs are superimposed. The first graph is a combination of the hyperparameters $a = 1,\ b = 1,$ and is drawn on the basis of the mean and median. The second graph is a combination of the hyperparameters $a = 1,\ b = 5,$ and is drawn on the basis of the mean and median. When the hyperparameter a is fixed and b varies, one would observe the predictive variability given the location parameter.

Given the location parameter median, one would obtain a narrower predictive interval and given the location parameter mean, one would obtain a wider predictive interval.

In Figure 4, two graphs are superimposed. The first graph is a combination of the hyperparameters $a = 1,\ b = 5,$ and is drawn on the basis of the location parameter mean and median. The second graph is a combination of the hyperparameters $a = 2,\ b = 5,$ and is drawn on the basis of the mean and median. In this case, the graphs look almost the same, but the second graph's variability is slightly higher than that of the first graph since the hyperparameter *a* widens the interval.

Table 1 includes the estimates of the predictive variance with the location parameter mean as well as median with respect to several values of the hyperparameters. Comparing all the results, it may be commented that the predictive interval becomes wider with an increased value of *a*. For larger values of *b*, the predictive interval becomes narrower.

The relative efficiency for a single future response is calculated on the basis of the ratio of the two variances, i.e.,

$$\frac{\text{Variance}(z_{\text{mean}})}{\text{Variance}(z_{\text{median}})}.$$

It is noted that as *a* increases with a fixed *b*, the relative efficiency slowly decreases. When *a* is fixed and *b* increases, the relative efficiency increases.

Table 2 presents the estimates of the average absolute deviation from the mean as well as median with respect to several values of the hyperparameters. It may be commented that the estimated absolute deviation is minimum when it is measured from the median.

Table 1. Comparison of variances for predictive densities with several values of the hyperparameters when the location parameter is considered mean as well as median

Hyperparameters		Variance (z_{mean})	Variance (z_{median})	Relative efficiency
a	b			
1	1	0.487820	0.366278	1.33183
2	1	0.496841	0.373932	1.32869
3	1	0.504421	0.381586	1.32190
4	1	0.511998	0.389239	1.31538
1	2	0.451763	0.338117	1.33611
1	3	0.419559	0.313956	1.33636
1	4	0.391618	0.293006	1.33655
1	5	0.366992	0.274520	1.33685
3	2	0.465787	0.352256	1.32230
4	2	0.472798	0.359327	1.31579
5	2	0.479807	0.366398	1.30952
6	2	0.486815	0.373469	1.30350
2	3	0.426077	0.320521	1.329330
2	4	0.397704	0.299132	1.329530
2	5	0.372860	0.280411	1.329690
2	6	0.350931	0.263891	1.329830

Table 2. Comparison of the deviation measured from mean as well median for predictive densities with several values of the hyperparameters

Hyperparameters		$E\lvert z - \text{mean}\rvert$	$E\lvert z - \text{median}\rvert$	$\Psi_1(\mathbf{x})$	$\Phi_1(\mathbf{x})$
a	b				
1	1	0.853860	0.717214	2.05482×10^{-48}	2.87460×10^{-45}
2	1	0.859895	0.723970	1.38982×10^{-48}	1.70323×10^{-45}
3	1	0.865982	0.730645	9.45499×10^{-49}	1.01988×10^{-45}
4	1	0.872016	0.737243	6.46937×10^{-49}	6.16913×10^{-46}
1	2	0.823302	0.692195	4.38945×10^{-52}	1.10512×10^{-48}
1	3	0.795940	0.669693	9.45265×10^{-56}	4.28312×10^{-52}
1	4	0.771179	0.649311	2.05014×10^{-59}	1.67173×10^{-55}
1	5	0.748632	0.630732	4.47419×10^{-63}	6.56516×10^{-59}
3	2	0.835132	0.705201	1.89988×10^{-52}	3.61411×10^{-49}
4	2	0.840971	0.711591	1.26172×10^{-52}	2.10143×10^{-49}
5	2	0.846760	0.717908	8.42975×10^{-53}	1.23478×10^{-49}
6	2	0.852502	0.724154	5.66506×10^{-53}	7.32917×10^{-50}
2	3	0.801699	0.676041	6.01126×10^{-56}	2.33734×10^{-52}
2	4	0.776772	0.655481	1.26427×10^{-59}	8.75519×10^{-56}
2	5	0.754072	0.636739	2.67556×10^{-63}	3.29978×10^{-59}
2	6	0.733284	0.619561	5.69346×10^{-67}	1.25045×10^{-62}

