Chapter 1
On Bayesian Inference of $P(Y < X)$ for Weibull Distribution

Abstract In this paper we consider the Bayesian inference on the stress strength parameter $R = P(Y < X)$, when X and Y follow independent Weibull distributions. We have considered different cases. It is assumed that the random variables X and Y have different scale parameters and (a) a common shape parameter or (b) different shape parameters. Moreover, both stress and strength may depend on some known covariates also. When the two distributions have a common shape parameter, Bayesian inference on R is obtained based on the assumption that the shape parameter has a log-concave prior, and given the shape parameter, the scale parameters have Dirichlet-Gamma prior. The Bayes estimate cannot be obtained in closed form, and we propose to use Gibbs sampling method to compute the Bayes estimate and also to compute the associated highest posterior density (HPD) credible interval. The results have been extended when the covariates are also present. We further consider the case when the two shape parameters are different. Simulation experiments have been performed to see the effectiveness of the proposed methods. One data set has been analyzed for illustrative purposes and finally we conclude the paper.

1.1 Introduction

The problem of estimating the stress-strength parameter $R = P(Y < X)$ and its associated inference, when X and Y are two independent random variables, have received considerable attention since the pioneering work of Birnbaum (1956). In spite of its apparent simplicity, extensive work has been done since then both from the parametric and non-parametric points of view. The estimation of R occurs quite naturally in the statistical reliability literature. For example, in a reliability study, if X is the strength of a system, which is subject to stress Y, then R is the measure of system performance and arises quite frequently in the mechanical reliability of a system. Alternatively, in survival analysis, the area under the receiving operating characteristic (ROC) curve, which is a plot of the sensitivity versus, 1 - specificity, at different cut-off points of the range of possible test values, is equal to R. A book length treatment on estimation of R, and its associated inference can be obtained in Kotz et al. (2003). See also Zhou (2008), Kundu and Raqab (2009), Ventura and Racugno (2011), Kizilaslan and Nadar (2016) and the references cited therein, for recent developments.

Although, extensive work has been done on this particular problem under different assumptions on X and Y, not much attention has been paid when both X and Y depend on some covariates. Guttman et al. (1988) and Weerahandi and Johnson (1992) considered the estimation of R, and also obtained the associated confidence interval of R, when both stress and strength depend on some known covariates. It has been assumed in both the papers that both X and Y are normally distributed, and their respective means depend on the covariates, but their variances do not depend on the covariates. Guttman and Papandonatos (1997) also considered the same model from the Bayesian point of view. They proposed different approximate Bayes estimates under non-informative priors. Interestingly, although it has been mentioned by Guttman et al. (1988) that it is important to develop statistical inferential procedure for the stress-strength parameter in case of covariates for non-normal distributions mainly for small sample sizes, nothing has been developed along that line till date. This is an attempt towards that direction.

The aim of this paper is to consider the inference on R, when both X and Y are independent Weibull random variables. They may or may not have the same shape parameters. It is assumed that both stress and strength distributions may depend on some known covariates. Here, we are proposing the covariates modelling in the following manner. It is assumed that for both X and Y, the scale parameter depends on some known covariates. Since the lifetime distributions are always non-negative, it is more natural to use Weibull, gamma, log-normal etc. rather than normal distribution for this purpose. Moreover, using the covariates through the scale parameter also ensures that both the mean and variance of the lifetime of an item/ individual depend on the covariates. In case of normal distribution, the covariates are modelled through mean only, see for example Guttman et al. (1988), Weerahandi and Johnson (1992) or Guttman and Papandonatos (1997), which may not be very reasonable.

Another important point should be mentioned that the statistical inference on $R = P(Y < X)$, when X and Y have independent Weibull random variables have been considered by many authors, see for example Kotz et al. (2003), Kundu and Gupta (2006), Kundu and Raqab (2009), and see the references cited therein. But in all these cases it has been assumed that X and Y have a common shape parameter. It has been assumed mainly due to analytical reason, although it may not be reasonable in many situation. In this paper we have also considered the case when X and Y may not have the same shape parameters. This is also another significant contribution of this paper.

Our approach in this paper is fully Bayesian. For the Bayesian inference we need to assume some priors on the unknown parameters. First it is assumed that X and Y have a common shape parameter and it is known. Assumptions of known shape parameter in case of a Weibull distribution is not very unreasonable, see for example Murthy et al. (2004). In many reliability experiments the experimenter has a prior knowledge about the shape parameter of the corresponding Weibull model. If the shape parameter is known, and there is no covariates, then the most convenient but quite general conjugate prior on the scale parameters is a Dirichlet-gamma prior, as suggested by Pena and Gupta (1990). The Bayes estimate cannot be obtained in explicit form, but it is quite convenient to generate samples from the posterior distribution of R. Hence, we use Gibbs sampling procedure to compute the Bayes estimate and the associated highest posterior density (HPD) credible interval of R.

When the covariates are also present, it is assumed that the covariate parameters have independent normal priors. In case of unknown common shape parameter, the conjugate prior on the shape parameter does not exist. In this case following the approach of Berger and Sun (1993), it is assumed that the prior on the shape parameter has the support on $(0,\infty)$, and it has a log-concave density function. Based on the above prior distributions, we obtain the joint posterior distribution function of the unknown parameters. The Bayes estimate of R cannot be obtained in explicit form and we have proposed an effective importance sampling procedure to compute the Bayes estimate and the associated HPD credible interval of the stress-strength parameter.

We further consider the case when the two shape parameters of the Weibull distributions are different and unknown. In this case the stress-strength parameter cannot be obtained in explicit form, and it can be obtained only in an integration form. In this case the priors on the scale parameters are same as before, and for the shape parameters it is assumed that the two prior distributions are independent, both of them have support on $(0,\infty)$ and they have log-concave density functions. We have used importance sampling technique to compute the Bayes estimate and the associated HPD credible interval of R. We perform some simulation experiments to see the effectiveness of the proposed methods. One small data set has been analyzed for illustrative purposes and to show how the proposed methods work for small sample sizes.

Rest of the paper is organized as follows. In Section 1.2, we provide the model description, and the prior assumptions. In Section 1.3, we provide the posterior analysis when there is no covariate. In Section 1.4, we consider the case when the covariates are present. Simulation results are provided in Section 1.5. The analysis of one data set is presented in Section 1.6. Finally, in Section 1.7 we conclude the paper.

1.2 Model Formulation and Prior Assumptions

1.2.1 Model Formulation & Available Data

A Weibull distribution with the shape parameter $\alpha > 0$ and scale parameter $\lambda > 0$, has the probability density function (PDF)

$$f_{WE}(x;\alpha,\lambda) = \alpha\lambda x^{\alpha-1}e^{-\lambda x^{\alpha}}; \quad x > 0, \tag{1.1}$$

and it will be denoted by $\mathrm{WE}(\alpha,\lambda)$.

Two Shape Parameters are Same: It is assumed that X follows ($\sim$) $\mathrm{WE}(\alpha,\lambda_1)$ and $Y \sim \mathrm{WE}(\alpha,\lambda_2)$, and they are independently distributed. The choice of a common α can be justified as follows. The logarithm of a Weibull random variable has an extreme value distribution, which belongs to a location scale family. The parameter α is the reciprocal of the scale parameter. In an analogy to the two-sample normal case, where it is customary to assume a common variance, we assume a common shape parameter α here.

