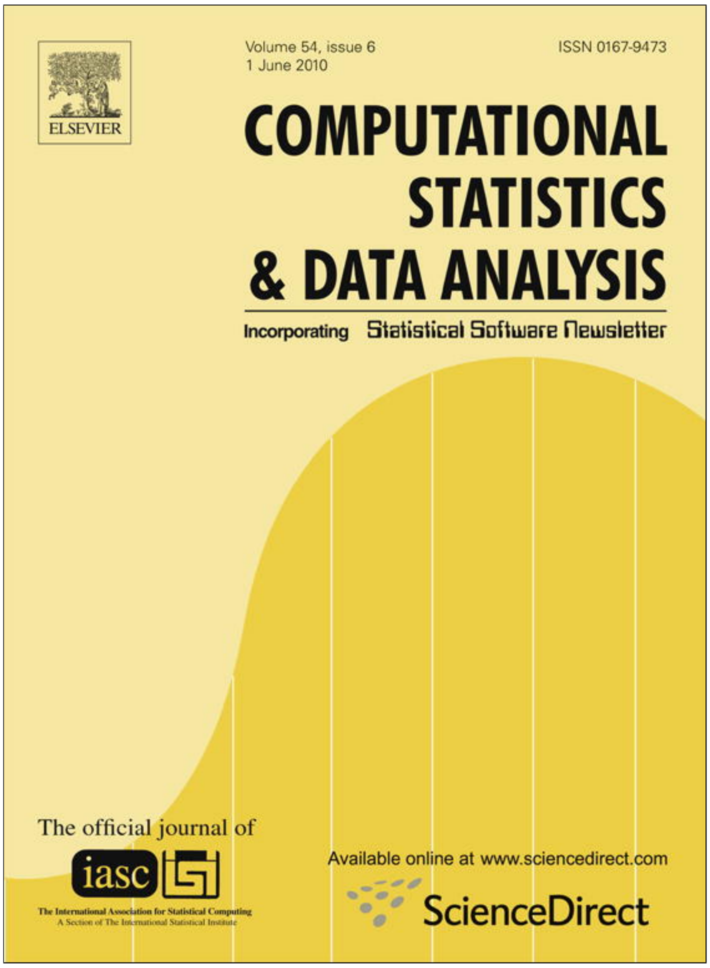
Volume 54, issue 6
1 June 2010
ISSN 0167-9473
ELSEVIER
COMPUTATIONAL STATISTICS & DATA ANALYSIS
Incorporating Statistical Software Newsletter
The official journal of
iasc
The International Association for Statistical Computing
A Section of The International Statistical Institute
Available online at www.sciencedirect.com
ScienceDirect

Computational Statistics and Data Analysis 54 (2010) 1547–1558

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Bayesian inference and prediction of the inverse Weibull distribution for Type-II censored data

Debasis Kundu[a,*], Hatem Howlader[b]

[a] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India
[b] Department of Mathematics and Statistics, University of Winnipeg, Winnipeg, Manitoba, Canada R3B 2E9

ARTICLE INFO

Article history:
Received 2 April 2009
Received in revised form 31 October 2009
Accepted 2 January 2010
Available online 13 January 2010

Keywords:
Bayes estimators
Asymptotic distribution
Type-I censoring
Type-II censoring
Markov Chain Monte Carlo
Posterior predictive density

ABSTRACT

This paper describes the Bayesian inference and prediction of the inverse Weibull distribution for Type-II censored data. First we consider the Bayesian inference of the unknown parameter under a squared error loss function. Although we have discussed mainly the squared error loss function, any other loss function can easily be considered. A Gibbs sampling procedure is used to draw Markov Chain Monte Carlo (MCMC) samples, and they have in turn, been used to compute the Bayes estimates and also to construct the corresponding credible intervals with the help of an importance sampling technique. We have performed a simulation study in order to compare the proposed Bayes estimators with the maximum likelihood estimators. We further consider one-sample and two-sample Bayes prediction problems based on the observed sample and provide appropriate predictive intervals with a given coverage probability. A real life data set is used to illustrate the results derived. Some open problems are indicated for further research.

1. Introduction

The Weibull distribution is one of the most popular distributions in analyzing the lifetime data. Much of the popularity of the Weibull distribution is due to the wide variety of shapes it can assume by varying its parameters. Extensive work has been done on this distribution, both from the frequentist and Bayesian points of view, see for example the excellent review by Johnson et al. (1995) and Kundu (2008) for some recent references. Although, the Weibull distribution has two parameters, in many practical problems, it is not unreasonable to assume one of them, the shape parameter, to be known, and to perform the necessary analysis under the assumption that only the scale parameter is unknown, see for example Nordman and Meeker (2002).

It is well known that the Weibull probability density function (PDF) can be decreasing or unimodal, and the hazard function (HF) can be either decreasing or increasing depending on the shape parameter. Because of the flexibility of the PDF and HF, the Weibull distribution has been used quite extensively when the data indicate a monotone HF. But it cannot be used at all if the data indicate a non-monotone and unimodal HF. In many practical situations, it is often known a priori that the hazard rate cannot be monotone. It may happen that the course of a disease is such that the mortality reaches a peak after some finite period, and then declines slowly. For example, in a study of curability of breast cancer, Langlands et al. (1979) found that the peak mortality occurred after about three years. Bennette (1983) analyzed the data from the Veterans Administration lung cancer trial presented by Prentice (1973) and showed that the empirical failure rates for both low and high-performance status groups were unimodal in nature. It is important to analyze such data sets with the

* Corresponding author.
E-mail address: kundu@iitk.ac.in (D. Kundu).

0167-9473/$ – see front matter © 2010 Published by Elsevier B.V.
doi:10.1016/j.csda.2010.01.003

appropriate models. If the empirical studies indicate that the hazard function might be unimodal, then the inverse Weibull (IW) distribution may be an appropriate model. A brief description of the two-parameter IW distribution is presented in Section 2.

In this paper first we consider the Bayesian inference of the scale parameter for Type-II censored data, under the assumption that the shape parameter is known. This assumption is not unrealistic. See for example, Nelson (1982) for several applications of the inverse Rayleigh distribution, in the reliability and survival analysis, which is a member of the IW distribution when the shape parameter is 2. We assume a gamma prior on the scale parameter and a squared error loss function. Other loss functions can also be easily handled. It is observed that the posterior density function is a mixture of gamma density functions, where the mixture coefficients may or may not be positive. The Bayes estimate can be obtained in explicit form, although the construction of a credible interval is not immediate from the posterior density function. We propose to use a Gibbs sampling procedure to generate MCMC samples and compute the Bayes estimate and construct the credible interval using an importance sampling technique.