Table 3. Summary results for the posterior parameters and the hyperparameters based on simulated data by making use of the software WinBUGS

Node	Mean	SD	MC Error	2.5%	Median	97.5%	Start	Sample
a	100.5	70.87	0.2954	12.34	84.50	279.8	1001	60000
b	19.96	14.11	0.057	2.473	16.75	55.77	1001	60000
μ	3.465	0.2522	$9.911E^{-4}$	2.968	3.465	3.965	1001	60000
σ	0.872	0.2841	0.001248	0.4086	0.8404	1.51	1001	60000

The Bayesian Markov chain Monte Carlo (MCMC) method is a useful alternative to obtain the posterior distributions of the parameters. The implementation of this method can be obtained by making use of the software WinBUGS. The BUGS (Bayesian Inference Using Gibbs Sampling) project is designed with a software package on the basis of carrying out MCMC computations for a wide variety of Bayesian models. MCMC methods are a class of algorithms for selecting samples from probability distributions on the basis of a Markov chain. In the case of the simulated data, there are 1,001 burn-in samples excluded, and the results are based on the additional 60,000 samples. The summary results of the parameters and hyperparameters are reported in Table 3. Figures 5 and 6 are the kernel densities, dynamic trace, and quantile plots for the parameters and hyperparameters.

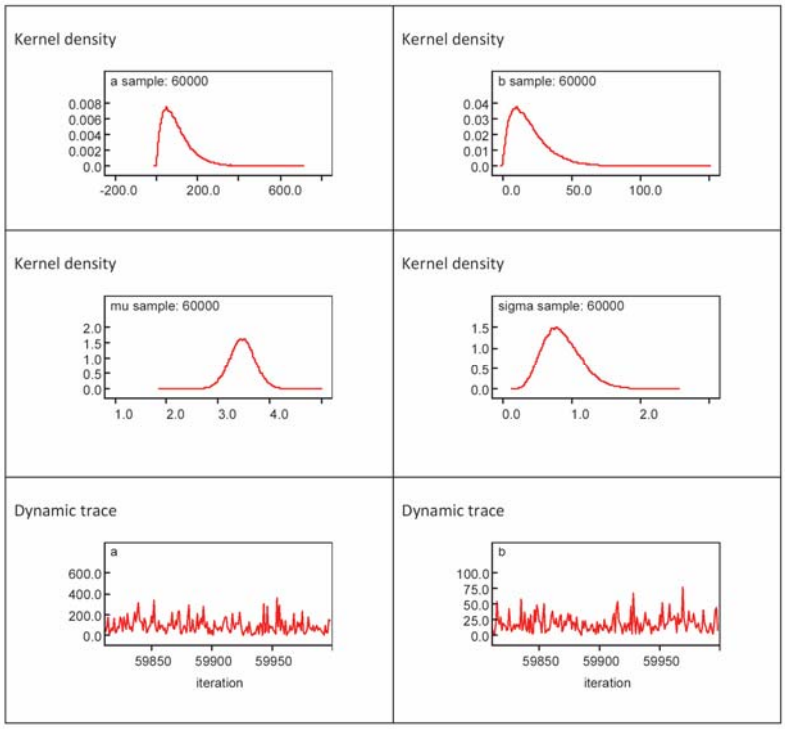

Figure 5. Posterior densities and dynamic trace plots for the parameters and hyperparameters based on simulated data.

