Statistical Theory and Related Fields

ISSN: 2475-4269 (Print) 2475-4277 (Online) Journal homepage: https://www.tandfonline.com/loi/tstf20

Bayesian analysis of series system with dependent causes of failure

Ancha Xu & Shirong Zhou

To cite this article: Ancha Xu & Shirong Zhou (2017) Bayesian analysis of series system with dependent causes of failure, Statistical Theory and Related Fields, 1:1, 128-140, DOI: 10.1080/24754269.2017.1348708

To link to this article: https://doi.org/10.1080/24754269.2017.1348708

Published online: 20 Jul 2017.

STATISTICAL THEORY AND RELATED FIELDS, 2017
VOL. 1, NO. 1, 128–140
https://doi.org/10.1080/24754269.2017.1348708

Bayesian analysis of series system with dependent causes of failure

Ancha Xu and Shirong Zhou

Department of Statistics, College of Mathematics and Information Science, Wenzhou University, Zhejiang, China

ABSTRACT
Most studies of series system assume the causes of failure are independent, which may not hold in practice. In this paper, dependent causes of failure are considered by using a Marshall–Olkin bivariate Weibull distribution. We derived four reference priors based on several grouping orders. Gibbs sampling combined with the rejection sampling algorithm and Metropolis–Hastings algorithm is developed to obtain the estimates of the unknown parameters. The proposed approach is compared with the maximum-likelihood method via simulation. We find that the root-mean-squared errors of the Bayesian estimates are much smaller for the case of small sample size, and that the coverage probabilities of the Bayesian estimates are much closer to the nominal levels. Finally, a real data-set is analysed for illustration.

ARTICLE HISTORY
Received 2 January 2017
Accepted 27 June 2017

KEYWORDS
Reference prior; Bayesian analysis; Weibull distribution; Gibbs sampling; Metropolis–Hastings algorithm

1. Introduction

In reliability and survival analysis, it is quite common that a series system or an individual might fail because of one of several causes of failure (or competing risks). A researcher is interested in a specific cause in the presence of other causes. In the statistical literature, this problem is known as the competing risks model, and the series system is a typical competing risk model. Data for a competing risks model may consist of failure times and other variables indicating the causes of failure, which may be dependent or independent. Most of the literature assumes the causes of failure are independent. See, for example, Kozumi (2004), Kundu, Kannan, and Balakrishnan (2003), Pareek, Kundu, and Kumar (2009), Mazucheli and Achcar (2011), Cramer and Schmiedt (2011), Xu and Tang (2011), Xu, Basu, and Tang (2014) and AL-Hussaini, Abdel-Hamid, and Hashem (2015).

However, the independence assumption among the causes of failure is often unrealistic. For instance, in a tug of war, failure of a player causes additional pressure on the team and poses an increased risk of failure to the remaining members. In other words, a positive dependence between failure times is imposed. Thus, the causes of failure being dependent are more practical. Wang and Ghosh (2003) studied dependent competing risks model with absolutely continuous bivariate exponential lifetime distribution under two non-informative priors (Laplace prior and Jeffreys prior). Lindqvist and Skogsrud (2009) have focused on modelling dependent competing risks in reliability by considering first the passage times of Wiener processes. Dijoux and Gaudoin (2009) proposed an alert-delay model which is a new model of dependent competing risks for maintenance and reliability analysis. Dimitrova, Haberman, and Kaishev (2013) demonstrated how copula functions can be applied in modelling dependence between lifetime random variables in the context of competing risks and studied the impact of removing one or more causes of death on the overall survival. When there is a positive probability of simultaneous failure, the most widely used model is the Marshall–Olkin bivariate lifetime model. The Marshall–Olkin bivariate exponential (MOBE) model, proposed by Marshall and Olkin (1967), has been used extensively to analyse two dependent causes of failure. Guan, Tang, and Xu (2013) considered the objective Bayesian analysis of dependent competing risks model using MOBE distribution. However, the MOBE distribution is very limited, because the hazard function is constant or the marginal function is strictly decreasing. Then the Marshall–Olkin bivariate Weibull (MOBW) distribution which was raised by Marshall and Olkin (1967) can make up this deficiency, and has been long enjoyed popularity in reliability (see Kundu & Dey, 2009; Kundu & Gupta, 2013; and the references therein). Feizjavadian and Hashemi (2015) developed the maximum-likelihood estimators (MLEs) and approximated MLEs of the unknown parameters when the lifetime of the two causes of failure follows MOBW distribution.

In this paper, Bayesian method will be used to analyse dependent competing risks model when the MOBW distribution is assumed. Both subjective Bayesian analysis and objective Bayesian analysis are popular in practice. Subjective Bayesian analysis is based on informative priors, i.e. conjugate prior. In such a case, the prior information is used to specify the hyperparameters in the prior distribution. However, prior distribution is not

CONTACT Ancha Xu xuancha@wzu.edu.cn

easy to elicit, especially in complex models. Then objective Bayesian analysis by using non-informative priors is preferred. The three most used non-informative priors are Laplace prior (Laplace, 1812), Jeffreys prior (Jeffreys, 1961) and reference prior (Bernardo, 1979). Laplace prior uses a constant prior distribution for the unknown parameters, which is quite easy to specify. However, it is not invariant under reparameterisation, and usually leads to an improper posterior distribution. Jeffreys prior, which is proportional to the square root of the determinant of the Fisher information matrix, has been proved to be successful in single-parameter problems. However, Jeffreys prior is often seriously deficient in multi-parameter problems. To overcome the deficiencies of the Jeffreys prior, Bernardo (1979) proposed the reference prior which works well in multi-parameter problems. It is invariant under reparameterisation and typically produces proper posteriors. The reference prior is the Jeffreys prior in usual single-parameter problems. This approach is very successful in various practical problems (see Xu & Tang, 2010; Xu, Tang, & Sun, 2015). Thus, we will consider noninformative priors for the unknown parameters in the proposed model. To the best of our knowledge, we are the first to develop Bayesian approach, especially objective Bayesian method, to analyse dependent competing risks model using the MOBW distribution. In particular, the reference priors are formally derived, and the properties of these reference priors are studied.

The paper is organised as follows. In Section 2, we describe the MOBW model. The Fisher information matrices under the original and transformed parameters are presented in Section 3. Section 4 is devoted to derive four reference priors under different grouping orders. In Section 5, the propriety of the posterior distribution under different reference priors are proved, and the Gibbs sampling procedures are given to obtain the Bayesian estimates. In Section 6, a simulation is given for illustration. A real data-set is analysed in Section 7. Finally, some concluding remarks are made in Section 8.

2. Marshall–Olkin bivariate Weibull distribution

The MOBW model can be described as follows. It is assumed that U_0, U_1 and U_2 are three independent random variables, and

$$U_0 \sim WE(\alpha, \lambda_0),\ U_1 \sim WE(\alpha, \lambda_1),\ U_2 \sim WE(\alpha, \lambda_2),$$

where $\sim$ means following in distribution and $WE(\alpha, \lambda)$ denotes a Weibull distribution with the parameters $\alpha > 0$ and $\lambda > 0$. For a $WE(\alpha, \lambda)$ distribution, the probability density function (PDF), the cumulative distribution function (CDF) and the survival function (SF) for $x > 0$ are

$$f_{WE}(x;\alpha,\lambda) = \alpha\lambda x^{\alpha-1}e^{-\lambda x^{\alpha}},$$
$$F_{WE}(x;\alpha,\lambda) = 1 - e^{-\lambda x^{\alpha}} \text{ and } S_{WE}(x;\alpha,\lambda) = e^{-\lambda x^{\alpha}},$$

respectively.