In this case

$$R = P(Y < X) == \int_0^\infty \alpha\lambda_2 y^{\alpha-1} e^{-(\lambda_1+\lambda_2)y^\alpha} dy = \frac{\lambda_2}{\lambda_1+\lambda_2}.$$

If both X and Y depend on some covariates, $u = (u_1,\ldots,u_p)^T$ and $v = (v_1,\ldots,v_q)^T$, respectively, as follows;

$$\lambda_1 = \theta_1 e^{\beta_1^T u} \quad \text{and} \quad \lambda_2 = \theta_2 e^{\beta_2^T v},$$

where $\beta_1 = (\beta_{11},\ldots,\beta_{1p})^T$ and $\beta_2 = (\beta_{21},\ldots,\beta_{2q})^T$ are p and q dimensional regression parameters, (p may not be equal to q), then

$$R(u,v) = P(Y < X|u,v) = \frac{\theta_2 e^{\beta_2^T v}}{\theta_1 e^{\beta_1^T u} + \theta_2 e^{\beta_2^T v}}. \tag{1.2}$$

From now on we use R instead of $R(u,v)$ for brevity. But if it is needed we shall make it explicit.

Two Shape Parameters are Different: Now we discuss the case when the two shape parameters are different. In this case, it is assumed that $X \sim \mathrm{WE}(\alpha_1,\lambda_1)$ and $Y \sim \mathrm{WE}(\alpha_2,\lambda_2)$, and they are independently distributed. In this case R cannot be obtained in explicit form. It can be obtained only in the integration form as follows:

$$R = P(Y < X) = \alpha_2\lambda_2 \int_0^\infty y^{\alpha_2-1} e^{-(\lambda_1 y^{\alpha_1}+\lambda_2 y^{\alpha_2})} dy. \tag{1.3}$$

Moreover, when X and Y depend on some covariates as before, then

$$R(u,v) = P(Y < X|u,v) = \alpha_2\theta_2 \int_0^\infty e^{\beta_2^T v} y^{\alpha_2-1} e^{-(\theta_1 e^{\beta_1^T u} y^{\alpha_1}+\theta_2 e^{\beta_2^T v} y^{\alpha_2})} dy. \tag{1.4}$$

It is assumed that we have the following stress strength data as $\{x_i,u_{i1},\ldots,u_{ip}\}$, for $i = 1,\ldots,m$, and $\{y_j,v_{j1},\ldots,v_{jq}\}$, for $j = 1,\ldots,n$, respectively, and all the observations are independently distributed.

1.2.2 Prior Assumptions

We use the following notation. A random variable is said to have a gamma distribution with the shape parameter $\alpha > 0$ and scale parameter $\lambda > 0$, if it has the PDF

$$f_{GA}(x;\alpha,\lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}; \quad x > 0. \tag{1.5}$$

From now on it will be denoted by $\text{GA}(\alpha,\lambda)$. A random variable is said to have a Beta distribution with parameters $a > 0$ and $b > 0$, if it has the PDF

$$f_{BE}(x;a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}; \quad 0 < x < 1, \tag{1.6}$$

and it will be denoted by $\text{Beat}(a,b)$. A p-variate normal random vector, with mean vector a and dispersion matrix B, will be denoted by $\text{N}_p(a,B)$.

We make the following prior assumptions on the different parameters as follows. Let us assume $\theta = \theta_1 + \theta_2$, then similarly as in Pena and Gupta (1990), it is assumed that $\theta \sim \text{GA}(a_0,b_0)$, with $a_0 > 0$, $b_0 > 0$, and $\theta_1/(\theta_1+\theta_2) \sim \text{Beta}(a_1,a_2)$, with $a_1 > 0$, $a_2 > 0$, and they are independently distributed. After simplification, it can be easily seen that the joint PDF of θ_1 and θ_2 takes the following form;

$$\begin{aligned}\pi_{BG}(\theta_1,\theta_2|a_0,b_0,a_1,a_2) = \frac{\Gamma(a_1+a_2)}{\Gamma(a_0)} (b_0\theta)^{a_0-a_1-a_2} \times \frac{b_0^{a_1}}{\Gamma(a_1)}\, \theta_1^{a_1-1} e^{-b_0\theta_1} \times \\ \frac{b_0^{a_2}}{\Gamma(a_2)}\, \theta_2^{a_2-1} e^{-b_0\theta_2}.\end{aligned} \tag{1.7}$$

The joint PDF (1.7) is known as the PDF of a Beta-Gamma distribution, and it will be denoted by $\text{BG}(b_0,a_0,a_1,a_2)$. The above Beta-Gamma prior is a very flexible prior. If $a_0 = a_1 + a_2$, then θ_1 and θ_2 become independent. If $a_0 > a_1 + a_2$, then θ_1 and θ_2 are positively correlated, and for $a_0 < a_1 + a_2$, they are negatively correlated. The following result will be useful for posterior analysis.

Result: If $(\theta_1,\theta_2) \sim \text{BG}(b_0,a_0,a_1,a_2)$, then for $i = 1,2$,

$$E(\theta_i) = \frac{a_0 a_i}{b_0(a_1+a_2)} \quad \text{and} \quad V(\theta_i) = \frac{a_0 a_i}{b_0^2(a_1+a_2)} \times \left\{ \frac{(a_i+1)(a_0+1)}{a_1+a_2+1} - \frac{a_0 a_i}{a_1+a_2} \right\}. \tag{1.8}$$

Moreover, Kundu and Pradhan (2011) suggested a very simple method to generate samples from a Beta-Gamma distribution, and that will be useful to generate samples from the posterior distribution. When the common shape parameter is unknown, it is assumed that the prior on α, $\pi_2(\alpha)$, is absolute continuous, and it has log-concave PDF.

In presence of covariates, we make the following prior assumptions on the unknown parameters:

$$(\theta_1,\theta_2) \sim \pi_1(\theta_1,\theta_2) = \text{BG}(b_0,a_0,a_1,a_2)$$

$$\begin{aligned}\alpha &\sim \pi_2(\alpha)\\ \beta_1 &\sim \pi_3(\beta_1) = N_p(0, \Sigma_1)\\ \beta_2 &\sim \pi_4(\beta_2) = N_q(0, \Sigma_2),\end{aligned}$$

and they are all independently distributed. It is further assumed that $\pi_2(\alpha)$ is absolute continuous and has a log-concave probability density function. If the two shape parameters, α_1 and α_2, are different, then for (θ_1, θ_2), β_1, β_2, we assume the same prior as before, and for α_1 and α_2, it is assumed that

$$\alpha_1 \sim \pi_{21}(\alpha_1) \quad \text{and} \quad \alpha_2 \sim \pi_{22}(\alpha_2), \tag{1.9}$$

and they are independently distributed. Here $\pi_{21}(\alpha_1)$ and $\pi_{22}(\alpha_2)$ are both absolute continuous and having log-concave PDFs.

1.3 Posterior Analysis: No Covariates

In this section, we provide the Bayesian inference of the unknown stress strength parameter based on the prior assumptions provided in Section 1.2, and when covariates are absent. We consider the case in presence of covariates in the next section. In developing the Bayes estimate, it is assumed that the error is squared error, although any other loss function can be easily incorporated. In this section we consider two cases separately, namely when (i) common shape parameter is known, and (ii) common shape parameter is unknown.