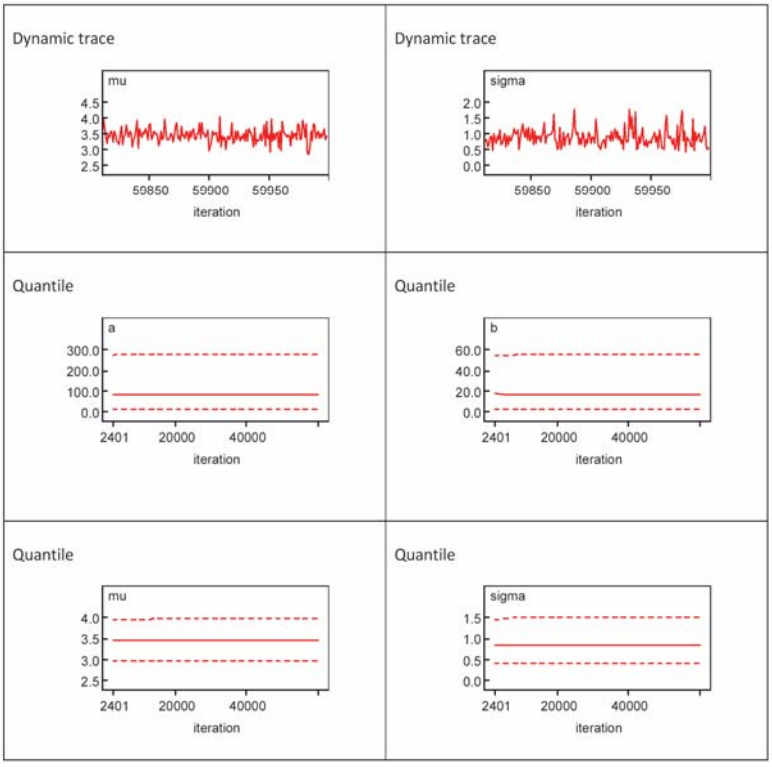

Figure 6. Trace and quantile plots for the parameters and hyperparameters based on simulated data.

3.2. Real data study

In this subsection, we consider a health related cancer data. Hepatocellular carcinoma (HCC) is the fifth most common cancer in the world with 80% of cases occurring in developing countries. The cancer is rapidly fatal in almost all cases with survival generally less than 1 year from diagnosis. The major risk factors for this cancer have been identified as chronic infection with hepatitis B (HBV) and hepatitis C (HCV) viruses and dietary exposure to aflatoxins. There is a safe and effective vaccine to prevent chronic HBV infection. Given estimates that approximately 70% of HCC in developing countries is attributable to HBV, then vaccination could prevent more than 250,000 cases per year in these areas of the world. A

major challenge now is to ensure the availability of vaccine in countries with endemic infection. Development of a vaccine against HCV is more problematic due to the genetic heterogeniety of the virus. However, with 24% of HCC in developing countries attributable to HCV (approximately 93,000 cases per year), a vaccine would make a major contribution to cancer prevention, Wild and Hall [19].

A survival analysis of 20 patients diagnosed with hepatocellular carcinoma of liver (HCC) in 2002 is performed. The data set is taken from a cancer registry at the University Hospital located in Newark, New Jersey. The survival days of 17 patients are recorded, and three out of 20 patients are lost to follow-up. They are the last three patients' survival days in order. These survival days can be formed as a type II censored sample.

This data set is composed of 17 patients' survival days and it contains a few extreme values. We consider several statistical probability models to fit the data. A Q-Q (quantile-quantile) plot is used to check whether or not a data-based sample comes from a specific population. We reported some Q-Q plots in Figures 7-9. The histogram of the survival days is shown in Figure 9. Based on the Q-Q plots, it is determined that the half-normal model is the appropriate fit (Anderson-Darling Test Statistics = 0.8521, $P = 0.6937$) for modeling the survival data of hepatocellular carcinoma of liver patients. The gamma model has the next lowest differences and the exponential model has the highest differences, see in Table 4.

Table 4. Estimated values and fitted distributions

Distributions	Anderson-Darling statistics	*P*-value
Half-normal	0.8521	0.6937
Gamma	1.1723	0.3410
Exponential	3.2370	0.2100