Let $X_1 = \min\{U_0, U_1\}$ and $X_2 = \min\{U_0, U_2\}$, then (X_1, X_2) has the MOBW distribution with the parameters $(\alpha, \lambda_0, \lambda_1, \lambda_2)$, and it will be denoted as MOBW $(\alpha, \lambda_0, \lambda_1, \lambda_2)$. Thus, the joint SF of MOBW $(\alpha, \lambda_0, \lambda_1, \lambda_2)$ can be written as

$$\begin{aligned}
&S_{X_1,X_2}(x_1,x_2)\\
&= P(X_1 > x_1, X_2 > x_2)\\
&= P(U_1 > x_1, U_2 > x_2, U_0 > \max\{x_1,x_2\})\\
&= S_{WE}(x_1;\alpha,\lambda_1)S_{WE}(x_2;\alpha,\lambda_2)S_{WE}(\max\{x_1,x_2\};\alpha,\lambda_0)\\
&= \begin{cases} S_{WE}(x_1;\alpha,\lambda_1)S_{WE}(x_2;\alpha,\lambda_0+\lambda_2), & \text{if } x_1 < x_2,\\ S_{WE}(x_1;\alpha,\lambda_0+\lambda_1)S_{WE}(x_2;\alpha,\lambda_2), & \text{if } x_1 > x_2,\\ S_{WE}(x;\alpha,\lambda_0+\lambda_1+\lambda_2), & \text{if } x_1 = x_2 = x. \end{cases}
\end{aligned} \tag{2.1}$$

Therefore, the joint PDF of X_1 and X_2 can be written as

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} f_1(x_1,x_2), & \text{if } x_1 < x_2,\\ f_2(x_1,x_2), & \text{if } x_1 > x_2,\\ f_0(x), & \text{if } x_1 = x_2 = x, \end{cases} \tag{2.2}$$

where

$$f_1(x_1,x_2) = f_{WE}(x_1;\alpha,\lambda_1)f_{WE}(x_2;\alpha,\lambda_0+\lambda_2),$$

$$f_2(x_1,x_2) = f_{WE}(x_1;\alpha,\lambda_0+\lambda_1)f_{WE}(x_2;\alpha,\lambda_2),$$

$$f_0(x) = \frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2}f_{WE}(x;\alpha,\lambda_0+\lambda_1+\lambda_2).$$

From the above equations, we note that the MOBW distribution has both an absolute continuous part ($f_1(x_1, x_2)$ and $f_2(x_1, x_2)$) and a singular part ($f_0(x)$). For more details about the MOBW distribution, we can refer to Bemis, Bain, and Higgins (1972) and Kundu and Dey (2009).

3. Data and Fisher information matrix

3.1. Data and likelihood function

Suppose that a series system (or a disease) has two causes of failure. The lifetime of the jth cause of failure is X_j ($j = 1, 2$), and $(X_1, X_2) \sim$ MOBW $(\alpha, \lambda_0, \lambda_1, \lambda_2)$. Then the distribution of the lifetime of the series system $T = \min(X_1, X_2)$ is $WE(\alpha, \lambda_0 + \lambda_1 + \lambda_2)$. Assume that there are n series systems in a lifetime experiment. Let (X_{1i}, X_{2i}) be the lifetime of the two causes of failure in the ith system. Therefore, the observed data are $(T_i, \delta_{1i}, \delta_{2i})$, $i = 1, 2, \ldots, n$, where $T_i = \min(X_{1i}, X_{2i})$, $\delta_{1i} = I(X_{1i} < X_{2i})$, $\delta_{2i} = I(X_{1i} > X_{2i})$, $I(A)$ denotes the indicator function of event A. Let $\lambda = \lambda_0 + \lambda_1 + \lambda_2$, we can obtain the following facts easily.

Lemma 3.1: *For $i = 1, \dots, n$, we have*

(1) $X_{1i} \sim WE(\alpha, \lambda_0 + \lambda_1)$, $X_{2i} \sim WE(\alpha, \lambda_0 + \lambda_2)$,
(2) T_i and $(\delta_{1i}, \delta_{2i})$ are independent,
(3) $(\delta_{1i}, \delta_{2i}) \sim multinomial\left(1; \frac{\lambda_1}{\lambda}, \frac{\lambda_2}{\lambda}\right)$.

Denote $n_1 = \sum_{i=1}^n \delta_{1i}$, $n_2 = \sum_{i=1}^n \delta_{2i}$ and $n_0 = \sum_{i=1}^n (1 - \delta_{1i} - \delta_{2i})$, given the observed data $(T_i, \delta_{1i}, \delta_{2i})$, $i = 1, 2, \dots, n$, the likelihood function is

$$L_1 = \alpha^n \lambda_0^{n_0} \lambda_1^{n_1} \lambda_2^{n_2} \exp\left\{-\lambda \sum_{i=1}^n T_i^\alpha\right\} \prod_{i=1}^n T_i^{\alpha-1}. \quad (3.1)$$

Then the log-likelihood function can be written as

$$l_1 = \ln L_1 = n \ln \alpha + n_0 \ln \lambda_0 + n_1 \ln \lambda_1 + n_2 \ln \lambda_2 + (\alpha - 1)\sum_{i=1}^n \ln T_i - (\lambda_0 + \lambda_1 + \lambda_2)\sum_{i=1}^n T_i^\alpha.$$

3.2. Fisher information matrix

The second-order partial derivatives of the log-likelihood function are as follows:

$$\frac{\partial^2 l_1}{\partial \alpha^2} = -\frac{n}{\alpha^2} - (\lambda_0 + \lambda_1 + \lambda_2)\sum_{i=1}^n T_i^\alpha (\ln T_i)^2,$$
$$\frac{\partial^2 l_1}{\partial \alpha \partial \lambda_q} = -\sum_{i=1}^n T_i^\alpha \ln T_i, \quad \frac{\partial^2 l_1}{\partial \lambda_q^2} = -\frac{n_q}{\lambda_q^2}, \quad \frac{\partial^2 l_1}{\partial \lambda_q \partial \lambda_\rho} = 0,$$

where $q = 0, 1, 2$, $\rho = 0, 1, 2$, $q \neq \rho$. Denote $Y_i = \lambda T_i^\alpha$, $i = 1, 2, \dots, n$, then Y_i follows the exponential distribution with hazard rate 1. For $\nu \geq 1$, let $r_\nu = \int_0^\infty (\ln y)^\nu e^{-y} dy$, which is the νth moment of $\ln Y_i$. Then,

$$E\left(T_i^\alpha \ln T_i\right) = \frac{1}{\alpha\lambda}(1 + r_1 - \ln\lambda),$$
$$E\left[T_i^\alpha (\ln T_i)^2\right] = \frac{1}{\alpha^2\lambda}\left[2r_1 + r_2 - 2(r_1 + 1)\ln\lambda + (\ln\lambda)^2\right].$$

After some algebraic calculations, we get

$$-E\left(\frac{\partial^2 l_1}{\partial \alpha^2}\right) = \frac{nk(\lambda)}{\alpha^2},$$
$$-E\left(\frac{\partial^2 l_1}{\partial \alpha \partial \lambda_q}\right) = \frac{n}{\alpha\lambda}(1 + r_1 - \ln\lambda),$$
$$-E\left(\frac{\partial^2 l_1}{\partial \lambda_q^2}\right) = \frac{n}{\lambda\lambda_q}, \quad -E\left(\frac{\partial^2 l_1}{\partial \lambda_q \partial \lambda_\rho}\right) = 0,$$

where $k(x) = 1 + 2r_1 + r_2 - 2(r_1 + 1)\ln x + (\ln x)^2$, $q = 0, 1, 2$, $\rho = 0, 1, 2$, $q \neq \rho$. Thus, the Fisher information matrix of $(\alpha, \lambda_0, \lambda_1, \lambda_2)$ has the following form:

$$\Sigma_0 = \begin{pmatrix} \frac{nk(\lambda)}{\alpha^2} & \frac{n}{\alpha\lambda}(1 + r_1 - \ln\lambda) & \frac{n}{\alpha\lambda}(1 + r_1 - \ln\lambda) & \frac{n}{\alpha\lambda}(1 + r_1 - \ln\lambda) \\ \frac{n}{\alpha\lambda}(1 + r_1 - \ln\lambda) & \frac{n}{\lambda\lambda_0} & 0 & 0 \\ \frac{n}{\alpha\lambda}(1 + r_1 - \ln\lambda) & 0 & \frac{n}{\lambda\lambda_1} & 0 \\ \frac{n}{\alpha\lambda}(1 + r_1 - \ln\lambda) & 0 & 0 & \frac{n}{\lambda\lambda_2} \end{pmatrix}.$$

3.3. Reparameterisation

In practice, the engineers may be interested in the hazard rate function of the series system, which is $\alpha\lambda t^{\alpha-1}$. λ affects the scale of the hazard rate function, and α will determine whether the hazard rate function is decreasing $(0 < \alpha < 1)$, constant $(\alpha = 1)$ or increasing $(\alpha > 1)$. Thus, we take the following transformation:

$$\theta_1 \equiv \lambda = \lambda_0 + \lambda_1 + \lambda_2, \theta_2 = \frac{\lambda_1}{\lambda}, \theta_3 = \frac{\lambda_2}{\lambda},$$

where θ_2 and θ_3 are the probabilities that the failure of series system is due to the first and second causes of failure, respectively. The transformation from $(\alpha, \lambda_0, \lambda_1, \lambda_2)$ to $(\alpha, \theta_1, \theta_2, \theta_3)$ is one-to-one with the inverse transformation

$$\alpha = \alpha, \lambda_1 = \theta_1\theta_2, \lambda_2 = \theta_1\theta_3, \lambda_0 = \theta_1(1 - \theta_2 - \theta_3).$$

After reparameterisation, the likelihood function (3.1) becomes

$$L_2 = \alpha^n \theta_1^n \theta_2^{n_1} \theta_3^{n_2} (1 - \theta_2 - \theta_3)^{n_0} \exp \times \left\{-\theta_1 \sum_{i=1}^n T_i^\alpha\right\} \prod_{i=1}^n T_i^{\alpha-1}. \quad (3.2)$$

The Jacobian matrix of the transformation is

$$H = \frac{\partial(\alpha, \lambda_0, \lambda_1, \lambda_2)}{\partial(\alpha, \theta_1, \theta_2, \theta_3)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - \theta_2 - \theta_3 & -\theta_1 & -\theta_1 \\ 0 & \theta_2 & \theta_1 & 0 \\ 0 & \theta_3 & 0 & \theta_1 \end{pmatrix},$$

where $0 < \theta_1 < \infty$, $0 < \theta_2 + \theta_3 < 1$.