1.3.1 Two Shape Parameter are Same: Known

Based on the observations, and the prior assumptions as provided in Section 2, we obtain the likelihood function as

$$l(data|\alpha, \lambda_1, \lambda_2) = \alpha^{m+n} \lambda_1^m \lambda_2^n \prod_{i=1}^{m} x_i^{\alpha-1} \prod_{j=1}^{n} y_j^{\alpha-1} e^{-\lambda_1 \sum_{i=1}^{m} x_i^{\alpha} - \lambda_2 \sum_{j=1}^{n} y_j^{\alpha}}. \tag{1.10}$$

Therefore, for known α, the posterior distribution of λ_1 and λ_2 becomes

$$\pi(\lambda_1, \lambda_2|\alpha, data) \propto \lambda_1^{m+a_1-1} \lambda_2^{n+a_2-1} (\lambda_1 + \lambda_2)^{a_0-a_1-a_2} e^{-\lambda_1 T_1(\alpha) - \lambda_2 T_2(\alpha)}, \tag{1.11}$$

where

$$T_1(\alpha) = b_0 + \sum_{i=1}^{m} x_i^{\alpha} \quad \text{and} \quad T_2(\alpha) = b_0 + \sum_{j=1}^{n} y_j^{\alpha}.$$

Now observe that for a given α, if $a_0 > 0$, $a_1 > 0$, $a_2 > 0$ and $b_0 > 0$

$$\int_0^\infty \int_0^\infty \lambda_1^{m+a_1-1} \lambda_2^{n+a_2-1} (\lambda_1+\lambda_2)^{a_0-a_1-a_2} e^{-\lambda_1 T_1(\alpha) - \lambda_2 T_2(\alpha)} d\lambda_1 d\lambda_2 \leq$$
$$\int_0^\infty \int_0^\infty \lambda_1^{m+a_1-1} \lambda_2^{n+a_2-1} (\lambda_1+\lambda_2)^{a_0-a_1-a_2} e^{-(\lambda_1+\lambda_2)\min\{T_1(\alpha),T_2(\alpha)\}} d\lambda_1 d\lambda_2 < \infty.$$

Hence, if the prior is proper, the corresponding posterior is also proper. Note that in this case even if $a_0 = a_1 = a_2 = b_0 = 0$, then also the posterior PDF becomes the product of two independent gamma PDFs, hence it is proper. The following result will be useful for further development.

Theorem 1: The posterior distribution of $R = \dfrac{\lambda_1}{\lambda_1+\lambda_2}$ given α is given by

$$\pi(R|data,\alpha) \propto \frac{R^{m+a_1-1}(1-R)^{n+a_2-1}}{(RT_1(\alpha)+(1-R)T_2(\alpha))^{m+n+a_1+a_2}}; \quad 0 < R < 1.$$

PROOF: Let us make the following transformation:

$$R = \frac{\lambda_1}{\lambda_1+\lambda_2} \quad \text{and} \quad \lambda = \lambda_1+\lambda_2.$$

Therefore, the joint posterior distribution of R and λ becomes

$$\pi(R,\lambda|data,\alpha) \propto R^{m+a_1-1}(1-R)^{n+a_2-1}\lambda^{m+n+a_1+a_2-1} e^{-\lambda(RT_1(\alpha)+(1-R)T_2(\alpha))}. \tag{1.12}$$

Now integrating out λ from 0 to ∞, the result follows. □

Therefore, the Bayes estimate of R with respect to the squared error loss function becomes

$$\widehat{R}_{Bayes} = E(R|data,\alpha) = K\int_0^1 \frac{R^{m+a_1}(1-R)^{n+a_2-1}}{(RT_1(\alpha)+(1-R)T_2(\alpha))^{m+n+a_1+a_2}}\, dR. \tag{1.13}$$

Here K is the normalizing constant, *i.e.*

$$K^{-1} = \int_0^1 \frac{R^{m+a_1-1}(1-R)^{n+a_2-1}}{(RT_1(\alpha)+(1-R)T_2(\alpha))^{m+n+a_1+a_2}}\, dR.$$

Clearly, $\widehat{R}_{Bayes}$ cannot be obtained in explicit forms. It is possible to use Lindley's approximation or Laplace's approximation to compute an approximate Bayes estimate of R. But it may not be possible to construct the highest posterior density (HPD) credible interval of R in that way. Observe that

$$\pi(R|data,\alpha) \leq K\frac{R^{m+a_1-1}(1-R)^{n+a_2-1}}{(\min\{T_1(\alpha),T_2(\alpha)\})^{m+n+a_1+a_2}}, \tag{1.14}$$

therefore, acceptance-rejection method can be easily used to generate samples directly from the posterior distribution of R, and the generated samples can be used to compute Bayes estimate and also to construct HPD credible interval of R.

1.3.2 Common Shape Parameter Unknown

In this case, the joint posterior density function of α, λ_1 and λ_2 can be written as

$$\pi(\lambda_1,\lambda_2,\alpha|data) \propto \lambda_1^{m+a_1-1}\lambda_2^{n+a_2-1}(\lambda_1+\lambda_2)^{a_0-a_1-a_2}e^{-\lambda_1 T_1(\alpha)-\lambda_2 T_2(\alpha)}\alpha^{m+n} S^{\alpha-1}\pi_2(\alpha), \quad (1.15)$$

where $S=\prod_{i=1}^{m} x_i \prod_{j=1}^{n} y_j$. Therefore, we have the similar result as Theorem 1, and whose proof also can be obtained along the same line. Hence it is avoided.

Theorem 2: The joint posterior distribution of $R=\dfrac{\lambda_1}{\lambda_1+\lambda_2}$ and α for $0<R<1$ and $0<\alpha<\infty$, is given by

$$\pi(R,\alpha|data) \propto \frac{R^{m+a_1-1}(1-R)^{n+a_2-1}}{(RT_1(\alpha)+(1-R)T_2(\alpha))^{m+n+a_1+a_2}} \times \alpha^{m+n}S^{\alpha-1}\pi_2(\alpha).$$

Therefore, in this case, the Bayes estimate of R with respect to the squared error loss function becomes

$$\widehat{R}_{Bayes} = K_1 \int_0^1 \int_0^\infty \frac{R^{m+a_1}(1-R)^{n+a_2-1}}{(RT_1(\alpha)+(1-R)T_2(\alpha))^{m+n+a_1+a_2}} \times \alpha^{m+n}S^{\alpha-1}\pi_2(\alpha)\; d\alpha\, dR, \quad (1.16)$$

here K_1 is the proportionality constant, and it can be obtained as

$$K_1^{-1} = \int_0^1 \int_0^\infty \frac{R^{m+a_1-1}(1-R)^{n+a_2-1}}{(RT_1(\alpha)+(1-R)T_2(\alpha))^{m+n+a_1+a_2}} \times \alpha^{m+n}S^{\alpha-1}\pi_2(\alpha)\; d\alpha\, dR.$$

Clearly, the Bayes estimate of R with respect to the squared error loss function cannot be obtained in closed form. Again, in this case also the Laplace's or Lindley's approximation may be used to approximate the Bayes estimate of R, but the associated credible interval cannot be obtained. We propose to use the importance sampling procedure to compute Bayes estimate and also the associated credible interval as follows. Let us re-write the joint posterior density of R and α as follows:

$$\pi(R,\alpha|data) = K_1 \times \text{Beta}(R; m+a_1, n+a_2) \times g_1(\alpha|data) \times h(\alpha,R)$$

Here $g_1(\alpha|data)$ is a proper density function, such that

$$g_1(\alpha|data) \propto \alpha^{m+n}S^{\alpha-1}\pi_2(\alpha) \quad (1.17)$$

and

$$h(\alpha,R) = \frac{1}{(RT_1(\alpha)+(1-R)T_2(\alpha))^{m+n+a_1+a_2}}.$$

Note that $g_1(\alpha|data)$ is log-concave, hence using the method of Devroye (1984), it is very easy to generate sample from $g_1(\alpha|data)$. The following algorithm can be used to compute Bayes estimate of R, and also to construct associated credible interval.