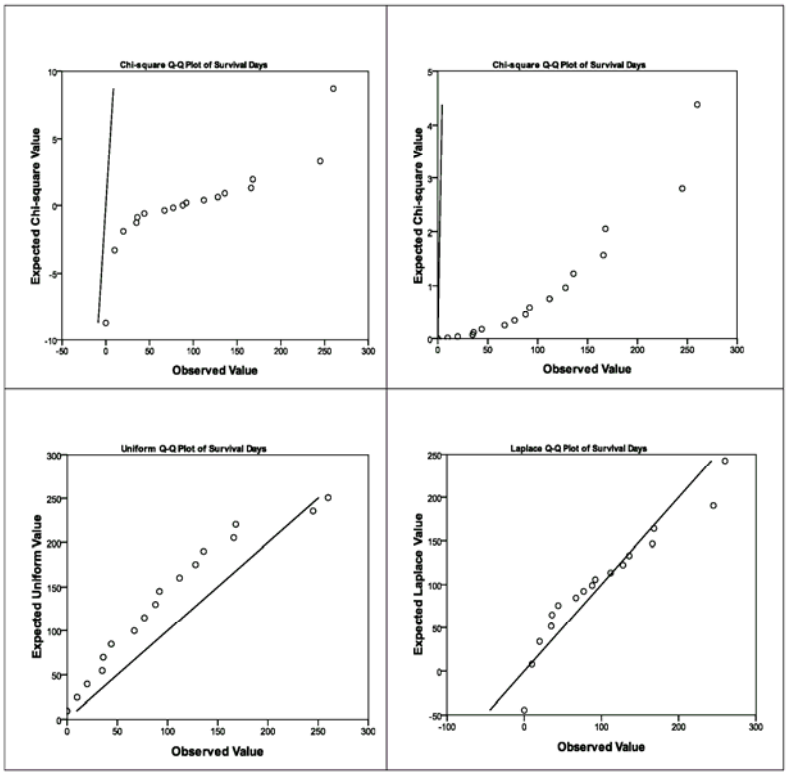

Figure 7. Q-Q plots for the survival days of HCC patients.

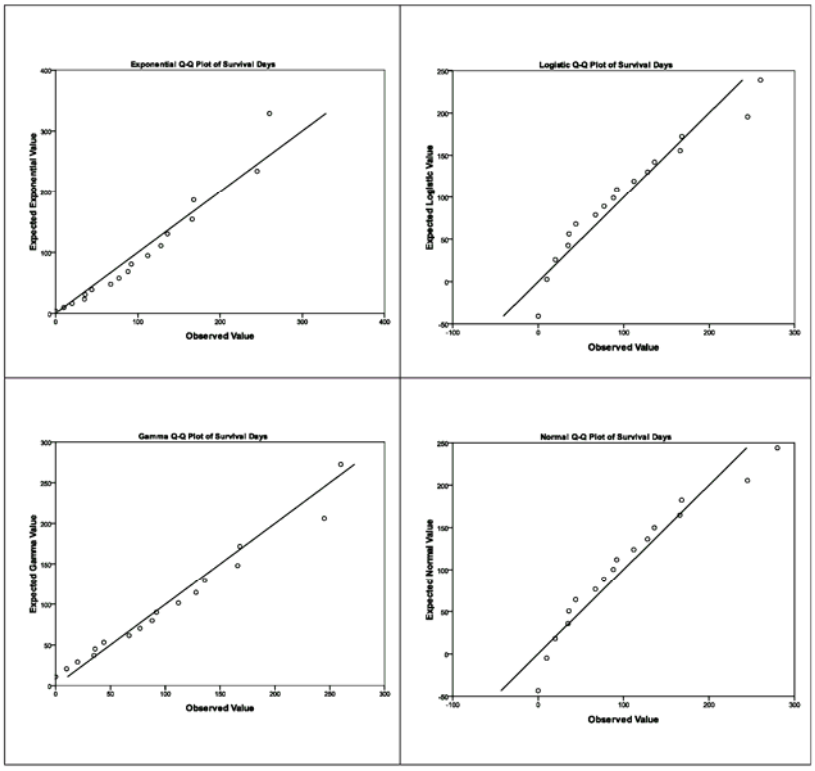

Figure 8. Q-Q plots for the survival days of HCC patients.

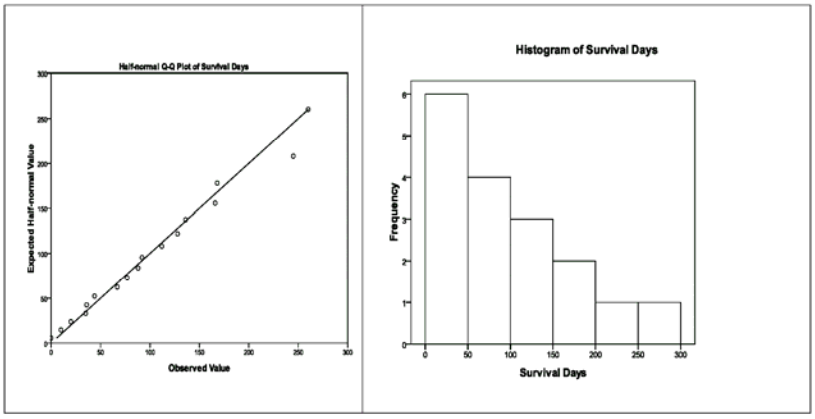

Figure 9. Q-Q plot and histogram for the survival days of HCC patients.