Lemma 3.2: *The Fisher information matrix of* $(\alpha, \theta_1, \theta_2, \theta_3)$ *has the following form:*

$$\Sigma = H^{'}\Sigma_0 H = \begin{pmatrix} \frac{nk(\theta_1)}{\alpha^2} & \frac{n(1+r_1-ln\theta_1)}{\alpha\theta_1} & 0 & 0 \\ \frac{n(1+r_1-ln\theta_1)}{\alpha\theta_1} & \frac{n}{\theta_1^2} & 0 & 0 \\ 0 & 0 & \frac{n}{\theta_2}+\frac{n}{1-\theta_2-\theta_3} & \frac{n}{1-\theta_2-\theta_3} \\ 0 & 0 & \frac{n}{1-\theta_2-\theta_3} & \frac{n}{\theta_3}+\frac{n}{1-\theta_2-\theta_3} \end{pmatrix}.$$

Due to the practical meanings of $(\alpha, \theta_1, \theta_2, \theta_3)$, all the derivations will be based on the new parameters in the following sections. If one is interested in the original parameters, the proposed method in this paper is also suitable.

4. Reference priors

Berger and Bernardo (1992) developed a general algorithm to derive the reference prior. The algorithm needs to divide the parameters into several groups according to the importance of interest. Jeffreys prior treats all the parameters equally, and is shown to yield a decidedly inferior posterior distribution in multivariate-parameter problems (Yang & Berger, 1994). Sometimes, Jeffreys prior will result in an inconsistent estimator of the parameters. The benefit of grouping the parameters is that the prior can be obtained in terms of inferential importance of parameters, and usually results in a proper posterior distribution. Thus, before deriving the reference priors, we should divide the parameters $\{\alpha, \theta_1, \theta_2, \theta_3\}$ according to the inferential importance. For example, the notation $\{\alpha, (\theta_1, \theta_2, \theta_3)\}$ will be used to represent the case that the parameters are divided into two groups, with α being the most important and $\theta_1, \theta_2, \theta_3$ being of equal importance. Similarly, $\{\alpha, \theta_1, (\theta_2, \theta_3)\}$ represents that the parameters are separated into three groups, with α being the most important and (θ_2, θ_3) being the least important. Ghosh and Mukerjee (1991) and Berger and Bernardo (1992) suggested that switching the role of the parameters of interest and nuisance parameters sometimes gives a reasonable reference prior. We will consider the grouping orders $\{(\alpha, \theta_1), (\theta_2, \theta_3)\}$, $\{(\theta_2, \theta_3), (\alpha, \theta_1)\}$, $\{\alpha, (\theta_1, \theta_2, \theta_3)\}$, $\{(\theta_2, \theta_3), \alpha, \theta_1\}$, $\{\alpha, \theta_1, \theta_2, \theta_3\}$, $\{\theta_3, \theta_2, \theta_1, \alpha\}$, $\{\theta_1, (\alpha, \theta_2, \theta_3)\}$, $\{(\theta_2, \theta_3), \theta_1, \alpha\}$.

Table 1. Possible reference prior of $(\alpha, \theta_1, \theta_2, \theta_3)$.

Ordered grouping	Reference prior for $(\alpha, \theta_1, \theta_2, \theta_3)$	Reference prior for $(\alpha, \lambda_0, \lambda_1, \lambda_2)$
$\{(\alpha,\theta_1),(\theta_2,\theta_3)\}$	$\omega_1(\alpha,\theta_1,\theta_2,\theta_3)$	$\pi_1(\alpha,\lambda_0,\lambda_1,\lambda_2)$
$\{(\theta_2,\theta_3),(\alpha,\theta_1)\}$	$\omega_1(\alpha,\theta_1,\theta_2,\theta_3)$	$\pi_1(\alpha,\lambda_0,\lambda_1,\lambda_2)$
$\{(\theta_2,\theta_3),\alpha,\theta_1\}$	$\omega_1(\alpha,\theta_1,\theta_2,\theta_3)$	$\pi_1(\alpha,\lambda_0,\lambda_1,\lambda_2)$
$\{\alpha,(\theta_1,\theta_2,\theta_3)\}$	$\omega_1(\alpha,\theta_1,\theta_2,\theta_3)$	$\pi_1(\alpha,\lambda_0,\lambda_1,\lambda_2)$
$\{\alpha,\theta_1,\theta_2,\theta_3\}$	$\omega_2(\alpha,\theta_1,\theta_2,\theta_3)$	$\pi_2(\alpha,\lambda_0,\lambda_1,\lambda_2)$
$\{\theta_3,\theta_2,\theta_1,\alpha\}$	$\omega_3(\alpha,\theta_1,\theta_2,\theta_3)$	$\pi_3(\alpha,\lambda_0,\lambda_1,\lambda_2)$
$\{\theta_1,(\alpha,\theta_2,\theta_3)\}$	$\omega_4(\alpha,\theta_1,\theta_2,\theta_3)$	$\pi_4(\alpha,\lambda_0,\lambda_1,\lambda_2)$
$\{(\theta_2,\theta_3),\theta_1,\alpha\}$	$\omega_4(\alpha,\theta_1,\theta_2,\theta_3)$	$\pi_4(\alpha,\lambda_0,\lambda_1,\lambda_2)$

Theorem 4.1: *The possible grouping ordering reference priors of* $(\alpha, \theta_1, \theta_2, \theta_3)$ *are listed in Table 1, and*

$$\omega_1(\alpha,\theta_1,\theta_2,\theta_3) = \frac{1}{\sqrt{\alpha^2\theta_1^2\theta_2\theta_3(1-\theta_2-\theta_3)}},$$
$$\pi_1(\alpha,\lambda_0,\lambda_1,\lambda_2) = \frac{1}{\sqrt{\alpha^2\lambda^3\lambda_0\lambda_1\lambda_2}},$$
$$\omega_2(\alpha,\theta_1,\theta_2,\theta_3) = \frac{1}{\sqrt{\alpha^2\theta_1^2\theta_2\theta_3(1-\theta_2)(1-\theta_2-\theta_3)}},$$
$$\pi_2(\alpha,\lambda_0,\lambda_1,\lambda_2) = \frac{1}{\sqrt{\alpha^2\lambda^2\lambda_0\lambda_1\lambda_2(\lambda_0+\lambda_2)}},$$
$$\omega_3(\alpha,\theta_1,\theta_2,\theta_3) = \frac{1}{\sqrt{\alpha^2\theta_1^2\theta_2\theta_3(1-\theta_3)(1-\theta_2-\theta_3)k(\theta_1)}},$$
$$\pi_3(\alpha,\lambda_0,\lambda_1,\lambda_2) = \frac{1}{\sqrt{\alpha^2\lambda^2\lambda_0\lambda_1\lambda_2(\lambda_0+\lambda_1)k(\lambda)}},$$
$$\omega_4(\alpha,\theta_1,\theta_2,\theta_3) = \frac{1}{\sqrt{\alpha^2\theta_1^2\theta_2\theta_3(1-\theta_2-\theta_3)k(\theta_1)}},$$
$$\pi_4(\alpha,\lambda_0,\lambda_1,\lambda_2) = \frac{1}{\sqrt{\alpha^2\lambda^3\lambda_0\lambda_1\lambda_2 k(\lambda)}}.$$

See the proof in the Appendix 1.

Remark 4.1: From Table 1, it can be noted that there are four different reference priors for the eight grouping orders. Actually, the number of grouping orders is much greater than eight. We choose the eight grouping orders because they have some practical meanings. For example, the grouping order $\{(\theta_2, \theta_3), (\alpha, \theta_1)\}$ can reflect our interest of the failure probabilities of series system due to the first and second causes of failure. While the grouping order $\{\alpha, (\theta_1, \theta_2, \theta_3)\}$ indicates that more attention is paid to the shape parameter of the MOBW distribution. Of course, someone could choose some other grouping orders. However, the closed form of the reference prior may not be obtained. For instance, when the grouping order is $\{\theta_1, \theta_2, \theta_3, \alpha\}$, the reference prior does not have closed form. If all the parameters are interested, the overall objective priors (Berger, Bernardo, & Sun, 2015) may be considered. However, the derivation of the overall objective priors is much different from the reference priors. and it is not the scope of our paper. In the following sections, we will study the properties of the four reference priors in detail.