We then consider the more important case when both the parameters are unknown. In this case, it is assumed that the scale parameter has the same gamma prior, and the shape parameter also has the gamma prior and they are independently distributed. As expected in this case also, the Bayes estimates cannot be obtained in closed form. We propose to use the Gibbs sampling procedure to generate MCMC samples, and then using the importance sampling methodology, we obtain the Bayes estimates and the HPD credible intervals of the unknown parameters. We perform some simulation experiments to see the behavior of the proposed Bayes estimators and compare their performances with the maximum likelihood estimators (MLEs).

Another important problem in life-testing experiments namely the prediction of unknown observables belonging to a future sample, based on the current available sample, known in the literature as the informative sample. For different application areas and for reference, readers are referred to Al-Hussaini (1999). In this paper we consider the prediction problem in terms of the estimation of the posterior predictive density of a future observation for both one-sample and two-sample schemes. We also construct a predictive interval for a future observation using the Gibbs sampling procedure. An illustrative example has been provided.

The rest of the paper is organized as follows. In Section 2, we provide a brief review of the IW distribution. Prior distribution, posterior analysis and Bayes estimates are provided in Section 3. Monte Carlo simulation results are presented in Section 4. Bayes predictions are provided in Section 5. Data analysis is provided in Section 6, and finally we conclude the paper in Section 7.

2. Inverse Weibull distribution; A brief review

If the random variable Y has a Weibull distribution with the PDF

$$f_Y(y;\alpha,\lambda) = \alpha\lambda y^{\alpha-1}\mathrm{e}^{-\lambda y^{\alpha}}, \quad y > 0, \tag{1}$$

then the random variable $X = \frac{1}{Y}$ has an IW distribution with the PDF

$$f_X(x;\alpha,\lambda) = \alpha\lambda \mathrm{e}^{-\lambda x^{-\alpha}} x^{-(\alpha+1)}, \quad x > 0. \tag{2}$$

The quantities $\alpha > 0$ and $\lambda > 0$ are the shape and scale parameters respectively. From now on it will be denoted by IW(α,λ). If X follows $(\sim)$ IW(α,λ), then the distribution function of X is given by

$$F_X(x;\alpha,\lambda) = \mathrm{e}^{-\lambda x^{-\alpha}}, \quad x > 0. \tag{3}$$

As in the Weibull distribution, the shape parameter α governs the shape of the PDF, the hazard function and the general properties of the IW distribution. As can be evidenced from Fig. 1 both the PDF and HF can be unimodal or decreasing depending on the choice of the shape parameter. In this respect the behavior of the IW distribution and the log-normal distribution is quite similar.

The k-th $(k \le \alpha)$ moment of X is

$$E(X^k) = \lambda^{\frac{k}{\alpha}}\Gamma\left(1-\frac{k}{\alpha}\right) \tag{4}$$

and for $k > \alpha$, the moments do not exist. It is clear that it is a heavy tail distribution and as $\alpha \to \infty$, the tail probability decreases. For $0 < \alpha \le 1$, the mean does not exist, and for $1 < \alpha \le 2$, the mean exists but the variance does not exist.

The IW model has been derived as a suitable model for describing the degradation phenomena of mechanical components, such as the dynamic components of diesel engines, see for example Murthy et al. (2004). The physical failure process given by Erto and Rapone (1984) also leads to the IW model. Erto and Rapone (1984) showed that the IW model provides a good fit to survival data such as the times to breakdown of an insulating fluid subject to the action of constant tension, see also Nelson (1982). Calabria and Pulcini (1994) provided an interpretation of the IW distribution in the context of a load-strength relationship for a component.

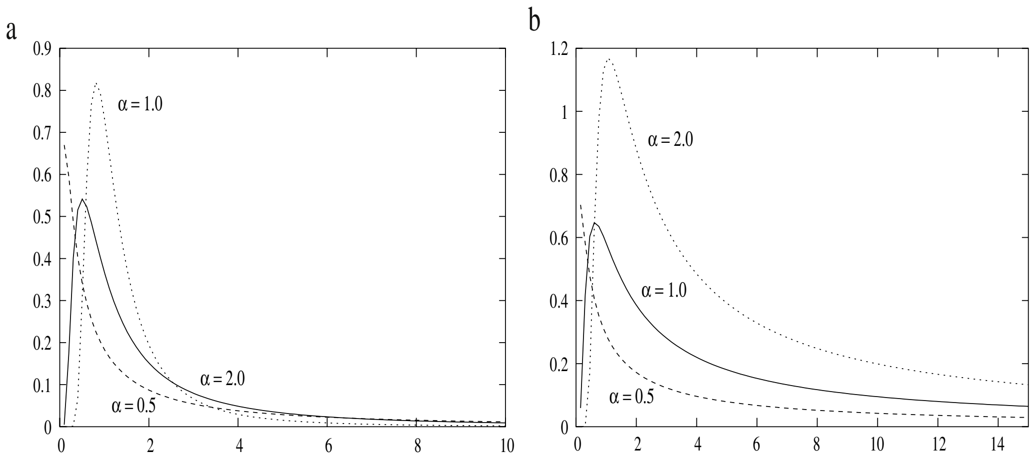

Fig. 1. (a) Probability density functions and (b) hazard functions of the inverse Weibull distribution for different values of α, when the scale parameter is 1.

3. Prior distribution and posterior analysis

In this section we discuss the Bayesian inference of the unknown parameters of the IW distribution, for Type-II censored samples.

Suppose n items are put on a life-testing experiment and we observe only the first r failure times, say $t_{(1)} < \cdots < t_{(r)}$. Under the assumptions that the lifetime distribution of the items are independent and identically distributed (*i.i.d.*) IW(α, λ) random variable, the likelihood function of the observed data without the multiplicative constant can be written as

$$L(\text{data}|\alpha, \lambda) = \alpha^r \lambda^r \mathrm{e}^{-\lambda \sum_{i=1}^{r} x_i^{\alpha}} \prod_{i=1}^{r} x_i^{\alpha+1} \left(1 - \mathrm{e}^{-\lambda x_r^{\alpha}}\right)^{n-r}, \tag{5}$$

here $x_i = t_{(i)}^{-1}$. We consider two cases separately.

3.1. Shape parameter known

In this subsection we provide the Bayes estimate and the associated HPD credible interval of the scale parameter, when the shape parameter is known. It is assumed that the scale parameter has a gamma prior distribution with the shape and scale parameters as c and d, respectively and it has the PDF

$$\pi(\lambda|c, d) \propto \lambda^{c-1} \mathrm{e}^{-d\lambda}, \quad \lambda > 0. \tag{6}$$

It is a natural conjugate prior (NCP). A family of priors is conjugate if the choice of a prior in that family generates a posterior distribution that belongs to the same family. In addition to the ease of mathematical manipulation of the posterior distribution generated from an NCP, the prior parameters can be chosen to suit the prior belief of the experimenter in terms of location and variability of the prior distribution. Moreover, Jeffrey's prior can be obtained as a special case of (6) by substituting $c = d = 0$.