In addition to checking the goodness of fit, we also compared the empirical distribution function (EDF) with the corresponding theoretical cumulative distribution function (CDF) given the sample of survival days. We consider two measures of discrepancy are being used to determine the best fit model. We considered the exponential, gamma, and half-normal models. The mean squared differences, i.e.,

$$\sum_{i=1}^{n} (F(x_i) - G(x_i))^2 / n,$$

and the absolute mean differences i.e.,

$$\sum_{i=1}^{n} | F(x_i) - G(x_i) | / n$$

are considered, where F is the empirical distribution function, G is the functional representation of the cumulative distribution function, and n is the observed sample size. It was obtained that the half-normal model is the best fit considering the lowest differences of mean squared differences and absolute mean differences. The gamma model has the next lowest differences and the exponential model has the highest differences, see in Table 5.

Table 5. Estimated values and fitted distributions

Distributions	$\sum_{i=1}^{17} (F(x_i) - G(x_i))^2/17$	$\sum_{i=1}^{17} \| F(x_i) - G(x_i) \|/17$
Half-normal	0.000067	0.023974
Gamma	0.004346	0.051685
Exponential	0.014260	0.104116

The predictive variances of a single future response are estimated given the 17 patients' survival days with respect to some values of the hyperparameters on the basis of the location parameter mean and median. The relative efficiency is estimated given each combination of the hyperparameters. Table 6 includes the estimate of the predictive variance based on the location parameter mean as well as median with respect to several values of the hyperparameters. Table 7 presents the estimate of the

average absolute deviation from the mean as well as median with respect to several values of the hyperparameters. It is obvious from Tables 6 and 7 that based on the median, the predictive estimates are better. The estimated absolute deviation is minimum when it is measured from the median.

Table 6. Comparison of variances for predictive densities with several values of the hyperparameters of the HCC patients survival days when the location parameter is considered mean as well as median

Hyperparameters		Variance (z_{mean})	Variance (z_{median})	Relative efficiency
a	b			
1	1	3450.86	1505.17	2.29267
2	1	3450.87	1505.18	2.29266
3	1	3450.88	1505.19	2.29265
4	1	3450.89	1505.20	2.29264
1	2	3225.93	1391.58	2.31818
1	3	3022.44	1292.11	2.33915
1	4	2838.88	1204.66	2.35658
1	5	2673.12	1127.44	2.37096
3	2	3225.79	1391.59	2.31806
4	2	3225.80	1391.60	2.31805
5	2	3225.81	1391.61	2.31804
6	2	3225.82	1391.62	2.31803
2	3	3022.45	1292.11	2.33916
2	4	2838.88	1204.66	2.35658
2	5	2673.12	1127.45	2.37094
2	6	2523.29	1058.91	2.38291

Table 7. Comparison of the deviation measured from mean as well median for predictive densities of the survival days of HCC patients with several values of the hyperparameters

Hyperparameters		$E\lvert z - \text{mean}\rvert$	$E\lvert z - \text{median}\rvert$	$\Psi_1(\mathbf{x})$	$\Phi_1(\mathbf{x})$
a	b				
1	1	77.9911	45.793	2.21664×10^{-134}	1.84105×10^{-123}
2	1	77.9912	45.793	2.21652×10^{-134}	1.84098×10^{-123}
3	1	77.9913	45.793	2.21652×10^{-134}	1.84075×10^{-123}
4	1	77.9914	45.793	2.21629×10^{-134}	1.84051×10^{-123}
1	2	75.1406	44.1389	9.8657×10^{-146}	4.70565×10^{-134}
1	3	72.5452	42.6432	4.40221×10^{-157}	1.2131×10^{-144}
1	4	70.1752	41.2854	1.9688×10^{-168}	3.15087×10^{-155}
1	5	68.0044	40.0473	8.82293×10^{-180}	8.23774×10^{-166}
3	2	75.1408	44.1391	9.86458×10^{-146}	4.70436×10^{-134}
4	2	75.1409	44.1392	9.86401×10^{-146}	4.70372×10^{-134}
5	2	75.1409	44.1393	9.86345×10^{-146}	4.70307×10^{-134}
6	2	75.1410	44.1394	9.86289×10^{-146}	4.70243×10^{-134}
2	3	72.5453	42.6433	4.40194×10^{-157}	1.21292×10^{-144}
2	4	70.1752	41.2855	1.96867×10^{-168}	3.15037×10^{-155}
2	5	68.0044	40.0473	8.82231×10^{-180}	8.23635×10^{-166}
2	6	68.0098	38.9137	3.96079×10^{-191}	2.16577×10^{-178}