Algorithm 1:

- Step 1: Generate $\alpha_1 \sim g_1(\alpha|data)$ and $R_1 \sim \text{Beta}(m+a_1, n+a_2)$.
- Step 2: Repeat Step 1, and obtain $(\alpha_1, R_1), \ldots, (\alpha_N, R_N)$.
- Step 3: A simulation consistent Bayes estimate of R can be obtained as

$$\widehat{R}_{Bayes} = \frac{\sum_{i=1}^{N} R_i h(\alpha_i, R_i)}{\sum_{j=1}^{N} h(\alpha_j, R_j)}.$$

- Step 4: Now to construct a $100(1-\gamma)\%$ HPD credible intervals, we propose the following method, see also Chen and Shao (1999) in this respect. let us denote

$$w_i = \frac{h(\alpha_i, R_i)}{\sum_{j=1}^{N} h(\alpha_j, R_j)}; \text{ for } i = 1, \ldots, N.$$

Rearrange, $\{(\alpha_1, R_1), \ldots, (\alpha_N, R_N)\}$, as follows $\{(R_{(1)}, w_{[1]}), \ldots, (R_{(N)}, w_{[N]})\}$, where $R_{(1)} < R_{(2)} < \ldots < R_{(N)}$. In this case, $w_{[i]}$'s are not ordered, they are just associated with $R_{(i)}$. Let N_p be the integer satisfying

$$\sum_{i=1}^{N_p} w_{[i]} \leq p < \sum_{i=1}^{N_p+1} w_{[i]},$$

for $0 < p < 1$. A $100(1-\gamma)\%$ credible interval of R can be obtained as $(R_{(N_\delta)}, R_{(N_{\delta+1-\gamma})}$, for $\delta = w_{[1]}, w_{[1]} + w_{[2]}, \ldots, \sum_{i=1}^{N_{1-\gamma}} w_{[i]}$. Therefore, a $100(1-\gamma)\%$ HPD credible interval of R becomes $(R_{(N_{\delta^*})}, R_{(N_{\delta^*+1-\gamma})})$, where

$$R_{(N_{\delta^*+1-\gamma})} - R_{(N_{\delta^*})} \leq R_{(N_{\delta+1-\gamma})} - R_{(N_\delta)} \quad \text{for all } \delta.$$

1.4 Inference on R: Presence of Covariates

1.4.1 Two Shape Parameters are Same

In this section we consider the Bayesian inference of R in presence of covariates (u, v), based on the observations and prior assumptions provided in Section 1.2. In this section we will make it explicit that $R = R(u, v) = P(Y < X|u, v)$. Let us use the following notations:

$$T_1(\alpha, \beta_1, U) = \sum_{i=1}^{m} x_i^{\alpha} e^{\beta_1^T u_i} \quad \text{and} \quad T_2(\alpha, \beta_2, V) = \sum_{j=1}^{n} y_j^{\alpha} e^{\beta_2^T v_j}.$$

Here U and V are covariate vectors of orders $m \times p$ and $n \times q$, respectively, as follows

$$U = \begin{bmatrix} u_{11} & \dots & u_{1p} \\ \vdots & \ddots & \vdots \\ u_{m1} & \dots & u_{mp} \end{bmatrix} = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} v_{11} & \dots & v_{1q} \\ \vdots & \ddots & \vdots \\ v_{n1} & \dots & v_{nq} \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

Therefore, based on the model assumptions, and the posterior distribution, the posterior distribution function of θ_1, θ_2, α, β_1 and β_2 can be written as

$$\begin{aligned}
&\pi(\theta_1,\theta_2,\alpha,\beta_1,\beta_2|data) \propto \\
&\quad \theta_1^{m+a_1-1} \exp\{-\theta_1(b_0+T_1(\alpha,\beta_1))\} \times \theta_2^{n+a_2-1} \exp\{-\theta_1(b_0+T_2(\alpha,\beta_2))\} \\
&\quad \times (\theta_1+\theta_2)^{a_0-a_1-a_2} \times \pi(\alpha)\alpha^{m+n}S^{\alpha} \times \exp\left\{\sum_{i=1}^{m} \beta_1^T u_i\right\} \prod_{k=1}^{p} \frac{1}{\sqrt{2\pi}\sigma_{1k}} \exp\left(-\frac{\beta_{1k}^2}{2\sigma_{1k}^2}\right) \\
&\quad \times \exp\left\{\sum_{j=1}^{n} \beta_2^T v_j\right\} \prod_{k=1}^{q} \frac{1}{\sqrt{2\pi}\sigma_{2k}} \exp\left(-\frac{\beta_{2k}^2}{2\sigma_{2k}^2}\right). \qquad (1.18)
\end{aligned}$$

Therefore, for a given u and v, with respect to squared error loss function, the Bayes estimate of R becomes

$$\begin{aligned}
\widehat{R}_B(u,v) = \int_0^{\infty}\int_0^{\infty}\int_0^{\infty}\int_{\mathbb{R}^p}\int_{\mathbb{R}^q} \frac{\theta_2 e^{\beta_2^T v}}{\theta_1 e^{\beta_1^T u} + \theta_2 e^{\beta_2^T v}} \pi(\theta_1,\theta_2,\alpha,\beta_1,\beta_2|data) \\
d\beta_1 d\beta_2\, d\theta_1\, d\theta_2\, d\alpha. \qquad (1.19)
\end{aligned}$$

Clearly, the Bayes estimate of R cannot be obtained in closed form. For a given (u,v), we use the following importance sampling technique to compute the Bayes estimate of R , and the associated HPD credible interval of R. Note that (1.18) can be written as

$$\begin{aligned}
\pi(\theta_1,\theta_2,\alpha,\beta_1,\beta_2|data) = K\, h(\theta_1,\theta_2)\pi_1(\theta_1|\alpha,\beta_1,data)\pi_2(\theta_2|\alpha,\beta_2,data) \times \\
\pi_3(\alpha|\beta_1,\beta_2,data)\pi_4(\beta_1|data)\pi_5(\beta_2|data). \qquad (1.20)
\end{aligned}$$

Here, K is the normalizing constant, $\pi_1(\cdot|\alpha,\beta_1,data)$, $\pi_2(\cdot|\alpha,\beta_2,data)$, $\pi_3(\cdot|\beta_1,\beta_2,data)$, $\pi_4(\cdot|data)$ and $\pi_5(\cdot|data)$ are proper density functions, and they are as follows:

$$\pi_1(\theta_1|\alpha,\beta_1,data) \sim \text{Gamma}(\theta_1|m+a_1,(b_0+T_1(\alpha,\beta_1))) \qquad (1.21)$$

$$\pi_2(\theta_2|\alpha,\beta_2,data) \sim \text{Gamma}(\theta_2|n+a_2,(b_0+T_2(\alpha,\beta_2))) \qquad (1.22)$$

$$\pi_3(\alpha|\beta_1,\beta_2,data) \propto \frac{\pi(\alpha)\alpha^{m+n}S^{\alpha}}{(b_0+T_1(\alpha,\beta_1))^{m+a_1} \times (b_0+T_2(\alpha,\beta_2))^{n+a_2}} \qquad (1.23)$$

$$\pi_4(\beta_1|data) = \prod_{k=1}^{p} N(c_k\sigma_{1k},\sigma_{1k}^2)$$

$$\pi_5(\beta_2|data) = \prod_{k=1}^{q} N(d_k\sigma_{2k}, \sigma_{2k}^2),$$

where $c_k = \sum_{i=1}^{m} u_{ik}$ for $k = 1,\ldots,p$, and $d_k = \sum_{j=1}^{n} v_{jk}$, for $k = 1,\ldots,q$, and

$$h(\theta_1, \theta_2) = (\theta_1 + \theta_2)^{a_0 - a_1 - a_2}.$$

The following result will be useful for further development.
Theorem 3: $\pi_3(\alpha|\beta_1, \beta_2, data)$ is log-concave.
PROOF: See in the Appendix.