Combining (5) and (6) the posterior density function of λ

$$\pi(\lambda|\text{data}) = \frac{L(\text{data}|\lambda)\pi(\lambda|c, d)}{\int_0^{\infty} L(\text{data}|\lambda)\pi(\lambda|a, b)\mathrm{d}\lambda}, \tag{7}$$

takes the form

$$\pi(\lambda|\text{data}) = \kappa \frac{(b + T_1)^m}{\Gamma(m)} \lambda^{m-1} \mathrm{e}^{-\lambda(b+T_1)} \times \left(1 - \mathrm{e}^{-\lambda T_2}\right)^{n-r}, \tag{8}$$

where $\kappa^{-1} = \frac{(b+T_1)^m}{\Gamma(m)} \int_0^{\infty} \lambda^{m-1} \mathrm{e}^{-\lambda(b+T_1)} \times (1 - \mathrm{e}^{-\lambda T_2})^{n-r} \mathrm{d}\lambda$, and $T_1 = \sum_{i=1}^{r} x_i^{\alpha}$ and $T_2 = x_r^{\alpha}$. Note that both T_1 and T_2 depend on α, but we do not make it explicit for brevity. Using the binomial expansion namely

$$\left(1 - \mathrm{e}^{-\lambda T_2}\right)^{n-r} = \sum_{j=0}^{n-r} (-1)^j \binom{n-r}{j} \mathrm{e}^{-j\lambda T_2},$$

κ^{-1} takes the form

$$\kappa^{-1} = (b+T_1)^m \sum_{j=0}^{n-r} \binom{n-r}{j} \frac{(-1)^j}{d_j^m},$$

with $m = c + r$, and $d_j = d + T_1 + jT_2$. Thus the posterior PDF (8) can be written as

$$\pi(\lambda|\text{data}) = \sum_{j=0}^{n-r} p_j \, \text{gamma}(\lambda; m, d_j), \tag{9}$$

here

$$p_j = (-1)^j \frac{\binom{n-r}{j} \Big/ d_j^m}{\sum_{i=0}^{n-r} (-1)^i \binom{n-r}{i} \Big/ d_i^m}. \tag{10}$$

It is clear that the posterior density function is a mixture of gamma density functions, but the mixture coefficients are not non-negative. From the expression (9) the Bayes estimate of λ with respect to the squared error loss function is

$$\widehat{\lambda}_B = m \sum_{j=0}^{n-r} \frac{p_j}{d_j}. \tag{11}$$

Note that when $n = r$, the Bayes estimate of λ becomes

$$\widehat{\lambda}_B = \frac{c+n}{b+T_1}. \tag{12}$$

Although, the Bayes estimate (11) can be computed explicitly in this case, but the corresponding credible interval cannot be obtained in explicit form from (9). We propose to use the Gibbs sampling technique to generate MCMC samples, and then use an importance sampling technique for constructing the Bayes estimate and also the corresponding credible interval.

Now we provide an algorithm to draw MCMC samples from the posterior distribution (8). Since

$$\pi(\lambda|\text{data}) \le \kappa \, \text{gamma}(\lambda; m, b+T_1), \tag{13}$$

it is possible to use the acceptance rejection method to generate samples from $\pi(\lambda|\text{data})$, by using gamma generation, and we use the Algorithm 1 below to generate a Gibbs sample from the posterior density function of λ.

Algorithm 1.
- Step 1: Generate λ from gamma$(m, b+T_1)$, and U from Uniform $(0, 1)$.
- Step 2: If $U \le (1 - e^{-\lambda T_2})^{n-r}$ then accept λ, otherwise go back to Step 1.
- Step 3: Generate $\lambda_1 \cdots \lambda_M$.
- Step 4: Obtain the Bayes estimate of λ under the squared error loss function as the posterior mean, *i.e.*

$$\widehat{E}(\lambda|\text{data}) = \frac{1}{M} \sum_{i=1}^{M} \lambda_i.$$

- Step 5: Obtain the posterior variance of λ as

$$\widehat{V}(\lambda|\text{data}) = \frac{1}{M} \sum_{i=1}^{M} \left(\lambda_i - \widehat{E}(\lambda|\text{data})\right)^2.$$

- Step 6: To compute the credible interval for λ first we order λ_i's as, $\lambda_{(1)} < \cdots < \lambda_{(M)}$. Denoting $[x]$ as the greatest integer less than or equal to x, the $100(1-\beta)$% symmetric credible interval for λ becomes

$$\left[\lambda_{[(M/100)\times\beta/2]}, \lambda_{[(M/100)\times(1-\beta/2)]}\right].$$

3.2. Both parameters unknown

In this subsection we consider the case when both the parameters are unknown, and we compute the Bayes estimates and the associated HPD credible intervals of the shape and scale parameters. It is assumed that α and λ each have independent gamma(a, b), and gamma(c, d) priors respectively, for $a > 0$, $b > 0$, $c > 0$, $d > 0$, *i.e.*

$$\pi_1(\alpha|a, b) \propto \alpha^{a-1} e^{-b\alpha} \quad \text{and} \quad \pi_2(\lambda|c, d) \propto \lambda^{c-1} e^{-d\lambda}. \tag{14}$$

Based on the priors, the joint density function of the data, α and λ is

$$\pi(\alpha,\lambda|\text{data}) = \frac{L(\text{data}|\alpha,\lambda)\pi_1(\alpha|a,b)\pi_2(\lambda|c,d)}{\int_0^\infty\int_0^\infty L(\text{data}|\alpha,\lambda)\pi_1(\alpha|a,b)\pi_2(\lambda|c,d)\mathrm{d}\alpha\mathrm{d}\lambda}. \tag{15}$$

Therefore, the posterior density function of α and λ given the data can be written as

$$\pi(\alpha,\lambda|\text{data}) \propto g_1(\lambda|\alpha,\text{data})g_2(\alpha|\text{data})h(\alpha,\lambda|\text{data}), \tag{16}$$

here $g_1(\lambda|\alpha,\text{data})$ is a gamma density function with the shape and scale parameters as $r+c$ and $(d+\sum_{i=1}^r x_i^\alpha)$, respectively, $g_2(\alpha|\text{data})$ is a proper density function given by

$$g_2(\alpha|\text{data}) \propto \frac{1}{\left(d+\sum_{i=1}^r x_i^\alpha\right)^{r+c}}\alpha^{a+r-1}\mathrm{e}^{-b\alpha}\prod_{i=1}^r x_i^{\alpha+1}. \tag{17}$$

Moreover

$$h(\alpha,\lambda|\text{data}) = \left(1-\mathrm{e}^{-\lambda x_r^\alpha}\right)^{n-r}. \tag{18}$$

Therefore, the Bayes estimate of any function of α and λ, say $g(\alpha,\lambda)$ under the squared error loss function is

$$\widehat{g}_B(\alpha,\lambda) = \frac{\int_0^\infty\int_0^\infty g(\alpha,\lambda)g_1(\lambda|\alpha,\text{data})g_2(\alpha|\text{data})h(\alpha,\lambda|\text{data})\mathrm{d}\alpha\mathrm{d}\lambda}{\int_0^\infty\int_0^\infty g_1(\lambda|\alpha,\text{data})g_2(\alpha|\text{data})h(\alpha,\lambda|\text{data})\mathrm{d}\alpha\mathrm{d}\lambda}. \tag{19}$$

It is not possible to compute (19) analytically but Lindley's (1980) approximation may be used to compute the ratio of integrals of the form (19). It cannot however be used to construct credible intervals. We therefore, do not use it, instead, we propose to approximate (19) by using an importance sampling technique as suggested by Chen and Shao (1999) and also construct the corresponding credible intervals. The details are explained below.