In the survival data analysis of 17 HCC patients, there are 1,001 burn-in samples excluded, and the results are based on the additional 60,000 samples. The summary results of the parameters and hyperparameters are reported in Table 8. Figures 10 and 11 are the kernel densities, dynamic trace, and quantile plots for the parameters and hyperparameters of the survival days of

the HCC patients. Figure 12 presents a graphical comparison of the variance and relative efficiency on the basis of the simulated study and real data study.

Table 8. Summary results for the posterior parameters and the hyperparameters in the case of survival days of the HCC patients by making use of the software WinBUGS

Node	Mean	SD	MC Error	2.5%	Median	97.5%	Start	Sample
a	100.4	70.88	0.2771	12.32	84.44	280.0	1001	6000
b	19.99	14.12	0.061	2.453	16.81	55.67	1001	6000
μ	68.86	19.31	0.08898	28.05	69.96	103.9	1001	6000
σ	$1.493E^{-4}$	$5.709E^{-5}$	$2.662E^{-7}$	$5.941E^{-5}$	$1.419E^{-4}$	$2.794E^{-4}$	1001	6000

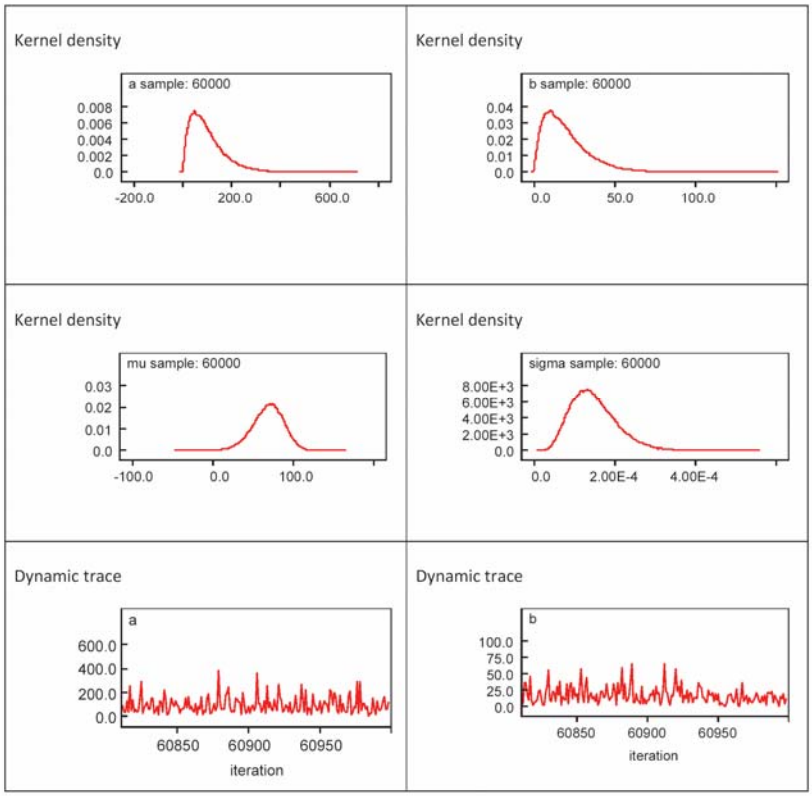

Figure 10. Posterior densities and dynamic trace plots for the parameters and hyperparameters of the survival days of the HCC patients.