Note that it is quite standard to generate random samples from $\pi_1(\theta_1|\alpha, \beta_1, data)$, $\pi_2(\theta_2|\alpha, \beta_2, data)$, $\pi_4(\beta_1|data)$, $\pi_5(\beta_2|data)$, and using the method of Devroye (1984) or Kundu (2008) random samples from $\pi_3(\alpha|\beta_1, \beta_2, data)$ also can be obtained quite conveniently.

The following Algorithm 2 can be used to compute a simulation consistent estimate of (1.19) and the associated HPD credible interval.
Algorithm 2:

- Step 1: Generate β_1 and β_2 from $\pi_4(\beta_1|data)$ and $\pi_5(\beta_2|data)$, respectively.
- Step 2: For a given β_1 and β_2, generate α from $\pi_3(\alpha|\beta_1, \beta_2, data)$, using the method suggested by Devroye (1984) or Kundu (2008).
- Step 3: Now for a given α, β_1 and β_2, generate θ_1 and θ_2, from $\pi_1(\theta_1|\alpha, \beta_1, data)$ and $\pi_2(\theta_2|\alpha, \beta_2, data)$, respectively.
- Step 4: Repeat Step 1 to Step 4, N times and obtain

$$\{(\alpha_i, \beta_{1i}, \beta_{2i}, \theta_{1i}, \theta_{2i}), i = 1,\ldots,N\},$$

and also compute for $i = 1,\ldots,N$,

$$R_i(u,v) = \frac{\theta_{2i}e^{\beta_{2i}^T v}}{\theta_{1i}e^{\beta_{1i}^T u} + \theta_{2i}e^{\beta_{2i}^T v}}.$$

- Step 5: A simulation consistent Bayes estimate of $R(u,v)$ can be obtained as

$$\widehat{R}_B(u,v) = \frac{\sum_{i=1}^{N} R_i(u,v)h(\theta_{1i}, \theta_{2i})}{\sum_{j=1}^{N} h(\theta_{1j}, \theta_{2j})}.$$

- Step 6: Now to construct a $100(1-\gamma)\%$ HPD credible interval of $R(u,v)$, first let us denote

$$R_i = R_i(u,v) \quad \text{and} \quad w_i = \frac{h(\theta_{1i}, \theta_{2i})}{\sum_{j=1}^{N} h(\theta_{1j}, \theta_{2j})}; \;\; \text{for} \;\; i = 1,\ldots,N.$$

Rearrange, $\{(R_1, w_1),\ldots,(R_N, w_N)\}$ as $\{(R_{(1)}, w_{[1]}),\ldots,(R_{(N)}, w_{[N]})\}$, where $R_{(1)} < \ldots, R_{(N)}$, and $w_{[i]}$'s are not ordered as before, and they are just associated with $R_{(i)}$. Let N_p be the integer satisfying

$$\sum_{i=1}^{N_p} w_{[i]} \le p < \sum_{i=1}^{N_p+1} w_{[i]},$$

for $0 < p < 1$. A $100(1-\gamma)\%$ credible interval of R can be obtained as $(R_{(N_\delta)}, R_{(N_{\delta+1-\gamma})}$, for $\delta = w_{[1]}, w_{[1]} + w_{[2]}, \ldots, \sum_{i=1}^{N_{1-\gamma}} w_{[i]}$. Therefore, a $100(1-\gamma)\%$ HPD credible interval of $R(u,v)$ becomes $(R_{(N_{\delta^*})}, R_{(N_{\delta^*+1-\gamma})})$, where

$$R_{(N_{\delta^*+1-\gamma})} - R_{(N_{\delta^*})} \le R_{(N_{\delta+1-\gamma})} - R_{(N_\delta)} \quad \text{for all } \delta.$$

1.4.2 Two Shape Parameters are Different

In this section we consider the case when the two shape parameters are not same. In this case it is assumed that $X \sim \text{WE}(\alpha_1, \lambda_1)$, $Y \sim \text{WE}(\alpha_2, \lambda_2)$, λ_1 and λ_2 follow similar relations as in Section 1.2. We make the following prior assumptions on the unknown parameters.

$$\begin{aligned}
(\theta_1, \theta_2) &\sim \pi_1(\theta_1, \theta_2) = \text{DG}(b_0, a_0, a_1, a_2) \\
\alpha_1 &\sim \pi_{21}(\alpha_1), \quad \alpha_2 \sim \pi_{22}(\alpha_2) \\
\beta_1 &\sim \pi_3(\beta_1) = N_p(0, \Sigma_1) \\
\beta_2 &\sim \pi_4(\beta_2) = N_q(0, \Sigma_2).
\end{aligned}$$

Here π_{21} and π_{22} have absolute continuous PDFs, and both the PDFs are log-concave, and all are independently distributed.

Therefore, based on the above prior assumptions, and using the same notations as in Section 1.4.1, the posterior distribution function of θ_1, θ_2, α_1, α_2, β_1 and β_2 can be written as

$$\begin{aligned}
&\pi(\beta_1, \beta_2, \theta_1, \theta_2, \alpha_1, \alpha_2 | data) \propto \\
&\quad \theta_1^{m+a_1-1} \exp\{-\theta_1(b_0 + T_1(\alpha_1, \beta_1))\} \times \theta_2^{n+a_2-1} \exp\{-\theta_2(b_0 + T_2(\alpha_2, \beta_2))\} \\
&\quad \times (\theta_1 + \theta_2)^{a_0 - a_1 - a_2} \times \pi_{21}(\alpha_1) \alpha_1^m S_1^{\alpha_1} \times \pi_{22}(\alpha_2) \alpha_2^n S_2^{\alpha_2} \\
&\quad \times \exp\left\{\sum_{i=1}^m \beta_1^T u_i\right\} \prod_{i=1}^p \frac{1}{\sqrt{2\pi}\sigma_{1i}} \exp\left(-\frac{\beta_{1i}^2}{2\sigma_{1i}^2}\right) \\
&\quad \times \exp\left\{\sum_{j=1}^n \beta_2^T v_j\right\} \prod_{j=1}^q \frac{1}{\sqrt{2\pi}\sigma_{2j}} \exp\left(-\frac{\beta_{2j}^2}{2\sigma_{2j}^2}\right). \qquad (1.24)
\end{aligned}$$