3.3. Importance sampling

We need the following theorem for further development.

Theorem 1. *$g_2(\alpha|\text{data})$ as given in (17) has a log-concave density function.*

Proof. See Appendix. ■

Since $g_2(\alpha|\text{data})$ has a log-concave density, using the idea of Devroye (1984) it is possible to generate a sample from $g_2(\alpha|\text{data})$. Moreover, since $g_1(\lambda|\alpha,\text{data})$ follows gamma, it is quite simple to generate from $g_1(\lambda|\alpha,\text{data})$. Now we would like to provide the importance sampling procedure to compute the Bayes estimates and also to construct the credible interval of $g(\alpha,\beta)=\theta$ (say). Using Theorem 1, a simulation based consistent estimate of $E(g(\alpha,\lambda))=E(\theta)$ can be obtained using Algorithm 2 as given below:

Algorithm 2.
- Step 1: Generate α from $g_2(\cdot|\lambda)$ using the method developed by Devroye (1984).
- Step 2: Generate λ from $g_1(\cdot|\alpha,\text{data})$.
- Step 3: Repeat Step 1 and Step 2 and obtain $(\alpha_1,\lambda_1),\ldots,(\alpha_M,\lambda_M)$.
- Step 4: An approximate Bayes estimate of θ under a squared error loss function can be obtained as

$$\widehat{g}_B(\alpha,\lambda) = \widehat{\theta} = \frac{\frac{1}{M}\sum_{i=1}^M \theta_i h(\alpha_i,\lambda_i|\text{data})}{\frac{1}{M}\sum_{i=1}^M h(\alpha_i,\lambda_i|\text{data})}.$$

- Step 5: Obtain the posterior variance of $\theta=g(\alpha,\lambda)$ as

$$\widehat{V}(g(\alpha,\lambda|\text{data})) = \frac{\frac{1}{M}\sum_{i=1}^M \left(\theta_i-\widehat{\theta}\right)^2 h(\alpha_i,\lambda_i|\text{data})}{\frac{1}{M}\sum_{i=1}^M h(\alpha_i,\lambda_i|\text{data})}.$$

We now obtain the credible interval of θ using the idea of Chen and Shao (1999). Let us denote $\pi(\theta|\text{data})$ and $\Pi(\theta|\text{data})$ as the posterior density and posterior distribution functions of θ, respectively. Also let $\theta^{(\beta)}$ be the β-th quantile of θ, *i.e.*

$$\theta^{(\beta)} = \inf\{\theta; \Pi(\theta|\text{data}) \geq \beta\}, \quad 0 < \beta < 1. \tag{20}$$

Observe that for a given θ^*,

$$\Pi(\theta^*|\text{data}) = E\left[1_{\theta \leq \theta^*}|\text{data}\right],$$

where $1_{\theta \leq \theta^*}$ is the indicator function defined as

$$1_{\theta \leq \theta^*} = \begin{cases} 1 & \text{if } \theta \leq \theta^* \\ 0 & \text{if } \theta > \theta^*. \end{cases}$$

Therefore, a simulation consistent estimator of $\Pi(\theta^*|\text{Data})$ can be obtained as

$$\widehat{\Pi}(\theta^*|\text{data}) = \frac{\frac{1}{M}\sum_{i=1}^{M} 1_{\theta \leq \theta^*} h(\alpha_i, \lambda_i|\text{data})}{\frac{1}{M}\sum_{i=1}^{M} h(\alpha_i, \lambda_i|\text{data})}.$$

For $i = 1, \ldots M$, let $\{\theta_{(i)}\}$ be the ordered values of θ_i, and

$$w_{(i)} = \frac{h(\alpha_{(i)}, \lambda_{(i)}|\text{data})}{\sum_{i=1}^{M} h(\alpha_i, \lambda_i|\text{data})}$$

be the associated weight, then we have

$$\widehat{\Pi}(\theta^*|\text{data}) = \begin{cases} 0 & \text{if } \theta^* < \theta_{(1)} \\ \sum_{j=1}^{i} w_{(j)} & \text{if } \theta_{(i)} \leq \theta^* < \theta_{(i+1)} \\ 1 & \text{if } \theta^* > \theta_{(M)}. \end{cases}$$

As before the credible interval for θ, can easily be constructed.

4. Monte Carlo simulations

In order to compare the proposed Bayes estimators with the MLEs, we perform a Monte Carlo Simulation study using different sample sizes (n), different effective sample sizes (r), and for different priors (non-informative and informative). All computations were performed using an Intel dual core processor. For random number generation we have used RAN2 of Press et al. (1991). The programs were written in FORTRAN 77. In computing the estimates we generated 1000 samples from the IW distribution with $\alpha = 2$ and $\lambda = 1$, and we replicated the process 1000 times. The averages and mean squared errors (MSE) in parentheses of estimators of α and λ are presented in Tables 1 and 2, respectively. For prior information we have used: Non-informative prior, Prior 1 with $a = b = c = d = 0$, and informative prior, Prior 2 with $a = 2$, $b = 1$, $c = d = 1$. For Prior 2 we have chosen the hyper-parameters in such a way that the prior mean became the expected value of the corresponding parameter.

It is clear from Tables 1 and 2 that the proposed Bayes estimators perform very well for different n and r. As expected, the performance in terms of average bias and the MSE of the Bayes estimators under Prior 1 and the MLE is very similar. The Bayes estimators under Prior 2 clearly outperform the MLEs in terms of average bias and MSE.

5. Bayes prediction

The Bayes prediction of an unknown observable, such as the value of a future sample based on the current sample, known as the informative sample is an important problem. Al-Hussaini (1999) provided a number of references on the applications of Bayes prediction in different areas of applied statistics.