5. Posterior analysis

5.1. Propriety of the posteriors

Since the reference priors listed in Table 1 are improper, we need to check the propriety of the posteriors before doing the Bayesian analysis.

Theorem 5.1: *When $n > 1$, the posterior distributions of $(\alpha, \theta_1, \theta_2, \theta_3)$ based on the priors $\omega_1(\alpha, \theta_1, \theta_2, \theta_3)$, $\omega_2(\alpha, \theta_1, \theta_2, \theta_3)$, $\omega_3(\alpha, \theta_1, \theta_2, \theta_3)$, $\omega_4(\alpha, \theta_1, \theta_2, \theta_3)$ are all proper.*

See the proof in the Appendix 2. Theorem 5.1 shows the conditions that the reference priors can be used when at least two failures are observed in the experiment.

5.2. Bayesian estimation

For the sake of convenience, we write the reference priors in a general way:

$$\omega(\alpha, \theta_1, \theta_2, \theta_3) = \frac{1}{\sqrt{\alpha^2\theta_1^2\theta_2\theta_3(1-\theta_2-\theta_3)(1-\theta_2)^{c_1}(1-\theta_3)^{c_2}\,(k(\theta_1))^{c_3}}}, \tag{5.1}$$

where c_1, c_2 and c_3 take the particular values when one of the reference priors is used. For example, when $\omega_1(\alpha, \theta_1, \theta_2, \theta_3)$ is used, $c_1 = c_2 = c_3 = 0$. Based on (3.2) and (5.1), the joint posterior density of $(\alpha, \theta_1, \theta_2, \theta_3)$ is

$$\begin{aligned} p(\alpha, \theta_1, \theta_2, \theta_3|\text{Data}) &\propto L_2 \times \omega(\alpha, \theta_1, \theta_2, \theta_3) \\ &\propto \alpha^{n-1}\theta_1^{n-1}\theta_2^{n_1-\frac{1}{2}}\theta_3^{n_2-\frac{1}{2}} \\ &\quad \times (1-\theta_2-\theta_3)^{n_0-\frac{1}{2}} \exp\left\{-\theta_1\sum_{i=1}^{n} T_i^{\alpha}\right\}\left(\prod_{i=1}^{n} T_i^{\alpha-1}\right) \\ &\quad \times (1-\theta_2)^{-c_1/2}(1-\theta_3)^{-c_2/2}\,(k(\theta_1))^{-c_3/2}. \end{aligned} \tag{5.2}$$

From (5.2), we see that (α, θ_1) and (θ_2, θ_3) are independent, which will make posterior sampling efficient. To obtain the Bayesian estimation of the parameters, the Gibbs sampling procedure can be implemented with the full conditional posterior distributions as follows.

(1) The conditional posterior density function of α, $p(\alpha|\theta_1, Data)$, is proportional to

$$\alpha^{n-1}\exp\left\{-\theta_1\sum_{i=1}^{n} T_i^{\alpha}\right\}\prod_{i=1}^{n} T_i^{\alpha},$$

which is log-concave. Thus, the adaptive rejection sampling algorithm (Gilks & Wild, 1992) can be used to generate the posterior samples of α.

(2) The conditional posterior density function of θ_1, $p(\theta_1|\alpha, \text{Data})$, is proportional to

$$\theta_1^{n-1}\exp\left\{-\theta_1\sum_{i=1}^{n} T_i^{\alpha}\right\}(k(\theta_1))^{-c_3/2}. \tag{5.3}$$

- When the reference priors $\omega_1(\alpha, \theta_1, \theta_2, \theta_3)$ and $\omega_2(\alpha, \theta_1, \theta_2, \theta_3)$ are used, $c_3 = 0$. Thus, given α and the observed data, the conditional posterior distribution of θ_1 is gamma distribution with the shape parameter n and the scale parameter $\sum_{i=1}^{n} T_i^{\alpha}$. Denote the gamma distribution as $\Gamma(n, \sum_{i=1}^{n} T_i^{\alpha})$.
- When the reference priors $\omega_3(\alpha, \theta_1, \theta_2, \theta_3)$ and $\omega_4(\alpha, \theta_1, \theta_2, \theta_3)$ are utilised, $c_3 = 1$. The rejection sampling algorithm is used to generate the posterior samples of θ_1. We choose $\Gamma(n, \sum_{i=1}^{n} T_i^{\alpha})$ as the proposal distribution. The algorithm proceeds in two steps:

Table 2. The RB(%), CPs and RMSEs of the parameters when $(\alpha, \lambda_0, \lambda_1, \lambda_2) = (0.5, 0.7, 1, 1.5)$.

n	Method		α	θ_1	θ_2	θ_3
15	MLE	RMSE	0.1337	1.8909	0.1196	0.1241
		RB(%)	9.9284	10.4440	0.1600	2.5209
		CP	0.9480	0.9710	0.8730	0.9410
	ω_1	RMSE	0.1310	2.0187	0.1088	0.1147
		RB(%)	9.0203	11.2454	0.7515	4.9180
		CP	0.9400	0.9300	0.9560	0.9600
	ω_2	RMSE	0.1310	2.0187	0.1128	0.1165
		RB(%)	9.0203	11.2454	3.8310	6.2795
		CP	0.9400	0.9300	0.9490	0.9550
	ω_3	RMSE	0.1248	1.7886	0.1084	0.1160
		RB(%)	7.5734	11.2454	0.7515	4.9180
		CP	0.9490	0.9530	0.9560	0.9460
	ω_4	RMSE	0.1248	1.7886	0.1088	0.1147
		RB(%)	7.5734	11.2454	0.7515	4.9180
		CP	0.9490	0.9530	0.9560	0.9600
30	MLE	RMSE	0.0857	0.9554	0.0871	0.0922
		RB(%)	5.1869	9.8592	0.5333	0.8213
		CP	0.9470	0.9640	0.9040	0.9340
	ω_1	RMSE	0.0845	0.9650	0.0830	0.0878
		RB(%)	4.6945	9.6570	0.1905	0.5934
		CP	0.9400	0.9400	0.9430	0.9580
	ω_2	RMSE	0.0845	0.9650	0.0844	0.0882
		RB(%)	4.6945	9.6570	1.4378	1.3578
		CP	0.9400	0.9400	0.9370	0.9440
	ω_3	RMSE	0.0828	0.9139	0.0830	0.0894
		RB(%)	4.0475	9.6570	0.1905	0.5934
		CP	0.9430	0.9450	0.9290	0.9470
	ω_4	RMSE	0.0828	0.9139	0.0830	0.0878
		RB(%)	4.0475	9.6570	0.1905	0.5934
		CP	0.9430	0.9450	0.9430	0.9580
50	MLE	RMSE	0.0614	0.6505	0.0663	0.0686
		RB(%)	3.1057	5.5929	0.0320	0.3648
		CP	0.9410	0.9660	0.9240	0.9610
	ω_1	RMSE	0.0610	0.6559	0.0644	0.0666
		RB(%)	2.8219	5.4257	0.2252	0.4872
		CP	0.9340	0.9460	0.9570	0.9610
	ω_2	RMSE	0.0610	0.6559	0.0652	0.0669
		RB(%)	2.8219	5.4257	1.1799	0.9393
		CP	0.9340	0.9460	0.9530	0.9570
	ω_3	RMSE	0.0601	0.6340	0.0643	0.0673
		RB(%)	2.4622	5.4257	0.2252	0.4872
		CP	0.9480	0.9470	0.9460	0.9560
	ω_4	RMSE	0.0601	0.6340	0.0644	0.0666
		RB(%)	2.4622	5.4257	0.2252	0.4872
		CP	0.9480	0.9470	0.9570	0.9610

Table 3. The RB(%), CPs and SRMSEs of the parameters when $(\alpha, \lambda_0, \lambda_1, \lambda_2) = (1, 0.7, 1, 1.5)$.