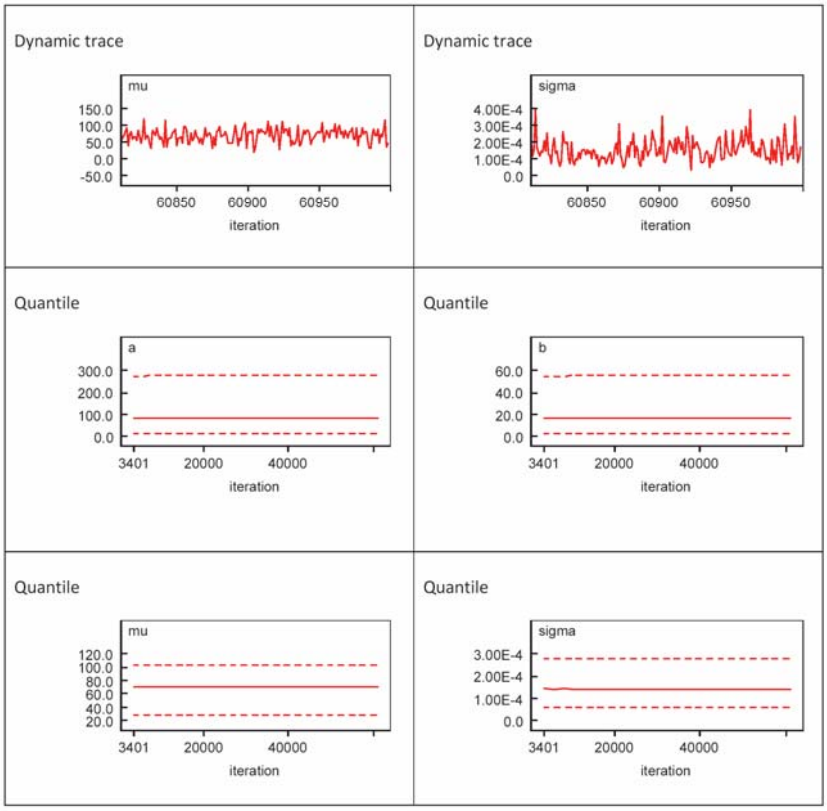

Figure 11. Trace and quantile plots for the parameters and hyperparameters of the survival days of the HCC patients.

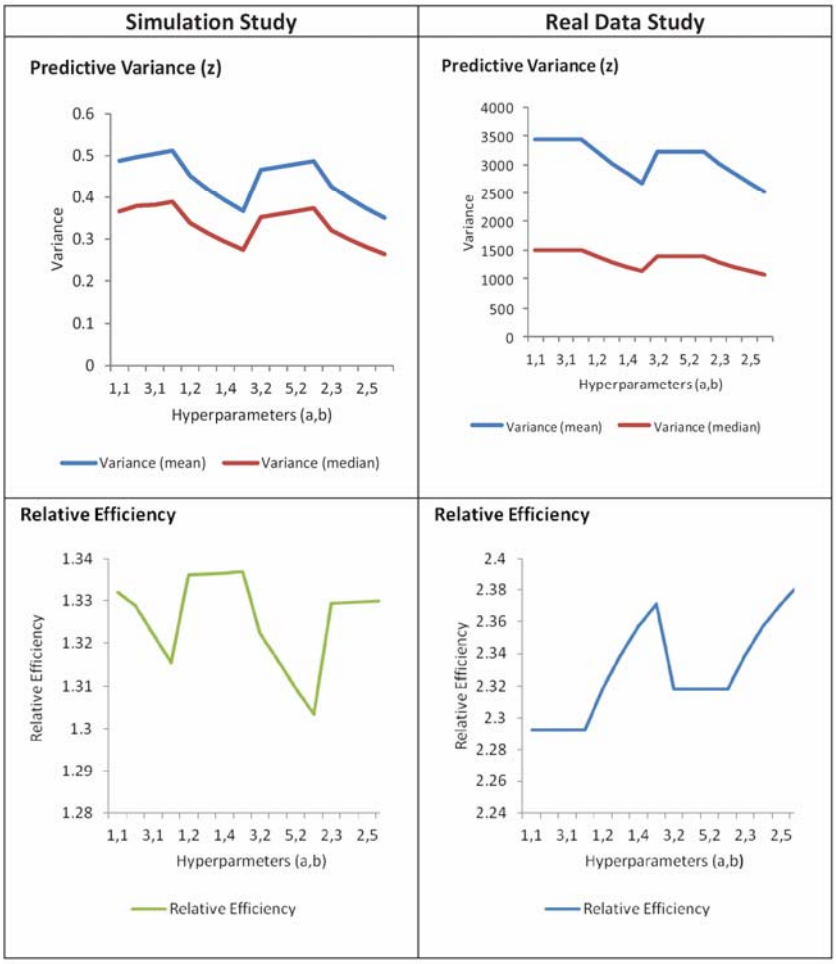

Figure 12. Predictive variability plots with respect to certain values of the hyperparameters for both studies.