There are two main types of prediction problem as indicated by Al-Hussaini (1999), namely (a) One-Sample Prediction and (b) Two-Sample Prediction. They can be briefly described as follows. Let $T_{(1)} < \cdots < T_{(r)}$ and $T_{(r+1)} < \cdots < T_{(n)}$ represent the informative sample and a future sample, respectively. A one-sample prediction problem involves the prediction and associated inference of the future order statistics $T_{(k)}$ for $r < k \leq n$. On the other hand, let $T_{(1)} < \cdots < T_{(r)}$ and $Y_{(1)} < \cdots < Y_{(m)}$ represent the informative sample from a random sample of size n, and a future ordered sample of size m, respectively. It is further assumed that the two samples are independent and each of their corresponding random samples is obtained from the same distribution function. Then a two-sample prediction problem involves the prediction and associated inference of the order statistics $Y_{(1)} < \cdots < Y_{(m)}$ of a future sample from the same distribution function.

The aim of this section is to provide the Bayes prediction of the k-th observation, $r < k \leq n$ for a one-sample prediction problem; and $1 \leq k \leq m$ for a two-sample prediction problem, and associated inference based on the available data, namely $t_{(1)} < \cdots < t_{(r)}$. Specifically, we wish to provide an estimate of the posterior density function of $T_{(k)}$ given the data, and also construct a $100(1-\gamma)\%$ predictive interval of $T_{(k)}$. We consider these two cases separately.

Table 1
Average estimates of α and the associated MSEs.

n	r	Bayes		MLE
		Prior 1	Prior 2	
20	10	1.9311 (0.0217)	2.0954 (0.0121)	1.9269 (0.0219)
20	15	1.9318 (0.0214)	2.0876 (0.0118)	1.9287 (0.0217)
25	15	1.9200 (0.0229)	2.1014 (0.0143)	1.9209 (0.0232)
25	20	1.9198 (0.0226)	2.0991 (0.0137)	1.9196 (0.0224)
40	15	1.9341 (0.0179)	2.0913 (0.0112)	1.9350 (0.0181)
40	20	1.9329 (0.0173)	2.0899 (0.0098)	1.9322 (0.0176)
40	30	1.9322 (0.0159)	2.0816 (0.0087)	1.9328 (0.0163)
40	35	1.9314 (0.0149)	2.0818 (0.0079)	1.9322 (0.0145)

Table 2
Average estimates of λ and the associated MSEs.

n	r	Bayes		MLE
		Prior 1	Prior 2	
20	10	0.9329 (0.0137)	1.0215 (0.0093)	0.9319 (0.0134)
20	15	0.9376 (0.0119)	1.0210 (0.0081)	0.9381 (0.0121)
25	15	0.9399 (0.0111)	1.0178 (0.0073)	0.9434 (0.0109)
25	20	0.9398 (0.0109)	1.0156 (0.0035)	0.9427 (0.0104)
40	15	0.9499 (0.0075)	1.0131 (0.0032)	0.9544 (0.0071)
40	20	0.9539 (0.0065)	1.0126 (0.0027)	0.9545 (0.0068)
40	30	0.9578 (0.0061)	1.0119 (0.0019)	0.9544 (0.0067)
40	35	0.9579 (0.0059)	1.0101 (0.0012)	0.9545 (0.0066)

5.1. One Sample Prediction

In this case we are interested in the posterior predictive density of $T_{(k)}$ given the data, which is

$$\pi_{T_{(k)}}(y|\text{data}) = \int_0^\infty \int_0^\infty f_{T_{(k)}|\text{Data}}(y|\alpha,\lambda)\pi(\alpha,\lambda|\text{data})\mathrm{d}\alpha\mathrm{d}\lambda, \quad y > t_{(r)}. \tag{21}$$

Here $f_{T_{(k)}|\text{Data}}(\cdot|\alpha,\lambda)$ is the conditional density of $T_{(k)}$ given $t_{(1)} < \cdots < t_{(r)}$. See for example Chen et al. (2000). Because of the Markov property of the conditional order statistics,

$$f_{T_{(k)}|\text{Data}}(y|\alpha,\lambda) = f_{T_{(k)}|T_{(r)}=t_{(r)}}(y|\alpha,\lambda), \quad y > t_{(r)}. \tag{22}$$

For notational simplicity, let us call $k = r + 1$. Therefore,

$$\begin{aligned} f_{T_{(r+1)}|T_{(r)}=t_{(r)}}(y|\alpha,\lambda) &= \frac{(n-r)f(y|\alpha,\lambda)(1-F(y|\alpha,\lambda))^{n-r+1}}{(1-F(t_{(r)}|\alpha,\lambda))^{n-r}}, \quad y > t_{(r)} \\ &= \frac{(n-r)\alpha\lambda \mathrm{e}^{-\lambda y^{-\alpha}} y^{-(\alpha+1)}\left(1-\mathrm{e}^{-\lambda y^{-\alpha}}\right)^{n-r-1}}{\left(1-\mathrm{e}^{-\lambda t_{(r)}^{-\alpha}}\right)^{n-r}}, \quad y > t_{(r)}. \end{aligned} \tag{23}$$

The predictive density of $T_{(r+1)}$ is then

$$f^*_{T_{(r+1)}|\text{Data}}(y) = \int_0^\infty \int_0^\infty f_{T_{(r+1)}|T_{(r)}=t_{(r)}}(y|\alpha,\lambda)\pi(\alpha,\lambda|\text{Data})\,\mathrm{d}\alpha\mathrm{d}\lambda, \tag{24}$$

and the predictive survival function is

$$S^*_{T_{(r+1)}|\text{Data}}(y) = \int_0^\infty \int_0^\infty S_{T_{(r+1)}|T_{(r)}=t_{(r)}}(y|\alpha,\lambda)\pi(\alpha,\lambda|\text{Data})\,\mathrm{d}\alpha\mathrm{d}\lambda. \tag{25}$$

Here

$$S_{T_{(r+1)}|T_{(r)}=t_{(r)}}(y|\alpha,\lambda) = \frac{\left(1-\mathrm{e}^{-\lambda y^{-\alpha}}\right)^{n-r}}{\left(1-\mathrm{e}^{-\lambda t_{(r)}^{-\alpha}}\right)^{n-r}}, \quad y > t_{(r)}. \tag{26}$$

A simulation based consistent estimator of $f^*_{T_{(r+1)}|\text{Data}}(y)$ and $S^*_{T_{(r+1)}|\text{Data}}(y)$ can be obtained by using the Gibbs sampling procedure as described in Section 3. Suppose $\{(\alpha_i,\lambda_i);\, i = 1\cdots M\}$ are MCMC samples obtained from $\pi(\alpha,\lambda|\text{Data})$, using the Gibbs sampling technique, the simulation consistent estimators of $f^*_{T_{(r+1)}|\text{Data}}(y)$ and $S^*_{T_{(r+1)}|\text{Data}}(y)$ can be obtained as

$$\widehat{f}^*_{T_{(r+1)}|\text{Data}}(y) = \sum_{i=1}^{M} f_{T_{(r+1)}|T_{(r)}=t_{(r)}}(y|\alpha_i,\lambda_i)w_i \tag{27}$$

and

$$\widehat{S}^*_{T_{(r+1)}|\text{Data}}(y) = \sum_{i=1}^{M} S_{T_{(r+1)}|T_{(r)}=t_{(r)}}(y|\alpha_i,\lambda_i)w_i \tag{28}$$

respectively, where

$$w_i = \frac{h(\alpha_i,\lambda_i|\text{data})}{\sum_{i=1}^{M} h(\alpha_i,\lambda_i|\text{data})}; \quad i = 1,\ldots,M. \tag{29}$$