n	Method		α	θ_1	θ_2	θ_3
	MLE	RMSE	0.2711	1.9711	0.1186	0.1246
		RB(%)	9.7428	12.5952	0.7147	1.5538
		CP	0.9510	0.9630	0.8770	0.9140
	ω_1	RMSE	0.2654	2.0968	0.1078	0.1147
		RB(%)	8.7483	14.8890	0.0436	4.0388
		CP	0.9500	0.9290	0.9540	0.9650
	ω_2	RMSE	0.2654	2.0968	0.1121	0.1169
		RB(%)	8.7483	14.8890	3.2041	5.4434
15		CP	0.9500	0.9290	0.9480	0.9580
	ω_3	RMSE	0.2517	1.9172	0.1082	0.1171
		RB(%)	7.3402	14.8890	0.0436	4.0388
		CP	0.9520	0.9380	0.9520	0.9550
	ω_4	RMSE	0.2517	1.9172	0.1078	0.1147
		RB(%)	7.3402	14.8890	0.0436	4.0388
		CP	0.9520	0.9380	0.9540	0.9650
	MLE	RMSE	0.1587	0.8689	0.0848	0.0923
		RB(%)	3.5935	8.2261	0.8747	0.6436
		CP	0.9480	0.9650	0.9140	0.9130
	ω_1	RMSE	0.1574	0.8757	0.0807	0.0879
		RB(%)	3.1080	7.9506	0.5156	0.7628
		CP	0.9440	0.9490	0.9500	0.9590
	ω_2	RMSE	0.1574	0.8757	0.0822	0.0884
		RB(%)	3.1080	7.9506	1.0826	1.5091
30		CP	0.9440	0.9490	0.9490	0.9400
	ω_3	RMSE	0.1539	0.8313	0.0807	0.0895
		RB(%)	2.4891	7.9506	0.5156	0.7628
		CP	0.9470	0.9500	0.9430	0.9440
	ω_4	RMSE	0.1539	0.8313	0.0807	0.0879
		RB(%)	2.4891	7.9506	0.5156	0.7628
		CP	0.9470	0.9500	0.9500	0.9590
	MLE	RMSE	0.1116	0.6033	0.0636	0.0694
		RB(%)	2.2050	4.8884	0.3072	0.1813
		CP	0.9550	0.9540	0.9290	0.9580
	ω_1	RMSE	0.1113	0.6034	0.0617	0.0675
		RB(%)	1.8847	4.6502	0.4924	0.6654
		CP	0.9520	0.9410	0.9560	0.9580
	ω_2	RMSE	0.1113	0.6034	0.0625	0.0679
50		RB(%)	1.8847	4.6502	0.5682	1.1268
		CP	0.9520	0.9410	0.9550	0.9520
	ω_3	RMSE	0.1098	0.5866	0.0618	0.0682
		RB(%)	1.5879	4.6502	0.4924	0.6654
		CP	0.9540	0.9380	0.9530	0.9490
	ω_4	RMSE	0.1098	0.5866	0.0617	0.0675
		RB(%)	1.5879	4.6502	0.4924	0.6654
		CP	0.9540	0.9380	0.9560	0.9580

Table 4. The RB(%), CPs and RMSEs of the parameters when $(\alpha, \lambda_0, \lambda_1, \lambda_2) = (2, 0.7, 1, 1.5)$.

n	Method		α	θ_1	θ_2	θ_3
	MLE	RMSE	0.5570	2.8778	0.1170	0.1265
		RB(%)	10.7944	10.9046	0.4160	1.2409
		CP	0.9440	0.9760	0.8820	0.9070
	ω_1	RMSE	0.5308	2.7675	0.1064	0.1162
		RB(%)	9.6667	10.7426	0.2279	3.7543
		CP	0.9310	0.9410	0.9560	0.9600
	ω_2	RMSE	0.5308	2.7675	0.1106	0.1177
		RB(%)	9.6667	10.7426	3.4277	5.1166
15		CP	0.9310	0.9410	0.9520	0.9500
	ω_3	RMSE	0.5079	2.3783	0.1063	0.1187
		RB(%)	8.2546	10.7426	0.2279	3.7543
		CP	0.9430	0.9450	0.9570	0.9420
	ω_4	RMSE	0.5079	2.3783	0.1064	0.1162
		RB(%)	8.2546	10.7426	0.2279	3.7543
		CP	0.9430	0.9450	0.9560	0.9600
	MLE	RMSE	0.3251	0.9553	0.0837	0.0878
		RB(%)	5.0828	9.9498	0.6400	0.4587
		CP	0.9600	0.9630	0.9210	0.9350
	ω_1	RMSE	0.3215	0.9729	0.0797	0.0840
		RB(%)	4.5786	9.7440	0.2921	1.8125
		CP	0.9550	0.9440	0.9640	0.9600
	ω_2	RMSE	0.3215	0.9729	0.0812	0.0846
		RB(%)	4.5786	9.7440	1.3427	2.5132
30		CP	0.9550	0.9440	0.9540	0.9570
	ω_3	RMSE	0.3126	0.9069	0.0796	0.0849
		RB(%)	3.9656	9.7440	0.2921	1.8125
		CP	0.9570	0.9500	0.9510	0.9590
	ω_4	RMSE	0.3126	0.9069	0.0797	0.0840
		RB(%)	3.9656	9.7440	0.2921	1.8125
		CP	0.9570	0.9500	0.9640	0.9600
	MLE	RMSE	0.2500	0.6531	0.0663	0.0739
		RB(%)	3.3462	5.7905	0.0960	0.2747
		CP	0.9460	0.9580	0.9280	0.9040
	ω_1	RMSE	0.2476	0.6539	0.0644	0.0719
		RB(%)	3.0103	5.5583	0.2874	0.9139
		CP	0.9520	0.9460	0.9540	0.9470
50	ω_2	RMSE	0.2476	0.6539	0.0652	0.0721
		RB(%)	3.0103	5.5583	1.2762	1.3791
		CP	0.9520	0.9460	0.9450	0.9380
	ω_3	RMSE	0.2444	0.6322	0.0643	0.0725
		RB(%)	2.6582	5.5583	0.2874	0.9139
		CP	0.9480	0.9460	0.9440	0.9390
	ω_4	RMSE	0.2444	0.6322	0.0644	0.0719
		RB(%)	2.6582	5.5583	0.2874	0.9139
		CP	0.9480	0.9460	0.9540	0.9470

(a) Sample $\theta_1^{(s)}$ at random from $\Gamma(n, \sum_{i=1}^{n} T_i^{\alpha})$.

(b) With probability $\left(k\left(\theta_1^{(s-1)}\right)\right)^{-1/2}$, accept $\theta_1^{(s)}$ as a draw from $p(\theta_1|\alpha, \text{Data})$, where $\theta_1^{(s-1)}$ is the value of θ_1 at the $(s-1)$th iteration in the sampling step. If the drawn $\theta_1^{(s)}$ is rejected, return to step (a).

(3) The joint marginal posterior distribution of (θ_2, θ_3) is proportional to

$$\theta_2^{n_1-\frac{1}{2}}\theta_3^{n_2-\frac{1}{2}}(1-\theta_2-\theta_3)^{n_0-\frac{1}{2}}$$
$$(1-\theta_2)^{-c_1/2}(1-\theta_3)^{-c_2/2}.$$

- When the reference priors $\omega_1(\alpha, \theta_1, \theta_2, \theta_3)$ and $\omega_4(\alpha, \theta_1, \theta_2, \theta_3)$ are used, $c_1 = c_2 = 0$. Thus, $(\theta_2, \theta_3)|\text{Data} \sim \text{Dir}(n_1 + 1/2, n_2 + 1/2, n_0 + 1/2)$, where $\text{Dir}(b_1, b_2, b_3)$ denotes the Dirichlet distribution with parameters b_1, b_2 and b_3.
- When the reference prior $\omega_2(\alpha, \theta_1, \theta_2, \theta_3)$ is utilised, $c_1 = 1$ and $c_2 = 0$. We use Metropolis–Hastings algorithm to generate the posterior samples of (θ_2, θ_3), and choose $\text{Dir}(n_1 + 1/2, n_2 + 1/2, n_0 + 1/2)$ as the proposal distribution. Thus, the procedure is as follows:

(a) Generate $(\theta_2^{(s)}, \theta_3^{(s)})$ at random from $\text{Dir}(n_1 + 1/2, n_2 + 1/2, n_0 + 1/2)$.

(b) Accept $(\theta_2^{(s)}, \theta_3^{(s)})$ with probability

$$\min\left\{1, \sqrt{(1-\theta_2^{(s-1)})/(1-\theta_2^{(s)})}\right\},$$

where $\theta_2^{(s-1)}$ is the value of θ_2 at the $(s-1)$th iteration in the Metropolis step. Otherwise, let $\theta_2^{(s)} = \theta_2^{(s-1)}$.

(c) Repeat the above two steps until the Markov chain converges.

Table 5. Minimum time to blindness in days and its causes for 71 patients with diabetic retinopathy.

i	T	δ^*	i	T	δ^*	i	T	δ^*	i	T	δ^*	i	T	δ^*
1	266	1	16	125	2	31	717	2	46	663	0	61	503	1
2	91	2	17	777	2	32	642	1	47	567	2	62	423	2
3	154	2	18	306	1	33	141	2	48	966	0	63	285	2
4	285	0	19	415	1	34	407	1	49	203	0	64	315	2
5	583	1	20	307	2	35	356	1	50	84	1	65	727	2
6	547	2	21	637	2	36	1653	0	51	392	1	66	210	2
7	79	1	22	577	2	37	427	2	52	1140	2	67	409	2
8	622	0	23	178	1	38	699	1	53	901	1	68	584	1
9	707	2	24	517	2	39	36	2	54	1247	0	69	355	1
10	469	2	25	272	0	40	667	1	55	448	2	70	1302	1
11	93	1	26	1137	0	41	588	2	56	904	2	71	227	2
12	1313	2	27	1484	1	42	471	0	57	276	1			
13	805	1	28	315	1	43	126	1	58	520	1			
14	344	1	29	287	2	44	350	2	59	485	2			
15	790	2	30	1252	1	45	350	1	60	248	2			

- When the reference prior $\omega_3(\alpha, \theta_1, \theta_2, \theta_3)$ is used, $c_1 = 0$ and $c_2 = 1$. The Metropolis–Hastings algorithm used here is similar to the above case.