4. Conclusion

We considered some arbitrary values of the hyperparameters to obtain the predictive variance with respect to the location parameter mean as well as median. Comparing all the predictive variability results, it is commented that the location parameter based on the median gives precise results. The WinBUGS software was used to obtain the summary results for the posterior

model parameters and hyperparameters. It is also used to graphically display the posterior densities, dynamic trace, and quantile plots for the parameters and hyperparameters. Furthermore, it was determined that the average predictive absolute deviation from the median is smaller than the average predictive absolute deviation from the mean with respect to the choices of the hyperparameters.

Acknowledgement

The author would like to thank the editor and the referees for their valuable comments and suggestions.

References

[1] S. E. Ahmed, Use of a priori information in the estimation of Poisson parameter, Soochow J. Math. 16 (1990), 185-192.

[2] S. E. Ahmed, Asymptotic estimation of reliability in a life-testing model, Journal of Industrial Mathematics Society 41 (1992), 7-18.

[3] S. E. Ahmed and A. K. Md. E. Saleh, Estimation of regression coefficients in an exponential regression model with censored observations, J. Japan Statist. Soc. 29 (1999), 55-64.

[4] S. E. Ahmed and M. H. Rahbar, Shrinkage and pretest nonparametric estimation of regression parameters from censored data with multiple observations at each level of covariate, Biom. J. 42 (2000), 511-525.

[5] M. Ahsanullah and S. E. Ahmed, Bayes and empirical Bayes estimates of survival and hazard functions of a class of distributions, Lecture Notes in Statist., 148 Springer, New York, 2001, pp. 81-87.

[6] L. J. Bain and M. Engelhardt, Statistical Analysis of Reliability and Life Testing Models: Theory and Methods, Marcel Dekker Inc., New York, 1991.

[7] A. Bakilzi and S. E. Ahmed, On the estimation of reliability function in a Weibull lifetime distribution, Statistics 42 (2008), 351-362.

[8] J. O. Berger, Statistical Decision Theory and Bayesian Analysis, Springer-Verlag, New York, 1985.

[9] J. M. Bernardo and A. F. M. Smith, Bayesian Theory, Wiley, New York, 1994.

[10] S. S. Buhamra, N. M. Al-Kandari and S. E. Ahmed, Inference concerning quantile for left truncated and right censored data, Comput. Statist. Data Anal. 46 (2004), 819-831.

[11] S. S. Buhamra, N. M. Al-Kandari and S. E. Ahmed, Nonparametric inference strategies for the quantile functions under left truncation and right censoring, J. Nonparametr. Stat. 19 (2007), 189-198.

[12] I. G. Evans and A. M. Nigm, Bayesian prediction for two-parameter Weibull lifetime models, Comm. Statist. Theory Methods A 9(6) (1980), 649-658.

[13] S. Geisser, The inferential use of predictive distributions, Foundation of Statistical Inference, V. P. Godambe and D. A. Sprott, eds., Holt, Toronto, 1971, pp. 456-469.

[14] S. Geisser, Predictive Inference: An Introduction, Chapman and Hall, New York, 1993.

[15] H. M. R. Khan, M. S. Haq and S. B. Provost, Predictive distributions for responses from the Weibull life testing model, J. Stat. Theory Appl. 3(1) (2004), 53-73.

[16] H. M. R. Khan, Predictive inference from the half-normal model given a type II censored sample, Comm. Statist. Theory Methods 42(1) (2012), 42-55.

[17] S. Khan, Predictive distribution of regression vector and residual sum of squares for normal multiple regression model, Comm. Statist. Theory Methods 33(10) (2004), 2423-2441.

[18] L. Thabane and M. S. Haq, Bayesian prediction from a compound statistical model: an actual application, Pakistan J. Statist. 16(2) (2000), 110-125.

[19] C. P. Wild and A. J. Hall, Primary prevention of hepatocellular carcinoma in developing countries, Mutat. Res. 462(2-3) (2000), 381-393.