Another important aspect of prediction is to construct a two-sided predictive interval for $T_{(r+1)}$. A symmetric $100\gamma\%$ predictive interval of $T_{(r+1)}$ can be obtained by solving the non-linear equations (30) and (31) simultaneously for the lower bound, L and upper bound, U:

$$\frac{1+\gamma}{2} = P(T_{(r+1)} > L|\text{Data}) \Rightarrow S^*_{T_{(r+1)}|\text{Data}}(L) = \frac{1+\gamma}{2} \tag{30}$$

and

$$\frac{1-\gamma}{2} = P(T_{(r+1)} > U|\text{Data}) \Rightarrow S^*_{T_{(r+1)}|\text{Data}}(U) = \frac{1-\gamma}{2}. \tag{31}$$

We need to apply a suitable numerical method as they cannot be solved analytically.

5.2. Two Sample Prediction

Let us consider a future sample $\{Y_1,\ldots,Y_m\}$ of size m, independent of the informative sample $\{X_1,\ldots,X_n\}$ and let $Y_{(1)} < \cdots < Y_{(r)} < \cdots Y_{(m)}$ be the order statistics of the future sample. Suppose we are interested in the predictive density of the order statistic $Y_{(k)}$ of the future sample, given the informative data set $\{x_1,\ldots,x_n\}$. The probability density function of the k th order statistic of the future sample is given by

$$g_{(k)}(y|\alpha,\lambda) = \frac{m!}{(k-1)!(m-k)!}\,[F(y|\alpha,\lambda)]^{k-1}\,[1-F(y|\alpha,\lambda)]^{m-k} f(y|\alpha,\lambda), \tag{32}$$

here $f(\cdot|\alpha,\lambda)$ is as given in (2) and $F(\cdot|\alpha,\lambda)$ denotes the corresponding cumulative distribution function of $f(\cdot|\alpha,\lambda)$, as given in (3). If we denote the the predictive density of $Y_{(k)}$ as $g^*_{(k)}(\cdot|\text{data})$, then

$$g^*_{(k)}(y|\text{data}) = \int_0^\infty \int_0^\infty g_{(k)}(y|\alpha,\lambda)\pi(\alpha,\lambda|\text{data})\mathrm{d}\alpha\mathrm{d}\lambda, \tag{33}$$

where $\pi(\alpha,\lambda|\text{data})$ is the joint posterior density of α and λ as given in (16). It is immediate that $g^*_{(k)}(y|\text{data})$ cannot be expressed in closed form and hence it cannot be evaluated analytically.

As before, based on the MCMC samples $\{(\alpha_i, \lambda_i), i = 1, \ldots, M\}$, a simulation consistent estimator of $g^*_{(k)}(y|\text{data})$, can be obtained as

$$\widehat{g}^*_{(k)}(y|\text{data}) = \sum_{i=1}^{M} g_{(k)}(y|\alpha_i, \lambda_i) w_i \tag{34}$$

and a simulation consistent estimator of the predictive distribution of $Y_{(k)}$, say $G^*_{(k)}(\cdot|\text{data})$ can be obtained as

$$\widehat{G}^*_{(k)}(y|\text{data}) = \sum_{i=1}^{M} G_{(k)}(y|\alpha_i, \lambda_i) w_i \tag{35}$$

w_i is same as defined in (29) and $G_{(k)}(y|\alpha, \lambda)$ denotes the distribution function corresponding to the density function $g_{(k)}(y|\alpha, \lambda)$, *i.e.*

$$\begin{aligned} G_{(k)}(y|\alpha, \lambda) &= \frac{m!}{(k-1)!(m-k)!} \int_0^y [F(z|\alpha, \lambda)]^{k-1} [1 - F(z|\alpha, \lambda)]^{m-k} f(z|\alpha, \lambda) \mathrm{d}z \\ &= \frac{m!}{(k-1)!(m-k)!} \int_0^{F(y|\alpha,\lambda)} u^{k-1}(1-u)^{m-k} \mathrm{d}u. \end{aligned} \tag{36}$$

It should be noted that the same MCMC samples $\{(\alpha_i, \lambda_i), i = 1, \ldots, M\}$ can be used to compute $\widehat{g}^*_{(k)}(y|\text{data})$ or $\widehat{G}^*_{(k)}(y|\text{data})$ for all y. Moreover, a symmetric $100\gamma\%$ predictive interval for $Y_{(r)}$ can be obtained by solving the non-linear equations (37) and (38), for the lower bound, L and upper bound, U:

$$\frac{1+\gamma}{2} = P\left[Y_{(k)} > L|\text{data}\right] = 1 - G^*_{(k)}(L|\text{data}) \Rightarrow G^*_{(k)}(L|\text{data}) = \frac{1}{2} - \frac{\gamma}{2} \tag{37}$$

$$\frac{1-\gamma}{2} = P\left[Y_{(k)} > U|\text{data}\right] = 1 - G^*_{(k)}(U|\text{data}), \Rightarrow G^*_{(k)}(U|\text{data}) = \frac{1}{2} + \frac{\gamma}{2}. \tag{38}$$

In this case it is also not possible to obtain the solutions analytically, and one needs a suitable numerical technique for solving these non-linear equations.

6. Illustrative example

In this section we consider a real life data set and illustrate the methods proposed in the previous sections. The data set is from Bjerkedal (1960), and it represents the survival times (in days) of guinea pigs injected with different doses of tubercle bacilli. It is known that guinea pigs have a high susceptibility to human tuberculosis and that is why they were used in this particular study. The regimen number is the common logarithm of the number of bacillary units in 0.5 ml. of challenge solution; i.e., regimen 6.6 corresponds to 4.0×10^6 bacillary units per 0.5 ml. ($\log(4.0 \times 10^6) = 6.6$). Corresponding to regimen 6.6, there were 72 observations listed below:
12, 15, 22, 24, 24, 32, 32, 33, 34, 38, 38, 43, 44, 48, 52, 53, 54, 54, 55, 56, 57, 58, 58, 59, 60, 60, 60, 60, 61, 62, 63, 65, 65, 67, 68, 70, 70, 72, 73, 75, 76, 76, 81, 83, 84, 85, 87, 91, 95, 96, 98, 99, 109, 110, 121, 127, 129, 131, 143, 146, 146, 175, 175, 211, 233, 258, 258, 263, 297, 341, 341, 376.