Here $S_1 = \prod_{i=1}^{m} x_i$ and $S_2 = \prod_{j=1}^{n} y_j$. Therefore, for a given u and v, with respect to squared error loss function, the Bayes estimate of R becomes

$$\widehat{R}_B(u,v) = \int_0^{\infty}\int_0^{\infty}\int_0^{\infty}\int_{\mathbb{R}^p}\int_{\mathbb{R}^q}\int_0^{\infty} g(\Gamma, y|u,v) \times \pi(\Gamma|data)\; dy d\Gamma, \qquad (1.25)$$

where $\Gamma = (\beta_1, \beta_2, \theta_1, \theta_2, \alpha_1, \alpha_2)$, $d\Gamma = (d\beta_1, d\beta_2, d\theta_1, d\theta_2, d\alpha_1, d\alpha_2)$ and

$$g(\Gamma, y|u,v) = \alpha_2\theta_2 e^{\beta_2^T v} y^{\alpha_2 - 1} e^{-(\theta_1 e^{\beta_1^T u} y^{\alpha_1} + \theta_2 e^{\beta_2^T v} y^{\alpha_2})}.$$

Clearly, (1.25) cannot be obtained in explicit form. We proceed as follows. Let us re-write

$$\widehat{R}_B(u,v) = E(g(Y, \Theta_1, \alpha_1, \Theta_2, \alpha_2, \beta_1, \beta_2; u, v)|data), \qquad (1.26)$$

here

$$g(y, \theta_1, \alpha_1, \theta_2, \alpha_2, \beta_1, \beta_2; u, v) = e^{-\theta_1 e^{\beta_1^T u} y^{\alpha_1} - \theta_2 e^{\beta_2^T v} y^{\alpha_2} + \beta_2^T v + \theta_2 y^{\alpha_2}},$$

and the joint PDF of Δ given the $data$, where $\Delta = (Y, \Theta_1, \alpha_1, \Theta_2, \alpha_2, \beta_1, \beta_2)$ can be written as

$$\begin{aligned}\pi(\Delta|data) = K &\times \pi_0(y|\alpha_2, \theta_2, data) \times \pi_1(\theta_1|\alpha_1, \beta_1, data) \times \pi_2(\theta_2|\alpha_2, \beta_2, data) \times \\ &\pi_3(\alpha_1|\beta_1, data) \times \pi_4(\alpha_2|\beta_2, data) \times \pi_5(\beta_1|data) \times \pi_6(\beta_2|data) \times \\ &h(\theta_1, \theta_2),\end{aligned}$$

and

$$\begin{aligned}
&\pi_0(y|\alpha_2, \theta_2, data) \sim \text{WE}(\alpha_2, \theta_2)\\
&\pi_1(\theta_1|\alpha_1, \beta_1, data) \sim \text{Gamma}(\theta_1|m + a_1, (b_0 + T_1(\alpha_1, \beta_1)))\\
&\pi_2(\theta_2|\alpha_2, \beta_2, data) \sim \text{Gamma}(\theta_2|n + a_2, (b_0 + T_2(\alpha_2, \beta_2)))\\
&\pi_3(\alpha_1|\beta_1, data) \propto \frac{\pi_{21}(\alpha_1)\alpha_1^m S_1^{\alpha_1}}{(b_0 + T_1(\alpha_1, \beta_1))^{m+a_1}}\\
&\pi_4(\alpha_2|\beta_2, data) \propto \frac{\pi_{22}(\alpha_2)\alpha_2^n S_2^{\alpha_2}}{(b_0 + T_2(\alpha_2, \beta_2))^{n+a_2}}
\end{aligned}$$

$$\begin{aligned}
\pi_5(\beta_1|data) &= \prod_{k=1}^{p} N(c_k\sigma_{1k}, \sigma_{1k}^2)\\
\pi_6(\beta_2|data) &= \prod_{k=1}^{q} N(d_k\sigma_{2k}, \sigma_{2k}^2),
\end{aligned}$$

where $c_k = \sum_{i=1}^{m} u_{ik}$ for $k = 1, \ldots, p$, and $d_k = \sum_{j=1}^{n} v_{jk}$, for $k = 1, \ldots, q$, and

$$h(\theta_1, \theta_2) = (\theta_1 + \theta_2)^{a_0 - a_1 - a_2}.$$

Following Theorem 3, it can be easily shown that $\pi_3(\alpha_1|\beta_1, data)$ and $\pi_4(\alpha_2|\beta_2, data)$ both are log-concave. Exactly similar algorithm like Algorithm 2, with an extra generation of $Y \sim \text{WE}(\alpha_2, \theta_2)$, can be used to compute a simulation consistent estimate of $\widehat{R}_B(u,v)$ and the associated HPD credible interval. The explicit details are not provided to avoid repetition.

1.5 Simulation Experiments

In this section we perform some simulation experiments mainly to compare the non-parametric estimates of R with the MLE and the Bayes estimates of R. We would also like to see whether the parametric estimates are robust or not with respect to the distributional assumptions. We have taken four different cases namely (i) $X \sim$ GA(2,1), $Y \sim$ GA(3,2), (ii) $X \sim$ GA(2,1), $Y \sim$ GA(4,2), (iii) $X \sim$ WE(2,1), $Y \sim$ WE(3,2), (iv) $X \sim$ WE(2,1), $Y \sim$ WE(4,2), and different sample sizes namely $m = n$ = 20, 40, 60, 80 and 100. In each case we estimate R non-parametrically (NPE) and also obtain the MLE and the Bayes estimate of R based on the assumption that X and Y follow Weibull distributions with different shape and scale parameters. To compute the Bayes estimate of R we have taken the following prior distributions of the unknown parameters:

$$\alpha_1 \sim \text{GA}(0.0001, 0.0001), \quad \lambda_1 \sim \text{GA}(0.0001, 0.0001),$$
$$\alpha_2 \sim \text{GA}(0.0001, 0.0001), \quad \lambda_2 \sim \text{GA}(0.0001, 0.0001).$$

We replicate the process 1000 times in each case and obtain the average estimates and the mean squared errors (MSEs). The results are reported in Table 1.1 and Table 1.2. In each box the first figure represents the average estimate and the corresponding MSE is reported within bracket below.

Some of the points are quite clear from this simulation experiments. The performances of the MLEs and the Bayes estimates based on the non-informative priors are very similar both in terms of biases and MSEs. It is observed that as the sample size increases in all the cases for the non-parametric estimates the biases and the MSEs decrease. It shows the consistency property of the non-parametric estimates. In case of MLEs and Bayes estimates, the MSEs decrease but when the model assumptions are not correct the biases increase. Therefore, it is clear that if we know the distributions of X and Y, it is better to use the parametric estimate otherwise it is better to use the non-parametric estimate, which are more robust with respect to the model assumptions, as expected.