After running the above Gibbs sampling procedure M times and discarding the initial B burn-in iterations, then we have $(M - B)$ iterations kept. Since the generated samples are not independent, we need to monitor the auto-correlations of the generated values and select a sampling lag $L > 1$ after which the corresponding auto-correlation is low, that is, the length of the thinning interval is L. Considering the length of the thinning interval, the final number of iterations kept is $M' = (M - B)/L$, and these independent samples will be used for posterior analysis. Then we can use the means of the posterior samples to estimate the parameters, and construct $100(1 - \gamma)\%$ credible intervals of the parameters via the quantiles of posterior samples, where $0 < \gamma < 1$.

6. Simulation study

In this section, we conduct some simulation studies to compare the Bayesian estimators and the MLEs for different parameter values and different sample sizes. We take the shape parameter $\alpha = 0.5, 1, 1.5$, the sample size $n = 15, 30, 50$, and $(\lambda_0, \lambda_1, \lambda_2) = (0.7, 1, 1.5)$. The relative bias (RB), the root-mean-squared error (RMSE) and also the coverage percentages (CPs) based on 95% credible (or confidence) intervals are calculated based on 1000 replications. The RB is defined as follows:

$$\mathrm{RB} = \frac{1}{1000} \sum_{i=1}^{1000} \frac{\hat{\xi}_i - \xi}{\xi} * 100\%,$$

where ξ denotes the true value, and $\hat{\xi}_i$ denotes its estimator of the ith iteration. The results are listed in Tables 2–4, where 'ω_u' denotes that the estimates are obtained under the reference prior $\omega_u(\alpha, \theta_1, \theta_2, \theta_3)$, $u = 1, 2, 3, 4$.

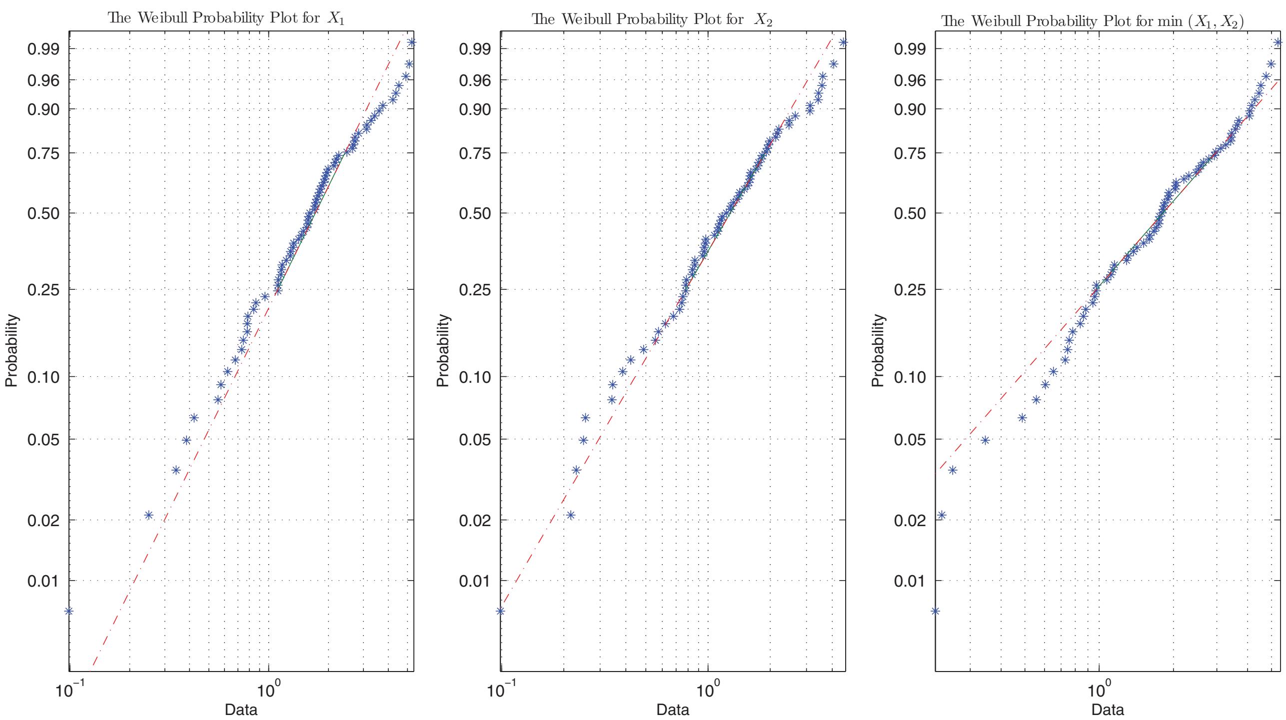

Figure 1. The Weibull probability plot.

Table 6. The estimates of the parameters for the DRS data.

Method	α	95%CI	θ_1	95%CI	θ_2	95%CI	θ_3	95%CP
MLE	1.558	(1.276,1.841)	0.469	(0.321,0.617)	0.394	(0.281,0.508)	0.465	(0.351,0.579)
ω_1	1.557	(1.286,1.846)	0.471	(0.337,0.632)	0.393	(0.285,0.507)	0.462	(0.349,0.577)
ω_2	1.557	(1.286,1.846)	0.471	(0.337,0.632)	0.396	(0.290,0.512)	0.460	(0.346,0.574)
ω_3	1.549	(1.285,1.831)	0.477	(0.341,0.633)	0.391	(0.279,0.506)	0.464	(0.351,0.579)
ω_4	1.549	(1.285,1.831)	0.477	(0.341,0.633)	0.393	(0.285,0.507)	0.462	(0.349,0.577)

Some of the points are quite clear from the simulation study. In all the cases, the estimates are slightly biased, mainly for small sample sizes, but the RBs and RMSEs decrease as the sample size increases. The CPs based on the maximum-likelihood method are a little far from the nominal level, especially when the sample size is small; this is because the asymptotic normality results are used to construct confidence intervals. While all the CPs based on the Bayesian method are close to the nominal level. However, for the case of large sample size, i.e. $n = 50$, the estimates based on both two methods are similar in terms of RBs and RMSEs. Among the four priors, it is hard to tell which one is better. When estimating α and θ_1, the reference priors ω_3 and ω_4 perform better. When estimating the parameters θ_2 and θ_3, ω_1, ω_3 and ω_4 are superior to ω_2, because the RBs of θ_2 and θ_3 based on ω_2 are the largest even the sample size is 50. As we have indicated before, the reference prior can be selected according to our inferential importance of the parameters. For example, when the parameter θ_3 is of interest, the reference prior ω_3 is preferred. As is shown in Tables 2–4, the CPs of θ_3 based on ω_3 are most close to the nominal level 0.95 for all the cases, although the RBs and RMSEs of θ_3 are close to those based on the maximum-likelihood method and other priors.

7. Real data analysis

The data come from the Diabetic Retinopathy Study (DRS) conducted by the National Eye Institute to estimate the effect of laser treatment in delaying the onset of blindness in patients with diabetic retinopathy. There are 71 patients involved in the study. At the beginning of the study, for each patient, one eye was randomly selected for laser treatment by one of three methods (argon laser, xenon arc or a combined treatment), while the other eye was given no treatment. Let X_1 represent the times to blindness of one eye under laser treatment, and X_2 represent the times to blindness of the other eye under no treatment. The observed data of (X_1, X_2) can be found in Csorgo and Welsh (1989). In Table 5, we list the minimum time to blindness ($T = \min\{X_1, X_2\}$) and the index (δ^*) for specifying the causes of failure for each patient, where

$$\delta^* = \begin{cases} 1, & \text{if } X_1 < X_2, \\ 2, & \text{if } X_1 > X_2, \\ 0, & \text{if } X_1 = X_2. \end{cases}$$

From Table 5, we note that some realisations of δ^* are 0, which indicates that blindness of two eyes happens simultaneously.