The mean, standard deviation and the coefficient of skewness are calculated as 99.82, 80.55 and 1.80, respectively. The measure of skewness indicates that the data are positively skewed. For computational ease, we have divided each data point by 1000.

Before progressing further we wish to examine the empirical hazard function of the observed data by applying the scaled Total Time on Test (TTT) plot, see Aarset (1987). This provides a very good idea about the shape of the hazard function of a distribution. For a family with the survival function $S(y) = 1 - F(y)$, the scaled TTT transform, with $H^{-1}(u) = \int_0^{F^{-1}(u)} S(y)\mathrm{d}y$ defined for $0 < u < 1$ is $g(u) = H^{-1}(u)/H^{-1}(1)$. The corresponding empirical version of the scaled TTT transform is given by $g_n(r/n) = H_n^{-1}(r/n)/H_n^{-1}(1) = [\sum_{i=1}^{r} y_{(i)} + (n-r)y_{(r)}]/(\sum_{i=1}^{n} y_{(i)})$, where $y_{(i)}$ denotes the i-th order statistic of the sample. It has been shown by Aarset (1987) that the TTT transform is convex (concave) if the hazard rate is decreasing (increasing); and for bathtub (unimodal) hazard rates, the scaled TTT transform is first convex (concave) and then concave (convex). The plot of the scaled TTT transform of the data, Fig. 2, indicates that the empirical hazard function is unimodal and therefore, it is reasonable to use an IW distribution to analyze the data.

We also wanted to check by using the Kolmogorov–Smirnov (K–S) statistic whether the IW model is suitable for this data. The maximum likelihood estimates of α and λ based on the complete sample are 1.4142 and 0.0169, respectively. The Bayes estimates of α and λ based on a complete sample and for a non-informative prior, *i.e.* $a = b = c = d = 0$, are 1.4086 and 0.0176, respectively. As expected the Bayes estimates under the non-informative prior, and the MLE are quite close to each other. In case of MLE, the K–S distance and the associated p-value are 0.1364 and 0.137, respectively, and for the Bayes estimates the corresponding values are 0.1277 and 0.191. Based on the p-values, the IW is found to fit the data very well.

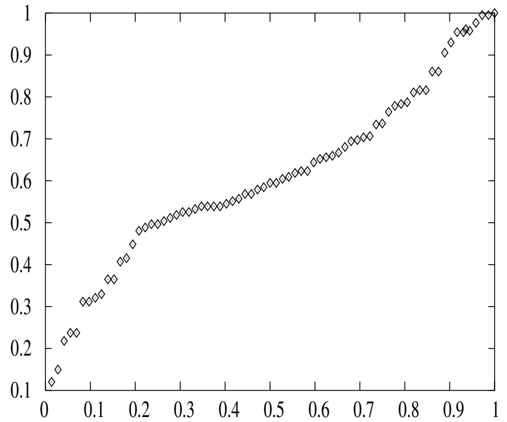

Fig. 2. Scaled TTT transform of the Guinea pigs data.

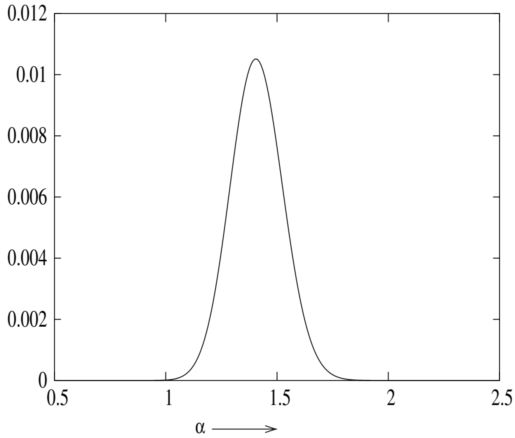

Fig. 3. Posterior density function of α.

Now we consider the case when the data are Type-II censored. It is assumed that we observe only the first 50 data points and the rest are censored. It is assumed that both the parameters are unknown. Since we do not have any prior information available, we use non-informative priors on both α and λ. The density function of $g_2(\alpha|\text{data})$ as given in (17) is plotted, Fig. 3. It is clearly log-concave. It can also be approximated by the normal distribution function, but we will not attempt the approach here. Now using Algorithm 2, we generate 10,000 MCMC samples and based on them we compute the Bayes estimates of α and λ as 1.4623 and 0.0137, respectively. Moreover, the 95% HPD credible intervals of α and λ are (1.3246, 1.6472), and (0.0071, 0.0189), respectively.

Now we consider the one sample prediction problem. In Fig. 4 we present the predictive density function and the predictive survival function of the 51-st order statistic based on the observed sample. The 95% predictive interval of the 51-st order statistic is (0.096, 0.107). Therefore, based on the observed sample the 51-st failure will occur between 96 and 107 days.

We now consider the two-sample prediction problem. Suppose we put 25 new guinea pigs on the same test, and we wish to find the predictive density and the predictive interval of the median of the future sample, based on the observed sample. The predictive density function and the distribution function of the median are plotted in Fig. 5. The two sided 95% predictive interval of the median is (0.0452, 0.1075). Therefore, based on the observed sample the median failure will occur between 45 and 108 days.

7. Conclusions

In this paper we have considered the Bayesian inference and prediction problems of the inverted Weibull distribution based on Type-II censored data. Since the moments of the inverse Weibull model do not always exist, the Bayesian inference seems to be the natural choice for the analysis and prediction of certain survival data. The prior belief of the model is represented by the independent gamma priors on the shape and scale parameters. The squared error loss function is used as it is appropriate when large errors of the estimation are considered to be more serious compared to other loss functions. It is observed even when the shape parameter is known and the HPD credible intervals cannot be obtained in explicit form. We used the Gibbs sampling technique to generate MCMC samples and then using importance sampling methodology we

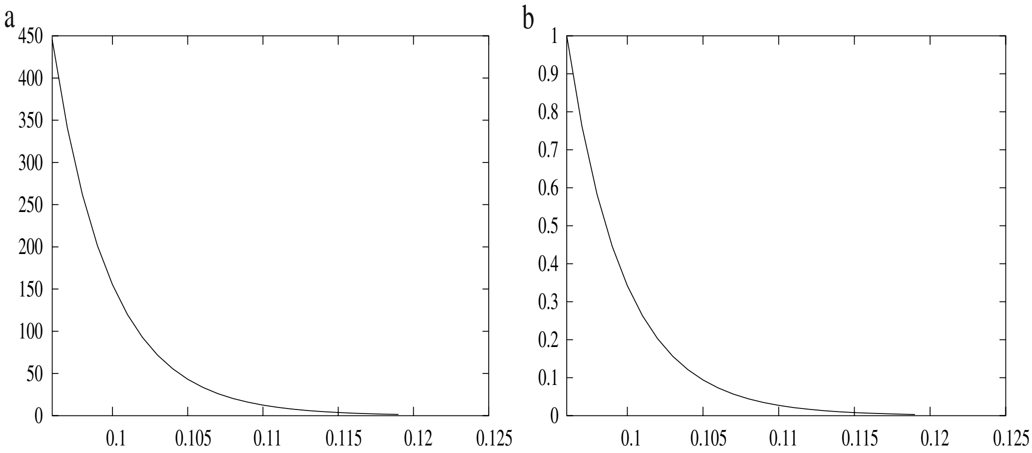

Fig. 4. (a) Predictive density function and (b) Predictive survival function of the 51-st order statistics.