Table 1.1 Comparison of the parametric and non-parametric estimates

—n—	$X \sim$ GA(2,1), $Y \sim$ GA(3,2) $R = 0.4074$			$X \sim$ WE(2,1), $Y \sim$ WE(3,2) $R = 0.3953$		
	NPE	MLE	Bayes	NPE	MLE	Bayes
20	0.4056 (0.0116)	0.4081 (0.0020)	0.4079 (0.0018)	0.3948 (0.0116)	0.3919 (0.0015)	0.3923 (0.0016)
40	0.4065 (0.0062)	0.4119 (0.0013)	0.4117 (0.0014)	0.3961 (0.0057)	0.3934 (0.0011)	0.3929 (0.0013)
60	0.4047 (0.0042)	0.4130 (0.0009)	0.4128 (0.0010)	0.3979 (0.0040)	0.3943 (0.0007)	0.3939 (0.0009)
80	0.4045 (0.0029)	0.4138 (0.0007)	0.4136 (0.0008)	0.3979 (0.0031)	0.3949 (0.0006)	0.3948 (0.0007)
100	0.4066 (0.0025)	0.4138 (0.0006)	0.4139 (0.0006)	0.3981 (0.0025)	0.3951 (0.0005)	0.3952 (0.0005)

Table 1.2 Comparison of the parametric and non-parametric estimates

—n—	$X \sim$ GA(2,1), $Y \sim$ GA(4,2) $R = 0.5391$			$X \sim$ WE(2,1), $Y \sim$ WE(4,2) $R = 0.4381$		
	NPE	MLE	Bayes	NPE	MLE	Bayes
20	0.5354 (0.0118)	0.4031 (0.0201)	0.4025 (0.0198)	0.4348 (0.0135)	0.4316 (0.0042)	0.4321 (0.0039)
40	0.5396 (0.0065)	0.4052 (0.0194)	0.4055 (0.0190)	0.4361 (0.0076)	0.4334 (0.0032)	0.4337 (0.0034)
60	0.5385 (0.0040)	0.4105 (0.0186)	0.4101 (0.0181)	0.4361 (0.0056)	0.4343 (0.0028)	0.4351 (0.0029)
80	0.5360 (0.0031)	0.4129 (0.0175)	0.4127 (0.0173)	0.4375 (0.0047)	0.4351 (0.0016)	0.4358 (0.0018)
100	0.5383 (0.0024)	0.4156 (0.0145)	0.4151 (0.0144)	0.4381 (0.0040)	0.4381 (0.0009)	0.4382 (0.0009)

1.6 Data Analysis

In this section we present the analysis of a real data set to illustrate different methods suggested in this paper. The data set represents the measurements of shear strength of spot welds for two different gauges of steel. In this case the weld diameter (in units of 0.001 in.) is the explanatory variable for both the cases. Let X and Y denote the weld strengths for 0.040 - 0.040- in and 0.060 - 0.060- in steel, respectively. The data are presented in Table 1.3 for convenience. The main interest here is on $R = P(Y < X|u,v)$.

Before progressing further, first we have made some preliminary analysis of the data. We look at the scaled TTT for both X and Y. It is well known that it provides some indication about the shape of the empirical hazard function, see for example Aarset (1987). The scaled TTT plots for X and Y are provided in Figure 1.1. From the scaled TTT plots it is clear that both X and Y variables have increasing hazard functions, hence Weibull distribution can be used to fit these data sets.

Table 1.3 Shear strength data of welds for two gauges of steel.

Obs. No.	X	u	Y	v
1.	350	380	680	190
2.	380	155	800	200
3.	385	160	780	209
4.	450	165	885	215
5.	465	175	975	215
6.	185	165	1025	215
7.	535	195	1100	230
8.	555	185	1030	250
9.	590	195	1175	265
10.	605	210	1300	250

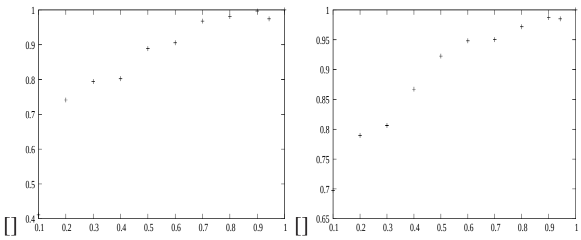

Fig. 1.1 Scaled TTT plot for (a) X and (b) Y

We have fitted WE(α_1, λ_1) to X and WE(α_2, λ_2) to Y. The maximum likelihood estimators (MLEs), the maximized log-likelihood value (MLL), the Kolmogorov-Smirnov (KS) distance between the fitted and empirical distribution function and the associated p value are presented in Table 1.4. From the KS distances and the

Table 1.4 The MLEs, KS distances and the associated p values for different cases.

Data	$\widehat{\alpha}$	$\widehat{\lambda}$	MLL	KS	p
X	4.4439	22.8226	7.0302	0.1574	0.9653
Y	5.9647	0.7434	2.7520	0.1216	0.9984

associated p values it is clear that the Weibull distribution fits both X and Y quite well. We would like to test the following hypothesis:

$$H_0 : \alpha_1 = \alpha_2 \quad \text{vs.} \quad H_1 : \alpha_1 \neq \alpha_2.$$

Under H_0, we obtain the MLEs of α, λ_1, λ_2 and MLL as $\widehat{\alpha}$ = 5.1335, $\widehat{\lambda}_1$ = 34.8707, $\widehat{\lambda}_2$ = 0.8189, MLL = 9.4476, respectively. Based on the likelihood ratio test (test

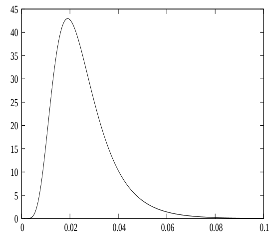

Fig. 1.2 Posterior PDF of R for known α.

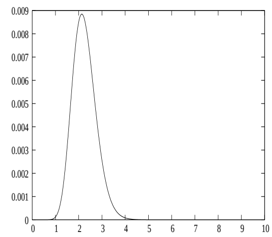

Fig. 1.3 Posterior PDF of α when the two shape parameters are equal

statistic value becomes 2(7.0302+2.7520-9.4476) = 0.6692), we cannot reject the null hypothesis as $p > 0.41$. Hence for this data set it is assumed that $\alpha_1 = \alpha_2 = \alpha$. Based on this assumption, the MLE of $R = P(Y < X)$ without the covariates becomes 0.0229.

At the beginning we consider the Bayesian inference of R ignoring the covariates. For illustrative purposes, first let us assume that the common shape parameter α is known, and we would like to obtain the Bayes estimate of R and the associated HPD credible interval. Since we do not have any prior knowledge, it is assumed throughout $a_0 = b_0 = a_1 = a_2 \approx 0$. We provide the posterior distribution of R with known α in Figure 1.2. Now we generate samples directly from $\pi(R|data, \alpha)$ using acceptance rejection principle, and based on 10000 samples we obtain the Bayes estimate with respect to the squared error loss function (posterior mean) as 0.0254, and the associated 95% HPD credible interval becomes (0.0123, 0.0502).

Now we would look at the case when the common shape parameter α is not assumed to be known. To compute the Bayes estimate of R we need to assume some specific form of $\pi_2(\alpha)$, and it is assumed that $\pi_2(\alpha) \sim \text{Gamma}(a,b)$. Since we do not have any prior knowledge about the hyper parameters, in this case also it is assumed that $a = b \approx 0$. Based on the above prior distribution, we obtain the posterior distribution of α, $g_1(\alpha|data)$, as in (1.17). The PDF of the posterior distribution of α is provided in Figure 1.3. We use Algorithm 1, and based on 10000, we obtain the Bayes estimate of R as 0.0238, and the associated 95% HPD credible intervals as (0.0115, 0.0518).

Now we want to estimate $R(u,v) = P(Y < X|u,v)$, when $u = 180$ and $v = 215$. These are the median values of the covariates of the respective groups, namely for X and Y. Here we have only one covariate each for both the groups. We obtain the MLEs of the unknown parameters, and they are as follows: $\widehat{\alpha} = 5.4318$, $\widehat{\theta}_1 = 3.9597$, $\widehat{\theta}_2 = 0.0406$, $\widehat{\beta}_1 = -0.0518$, $\widehat{\beta}_2 = -0.0372$. Hence we obtain the MLE of $R(u,v)$ as 0.0326. To compute the Bayes estimate of $R(u,v) = P(Y < X|u,v)$, we have further assumed $\beta_1 \sim$ N(0,0.01) and $\beta_2 \sim$ N(0,0.01). Based on the above assumption, using Algorithm 2, we obtain the Bayes estimate of $R(u,v) = P(Y < X|u,v)$ as 0.0298 and the associated 95% HPD credible interval as (0.0118,0.0594).