Before using the MOBW distribution to analyse the data, we will check the distribution of X_1, X_2 and $\min(X_1, X_2)$, respectively. In the following analysis, the original data are divided by 365 and computed in terms of year. The function *wblplot* in MATLAB software is used to graphically assess whether the data come from a Weibull distribution. If the data are from Weibull distribution, the plotted line will be linear. Other distributions might introduce curvature in the plot. From Figure 1, we can intuitively tell that X_1, X_2 and $\min(X_1, X_2)$ fit Weibull distribution. The Kolmogorov–Smirnov distances between the empirical CDF and the hypothesised CDF for X_1, X_2 and $\min(X_1, X_2)$ are 0.0744, 0.0912 and 0.0563, and the corresponding p values are 0.8124, 0.7671 and 0.9745, respectively. Thus, based on the p-values Weibull distribution cannot be rejected for the marginal and for the minimum also. In fact, the MLEs of the shape and scale parameters of the respective Weibull distribution for X_1, X_2 and $\min(X_1, X_2)$ are (1.6456, 0.2511), (1.6382, 0.2955) and (1.5582, 0.4691), and the confidence intervals of the shape parameter for X_1 and X_2 are [1.3699, 1.9768] and [1.3715, 1.9568], respectively, which means the hypothesis of the same-shape parameter of the Weibull distribution for X_1 and X_2 cannot be rejected.

Naturally, we can choose the MOBW distribution to analyse the data. The estimates and 95% confidence (credible) intervals for the parameters are summarised in Table 6. From Table 6, it can be seen that the estimates of parameters are close to each other. Table 6 also shows that the shape parameter $\alpha > 1$ which means the hazard rate function of the lifetime of the two eyes is increasing. Moreover, the failure probability of the laser-treated eye is much smaller than the untreated one, which indicates that the laser treatment has positive effect on delaying the onset of blindness in patients with diabetic retinopathy.

8. Conclusion

In this paper, we have considered the dependent competing risks model by using an MOBW distribution. The objective Bayesian method is proposed to estimate the parameters. Based on different grouping orders, four reference priors are derived. Then the Bayesian approaches are compared with the maximum-likelihood method in terms of RMSE, RB and CP via a simulation study. The simulation results show that the Bayesian estimates perform better in terms of the CPs and RMSEs, and that we can choose a suitable reference prior to estimate the parameters according to the inferential importance.

Acknowledgments

The authors thank the editor-in-chief Dongchu Sun and the referee for their constructive suggestions that greatly improved the article.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The research is supported by Natural Science Foundation of China [grant number 11671303], [grant number 11201345]; Natural Science Foundation of Zhejiang province [grant number LY15G010006]; China Postdoctoral Science Foundation [grant number 2015M572598].

Notes on contributors

Ancha Xu received the B.E. degree from the College of Mathematics and Statistics, Central China Normal University, Wuhan, China, and the Ph.D. degree from East China Normal University, Shanghai, China, in 2006 and 2011, respectively. He is currently an Associate Professor in the School of Mathematics and Information Science, Wenzhou University, Zhejiang, China. His research interests include Bayesian statistics, degradation modeling, and accelerated life testing.

Shirong Zhou is currently working toward the Master degree in tSchool of Mathematics and Information Science, Wenzhou University, Zhejiang, China. His research interests include failure data analysis, accelerated life testing, and Bayesian inference.

References

AL-Hussaini, E. K., Abdel-Hamid, A. H., & Hashem, A. F. (2015). Bayesian prediction intervals of order statistics based on progressively type-II censored competing risks data from the half-logistic distribution. *Journal of the Egyptian Mathematical Society, 23*, 190–196.

Bemis, B., Bain, L. J., & Higgins, J. J. (1972). Estimation and hypothesis testing for the parameters of a bivariate exponential distribution. *Journal of the American Statistical Association, 67*, 927–929.

Berger, J. O., & Bernardo, J. M. (1992). On the development of reference priors (with discussion). In J. M. Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian statistics* (Vol. *4*, pp. 35–60). Oxford: Clarendon Press.

Berger, J. O., Bernardo, J. M., & Sun, D. (2015). Overall objective priors (with discussion). *Bayesian Analysis, 10*, 189–221.

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference (with discussion). *Journal of the Royal Statistical Society B, 41*, 113–147.

Cramer, E., & Schmiedt, A. B. (2011). Progressively type-II censored competing risks data from Lomax distributions. *Computational Statistics and Data Analysis, 55*, 1285–1303.

Csorgo, S., & Welsh, A. H. (1989). Testing for exponential and Marshall-Olkin distribution. *Journal of Statistical Planning and Inference, 23*, 287–300.

Dijoux, Y. D., & Gaudoin, O. (2009). The alert-delay competing risks model for maintenance analysis. *Journal of Statistical Planning and Inference, 139*, 1587–1603.

Dimitrova, D. S., Haberman, S., & Kaishev, V. K. (2013). Dependent competing risks: Cause elimination and its impact on survival. *Insurance: Mathematics and Economics, 53*, 464–477.

Feizjavadian, S. H., & Hashemi, R. (2015). Analysis of dependent competing risks in the presence of progressive hybrid censoring using Marshall-Olkin bivariate Weibull distribution. *Computational Statistics and Data Analysis, 82*, 19–34.

Ghosh, J. K., & Mukerjee, R. (1991). Characterization of priors under which Bayesian and frequentist bartlett corrections are equivalent in the multiparameter case. *Journal of Multivariate Analysis, 38*, 385–393.

Gilks, W. R., & Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Applied Statistics, 41*, 337–348.

Guan, Q., Tang, Y. C., & Xu, A. C. (2013). Objective Bayesian analysis for bivariate Marshall-Olkin exponential distribution. *Computational Statistics and Data Analysis, 64*, 299–313.

Jeffreys, H. (1961). *Theory of probability*. London: Oxford University Press.

Kozumi, H. (2004). Posterior analysis of latent competing risk models by parallel tempering. *Computational Statistics and Data Analysis, 46*, 441–458.

Kundu, D., & Dey, A. K. (2009). Estimating the parameters of the Marshall-Olkin bivariate Weibull distribution by EM algorithm. *Computational Statistics and Data Analysis, 53*, 956–965.

Kundu, D., & Gupta, A. K. (2013). Bayes estimation for the Marshall-Olkin bivariate Weibull distribution. *Computational Statistics and Data Analysis, 57*, 271–281.

Kundu, D., Kannan, N., & Balakrishnan, N. (2003). Analysis of progressively censored competing risks data. In N. Balakrishnan & C. R. Rao (Eds.), *Handbook of statistics on survival analysis* (Vol. *23*, pp. 331–348). New York: Elsevier Publications.

Laplace, P. (1812). *Analytical theory of probability*. Paris: Courcier Press.

Lindqvist, B. H., & Skogsrud, G. (2009). Modeling of dependent competing risks by first passage times of Wiener processes. *IIE Transactions, 41*(1), 72–80.

Marshall, A. W., & Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association, 62*, 30–44.

Mazucheli, J., & Achcar, J. A. (2011). The Lindley distribution applied to competing risks lifetime data. *Computer Methods and Programs in Biomedicine, 104*, 188–192.

Pareek, B., Kundu, D., & Kumar, S. (2009). On progressively censored competing risks data for Weibull distributions. *Computational Statistics and Data Analysis, 53*, 4083–4094.

Wang, C. P., & Ghosh, M. (2003). Bayesian analysis of bivariate competing risks models with covariates. *Journal of Statistical Planning and Inference, 115*, 441–459.

Xu, A. C., Basu, S, & Tang, Y. C. (2014). A full Bayesian approach for masked data in step-stress accelerated life testing. *IEEE Transactions on Reliability, 63*, 798–806.

Xu, A. C, & Tang, Y. C (2010). Reference analysis for Birnbaum-Saunders distribution. *Computational Statistics & Data Analysis, 54*, 185–192.

Xu, A. C., & Tang, Y. C. (2011). Objective Bayesian analysis of accelerated competing failure models under type-I censoring. *Computational Statistics & Data Analysis, 55*, 2830–2839.

Xu, A. C, Tang, Y. C, & Sun, D. C. (2015). Objective Bayesian analysis for masked data under symmetric assumption. *Statistics and its Interface, 8*, 227–237.

Yang, R., & Berger, J. O. (1994). Estimation of a covariance matrix using the reference prior. *The Annals of Statistics, 22*, 1195–2111.

Appendices

Appendix 1

A1. The reference prior of the MOBW distribution

Following the notions of Berger and Bernardo (1992), according to the inferential importance, we order a multi-dimensional parameter $\tau = (\tau_1, \dots, \tau_l)$ and separate it into m groups of sizes $n_1, n_2, \dots, n_m$, and these groups are given by

$$\begin{aligned}\tau_{(1)} &= (\tau_1, \dots, \tau_{n_1}), \quad \tau_{(2)} \\ &= (\tau_{n_1+1}, \dots, \tau_{n_1+n_2}), \dots, \\ \tau_{(i)} &= (\tau_{N_{i-1}+1}, \dots, \tau_{N_i}), \dots, \tau_{(m)} \\ &= (\tau_{N_{m-1}+1}, \dots, \tau_l),\end{aligned}$$

where $N_j = \sum_{i=1}^{j} n_i$. Define

$$\begin{aligned}\tau_{[j]} &= (\tau_{(1)}, \dots, \tau_{(j)}) = (\tau_1, \dots, \tau_{N_j}), \; \tau_{[\sim j]} \\ &= (\tau_{(j+1)}, \dots, \tau_{(m)}) = (\tau_{N_j+1}, \dots, \tau_l),\end{aligned}$$

then $\tau_{[\sim 0]} = \tau$ and $\tau_{[0]}$ is vacuous.