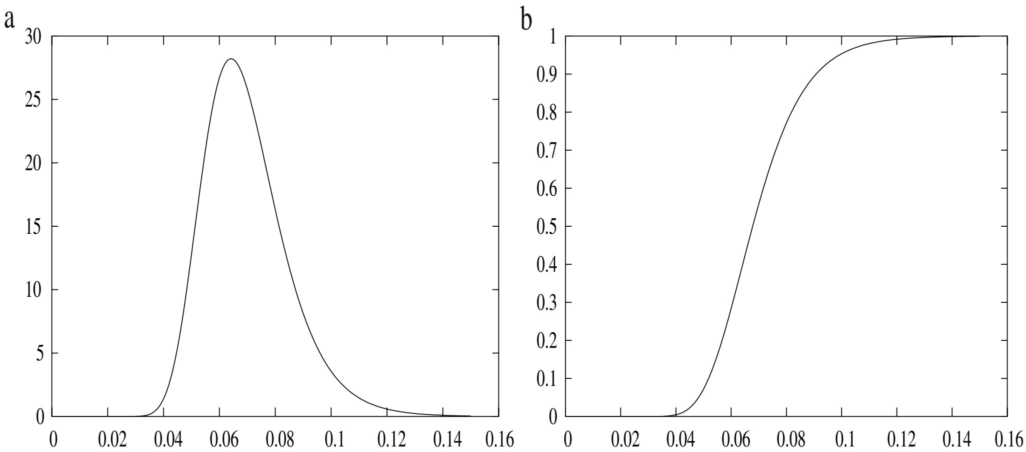

Fig. 5. (a) Predictive density function and (b) Predictive distribution function of the median of a future sample of size 25.

computed the Bayes estimates and constructed the HPD credible intervals. The same MCMC samples were used for one sample and two sample prediction problems. The details have been explained using a real life example.

An important problem will be to extend these results for other censoring schemes such as Type-I, hybrid and progressive censoring schemes. The work is in progress.

Acknowledgements

The authors would like to thank two referees for their suggestions which had improved significantly the earlier draft of the manuscript. Part of the work of the first author has been supported by a grant from the Department of Science and Technology, Government of India.

Appendix

Proof of Theorem 1. Since

$$g_2(\alpha|\text{data}) \propto \alpha^{a+r-1}\mathrm{e}^{-b\alpha}\prod_{i=1}^{r} x_i^{\alpha+1} \times \frac{1}{\left(d+\sum_{i=1}^{r} x_i^{\alpha}\right)^{r+c}}. \tag{39}$$

$\ln g_2(\alpha|\text{data})$ without the additive constant is

$$\ln g_2(\alpha|\text{data}) = (a+r-1)\ln\alpha - b\alpha + (\alpha+1)\sum_{i=1}^{r}\ln x_i - (r+c)\ln\left(d+\sum_{i=1}^{r} x_i^{\alpha}\right). \tag{40}$$

Suppose

$$u(\alpha) = d + \sum_{i=1}^{r} x_i^{\alpha} \Rightarrow u'(\alpha) = \sum_{i=1}^{r} x_i^{\alpha} \ln x_i \Rightarrow u''(\alpha) = \sum_{i=1}^{r} x_i^{\alpha} (\ln x_i)^2.$$

Observe that

$$\left(\sum_{i=1}^{r} x_i^{\alpha} (\ln x_i)^2\right)\left(\sum_{i=1}^{r} x_i^{\alpha}\right) - \left(\sum_{i=1}^{r} x_i^{\alpha} (\ln x_i)\right)^2 = \sum_{1 \le i < j \le r} x_i^{\alpha} x_j^{\alpha} \left(\ln x_i - \ln x_j\right)^2 \ge 0.$$

Therefore, for all $d \ge 0$,

$$u''(\alpha)u(\alpha) \ge (u'(\alpha))^2.$$

It implies

$$\frac{\mathrm{d}^2}{\mathrm{d}\alpha^2} \ln g_2(\alpha|\text{data}) < 0,$$

and that proves the theorem. ■

References

Aarset, M.V., 1987. How to identify a bathtub hazard function. IEEE Transactions on Reliability 36, 106–108.
Al-Hussaini, E.K., 1999. Predicting observables from a general class of distributions. Journal of Statistical Planning and Inference 79, 79–91.
Bennette, S., 1983. Log-logistic regression models for survival data. Applied Statistics 32, 165–171.
Bjerkedal, T., 1960. Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. American Journal of Hygiene 72, 130–148.
Calabria, R., Pulcini, G., 1994. Bayesian 2-sample prediction for the inverse Weibull distribution. Communications in Statistics—Theory and Methods 23, 1811–1824.
Chen, M.-H., Shao, Q.-M., 1999. Monte Carlo estimation of Bayesian credible and HPD intervals. Journal of Computational and Graphical Statistics 8 (1), 69–92.
Chen, M.-H., Shao, Q.-M., Ibrahim, J.G., 2000. Monte Carlo Methods in Bayesian Computation. Springer-Verlag, New York.
Devroye, L., 1984. A simple algorithm for generating random variates with a log-concave density function. Computing 33, 247–257.
Erto, P., Rapone, M., 1984. Non-informative and practical Bayesian confidence bounds for reliable life in the Weibull model. Reliability Engineering 7, 181–191.
Johnson, R., Kotz, S., Balakrishnan, N., 1995. Continuous Univariate Distribution, 2nd ed.. Wiley and Sons, New York.
Kundu, D., 2008. Bayesian inference and reliability sampling plan for Weibull distribution. Technometrics 50 (2), 144–154.
Langlands, A.O., Pocock, S.J., Kerr, G.R., Gore, S.M., 1979. Long term survival of patients with breast cancer: A study of curability of the disease. British Medical Journal 1247–1251.
Lindley, D.V., 1980. Approximate Bayes method. Trabajos de Estadistica 31, 223–237.
Murthy, D.N.P., Xie, M., Jiang, R., 2004. Weibull Models. Wiley, New York.
Nelson, W., 1982. Applied Lifetime Data Analysis. Wiley, New York.
Nordman, D., Meeker, W.Q., 2002. Weibull prediction intervals for a future number of failures. Technometrics 44, 15–23.
Prentice, R.L., 1973. Exponential survivals with censoring and explanatory variables. Biometrika 60, 279–288.
Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 1991. Numerical Recipes: The Art of Scientific Computing. Cambridge University Press, UK.