For illustrative purposes, we obtain the Bayes estimates of R and $R(u,v)$ without the assumption $\alpha_1 = \alpha_2$ of the same data set. We have already observed that without the covariates, the MLEs of the unknown parameters are as follows: $\widehat{\alpha}_1 = 4.4439$, $\widehat{\lambda}_1 = 22.8226$, $\widehat{\alpha}_2 = 5.9646$ and $\widehat{\lambda}_2 = 0.7434$. Hence, the MLE of R based on (1.3) can be obtained as 0.0131. In presence of the covariates, the MLEs of the unknown parameters are as follows: $\widehat{\alpha}_1 = 4.4724$, $\widehat{\theta}_1 = 3.3638$, $\widehat{\beta}_1 = -0.0453$, $\widehat{\alpha}_2 = 5.9982$, $\widehat{\theta}_2 = 0.0286$, $\widehat{\beta}_2 = -0.0449$. The MLE of $R(u,v)$ becomes 0.0234. To compute the Bayes estimate of $R(u,v)$, it is assumed that $\pi_{21}(\alpha_1) \sim$ Gamma((a_{21}, b_{21}), $\pi_{22}(\alpha_2) \sim$ Gamma((a_{22}, b_{22}) and $a_{21} = b_{21} = a_{22} = b_{22} \approx 0$. The other prior distributions have taken same as before. Using the algorithm similar to Algorithm 2, based on 10000 samples, we obtain the Bayes estimate of $R(u,v)$, for $u = 180$, $v = 250$ as 0.0289, and the associated 95% HPD credible interval as (0.0078, 0.0634).

1.7 Conclusions

In this paper we have considered the Bayesian inference of the stress-strength parameter $R = P(Y < X)$, when both X and Y have Weibull distribution. We have considered different cases namely when the two shape parameters are equal and when they are different. We have also addressed the issue when both X and Y may have some covariates also. We have assumed fairly general priors on the shape and scale parameters. In most of the cases the Bayes estimate and the corresponding HPD credible interval cannot be obtained in explicit form and we propose to use Gibbs sampling and importance sampling methods to compute simulation consistent Bayes estimate and the associate HPD credible interval. We have demonstrated different methods using one data set, and it is observed that the proposed methods work well in practice.

Acknowledgements:

The author would like to thank one unknown referee for his/ her many constructive suggestions which have helped to improve the paper significantly.

Appendix

Proof of Theorem 3: Note that

$$\ln \pi_3(\alpha|\beta_1,\beta_2,data) = const. + \ln\phi(\alpha) + (m+n)\ln\alpha + \alpha\ln S - (m+a_1)\ln(b_0 + \sum_{i=1}^{m} x_i^{\alpha} e^{\beta_1{}^T u_i}) - (n+a_2)\ln(b_0 + \sum_{j=1}^{n} y_j^{\alpha} e^{\beta_2{}^T v_j}).$$

It has been shown in the proof of Theorem 2 of Kundu (2008), that

$$\frac{d^2}{d\alpha^2}\ln(b_0 + \sum_{i=1}^{m} x_i^{\alpha} e^{\beta_1{}^T u_i}) \geq 0 \ \text{ and } \ \frac{d^2}{d\alpha^2}\ln(b_0 + \sum_{j=1}^{n} y_j^{\alpha} e^{\beta_2{}^T v_j}) \geq 0.$$

Since $\pi(\alpha)$ is log-concave, it immediately follows that

$$\frac{d^2}{d\alpha^2}\ln\pi_3(\alpha|\beta_1,\beta_2,data) \leq 0.$$

References

1. Aarset, M.V. (1987), "How to identify a bathtub failure rate?", *IEEE Transactions on Reliability*, vol. 36, 106 - 108.
2. Berger, J.O. and Sun, D. (1993), "Bayesian analysis for the Poly-Weibull distribution", *Journal of the American Statistical Association*, vol. 88, 1412 - 1418.
3. Birnbaum, Z.W. (1956), "On a use if Mann-Whitney statistics", *Proceedings of Third Berkeley Symposium in Mathematical Statistics and Probability*, vol. 1, 13 - 17, University of California Press, Berkeley, CA.
4. Chen, M.H. and Shao, Q.M. (1999), "Monte Carlo estimation of Bayesian credible and HPD intervals", *Journal of Computational and Graphical Statistics*, vol. 8, 69 - 92.
5. Devroye, L. (1984), "A simple algorithm for generating random variables using log-concave density function", *Computing*, vol. 33, 247 - 257.
6. Duncan, A.J. (1986), *Quality Control and Industrial Statistics*, 5-th edition, Richard D. Irwin, Homewood, Ill.
7. Guttman, I., Johnson, R.A., Bhattacharya, G.K. and Reiser, B. (1988), "Confidence limits for stress-strength models with explanatory variables", *Technometrics*, vol. 30, 161 - 168.
8. Guttman, I. and Papandonatos, G.D. (1997), "A Bayesian approach to a reliability problem; theory, analysis and interesting numerics", *Canadian Journal of Statistics*, vol. 25, 143 - 158.
9. Kizilaslan, F. and Nadar, M. (2016), "Estimation of reliability in a multicomponent stress-strength model based on a bivariate Kumaraswamy distribution", *Statistical Papers*, to appear, DOI: 10.1007/s00362-016-0765-8.
10. Kotz, S., Lumelskii, Y. and pensky, M. (2003), *The stress-strength model and its generalizations*, World Scientific Press, Singapore.
11. Kundu, D. (2008), "Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring", *Technometrics*, vol. 50, 144 - 154.
12. Kundu, D. and Gupta, R.D. (2006), "Estimation of $P(Y < X)$ for Weibull distribution", *IEEE Transactions on Reliability Theory*, vol. 55, 270 - 280.
13. Kundu, D. and Pradhan, B. (2011), "Bayesian analysis of progressively censored competing risks data", *Sankhya*, Ser. B., vol. 73, 276 - 296.

14. Kundu, D. and Raqab, M.Z. (2009), "Estimation of $R = P(Y < X)$ for three parameter Weibull distribution", *Statistics and Probability Letters*, vol. 79, 1839 - 1846.
15. Murthy, D.N.P., Xie, M. and Jiang, R. (2004), *Weibull Models*, John Wiley and Sons, New York.
16. Pena, E.A. and Gupta, A.K. (1990), "Bayes estimation for the Marshall-Olkin exponential distribution", *Journal of the Royal Statistical Society*, Ser. B, vol. 52, 379 - 389.
17. Ventura, L. and Racugno, W. (2011), "Recent advances on Bayesian inference for $P(X < Y)$", *Bayesian Analysis*, vol. 6, 1 - 18.
18. Weerahandi, S. and Johnson, R.A. (1992), "Testing reliability in a stress-strength model when X and Y are normally distributed", *Technometrics*, vol. 34, 83 - 91.
19. Zhou, W. (2008), "Statistical inference for $P(X < Y)$", *Statistics in Medicine*, vol. 27, 257 - 279.