The same as Berger and Bernardo (1992), the notation $I(\tau)$ is the Fisher information matrix and $S(\tau) = (I(\tau))^{-1}$. We will substitute I and S for $I(\tau)$ and $S(\tau)$ in the following sections.

Write S as

$$S = \begin{pmatrix} A_{11} & A_{21}^t & \cdots & A_{m1}^t \\ A_{21} & A_{22} & \cdots & A_{m2}^t \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mm}^t \end{pmatrix},$$

so that A_{ij} is a $n_i \times n_j$ matrix, and define S_j = upper left $(N_j \times N_j)$ corner of S, with $S_m = S$, and $H_j = S_j^{-1}$. Then the matrices

$$\begin{aligned}h_j &\equiv \text{lower right } (n_j \times n_j) \text{ corner of } H_j, \\ j &= 1, \dots, m,\end{aligned}$$

play an importance role in deriving the reference priors. In particular, $h_1 = H_1 = A_{11}^{-1}$ and, if S is a block diagonal matrix (i.e. $A_{ij} = 0$ for all $i \neq j$), then $h_j = A_{jj}^{-1}$, $j = 1, \dots, m$.

Lemma A.1: *If $|h_j(\tau)|$ depending only on $\tau_{[j]}$ holds, for $j = 1, \dots, m$, then the reference prior*

$$\pi(\tau) = \lim_{\ell \to \infty} \frac{\pi^\ell(\tau)}{\pi^\ell(\tau^*)},$$

where $\pi^\ell(\tau) = \left(\prod_{j=1}^{m} \frac{|h_j|^{1/2}}{\int_{\Theta(\tau_{[j-1]})} |h_j|^{1/2} d\tau_{(j)}}\right) I_\Theta(\tau)$, *$|h_j(\tau)|$ is the determinant of $h_j(\tau)$, τ^* is any fixed point in Θ, and Θ is a compact subset.*

One can refer to Berger and Bernardo (1992) for the proof. By using this lemma, the derivation of the m-group reference prior is greatly simplified.

There are four parameters in the MOBW distribution, in the following, we consider three cases, where the parameters are separated into two, three and four groups. We take $\{\alpha, (\theta_1, \theta_2, \theta_3)\}$, $\{(\theta_2, \theta_3), \alpha, \theta_1)\}$ and $\{\alpha, \theta_1, \theta_2, \theta_3\}$ as an example.

(I) The reference prior of $\{\alpha, (\theta_1, \theta_2, \theta_3)\}$ is $\omega_1(\alpha, \theta_1, \theta_2, \theta_3)$.

Proof: Based on the Fisher information matrix Σ, we can easily derive

$$S = \Sigma^{-1} = \begin{pmatrix} \frac{\alpha^2}{n(r_2 - r_1^2)} & -\frac{\alpha\theta_1(1+r_1-\ln\theta_1)}{n(r_2-r_1^2)} & 0 & 0 \\ -\frac{\alpha\theta_1(1+r_1-\ln\theta_1)}{n(r_2-r_1^2)} & \frac{\theta_1^2 k(\theta_1)}{n(r_2-r_1^2)} & 0 & 0 \\ 0 & 0 & \frac{\theta_2(1-\theta_2)}{n} & -\frac{\theta_2\theta_3}{n} \\ 0 & 0 & -\frac{\theta_2\theta_3}{n} & \frac{\theta_3(1-\theta_3)}{n} \end{pmatrix}.$$

According to S_j = upper left $(N_j \times N_j)$ corner of S, $H_j = S_j^{-1}$, we can obtain

$$\begin{aligned}S_1 &= \frac{\alpha^2}{n(r_2 - r_1^2)}, \quad S_2 = S, \\ H_1 &= \frac{n(r_2 - r_1^2)}{\alpha^2}, \quad H_2 = \Sigma.\end{aligned}$$

Due to h_j = lowerright $(n_j \times n_j)$ cornerof H_j, $j = 1, 2$, then

$$\begin{aligned}h_1 &= H_1 = \frac{n(r_2 - r_1^2)}{\alpha^2}, \quad h_2 \\ &= \begin{pmatrix} \frac{n}{\theta_1^2} & 0 & 0 \\ 0 & \frac{n}{\theta_2} + \frac{n}{1-\theta_2-\theta_3} & \frac{n}{1-\theta_2-\theta_3} \\ 0 & \frac{n}{1-\theta_2-\theta_3} & \frac{n}{\theta_3} + \frac{n}{1-\theta_2-\theta_3} \end{pmatrix}.\end{aligned}$$

Choose $\Theta_k = \Theta_{1k} \times \Theta_{234k} = \{\alpha | a_{1k} < \alpha < b_{1k}\} \times \{(\theta_1, \theta_2, \theta_3) | a_{2k} < \theta_1 < b_{2k}, a_{3k} < \theta_2, a_{4k} < \theta_3, \theta_2 + \theta_3 < d_k\}$, such that $a_{1k}, a_{2k}, a_{3k}, a_{4k} \longrightarrow 0$, $b_{1k}, b_{2k} \longrightarrow \infty$, $d_k \longrightarrow 1$. Note that h_1 and h_2 satisfy Lemma A.1, then after some calculations, the reference prior for $\{\alpha, (\theta_1, \theta_2, \theta_3)\}$ is

$$\begin{aligned}\omega_1(\alpha, \theta_1, \theta_2, \theta_3) &= \lim_{k \to \infty} \frac{\pi^k(\alpha, \theta_1, \theta_2, \theta_3)}{\pi^k(1, 1, 0.2, 0.3)} \\ &= \frac{1}{\sqrt{\alpha^2 \theta_1^2 \theta_2 \theta_3 (1 - \theta_2 - \theta_3)}}.\end{aligned}$$

(II) The reference prior of $\{(\theta_2, \theta_3), \alpha, \theta_1\}$ is $\omega_1(\alpha, \theta_1, \theta_2, \theta_3)$.

□

Proof: The Fisher information matrix of $\{(\theta_2,\theta_3),\alpha,\theta_1\}$ is

$$\Sigma_1=\begin{pmatrix} \frac{n}{\theta_2}+\frac{n}{1-\theta_2-\theta_3} & \frac{n}{1-\theta_2-\theta_3} & 0 & 0\\ \frac{n}{1-\theta_2-\theta_3} & \frac{n}{\theta_3}+\frac{n}{1-\theta_2-\theta_3} & 0 & 0\\ 0 & 0 & \frac{nk(\theta_1)}{\alpha^2} & \frac{n(1+r_1-\ln\theta_1)}{\alpha\theta_1}\\ 0 & 0 & \frac{n(1+r_1-\ln\theta_1)}{\alpha\theta_1} & \frac{n}{\theta_1^2}\end{pmatrix}.$$

Thus

$$S=\Sigma_1^{-1}=\begin{pmatrix} \frac{\theta_2(1-\theta_2)}{n} & \frac{-\theta_2\theta_3}{n} & 0 & 0\\ \frac{-\theta_2\theta_3}{n} & \frac{\theta_3(1-\theta_3)}{n} & 0 & 0\\ 0 & 0 & \frac{\alpha^2}{n(r_2-r_1^2)} & \frac{-\alpha\theta_1(1+r_1-\ln\theta_1)}{n(r_2-r_1^2)}\\ 0 & 0 & \frac{-\alpha\theta_1(1+r_1-\ln\theta_1)}{n(r_2-r_1^2)} & \frac{\theta_1^2k(\theta_1)}{n(r_2-r_1^2)}\end{pmatrix},$$

$$S_1=\begin{pmatrix} \frac{\theta_2(1-\theta_2)}{n} & \frac{-\theta_2\theta_3}{n}\\ \frac{-\theta_2\theta_3}{n} & \frac{\theta_3(1-\theta_3)}{n}\end{pmatrix},\ S_2=\begin{pmatrix} \frac{\theta_2(1-\theta_2)}{n} & \frac{-\theta_2\theta_3}{n} & 0\\ \frac{-\theta_2\theta_3}{n} & \frac{\theta_3(1-\theta_3)}{n} & 0\\ 0 & 0 & \frac{\alpha^2}{n(r_2-r_1^2)}\end{pmatrix},\ S_3=S.$$

According to $H_j=S_j^{-1}$, and $h_j=$ lower right $(n_j\times n_j)$ corner of H_j, $j=1,2,3$, we can obtained that $H_1=\begin{pmatrix} \frac{n}{\theta_2}+\frac{n}{1-\theta_2-\theta_3} & \frac{n}{1-\theta_2-\theta_3}\\ \frac{n}{1-\theta_2-\theta_3} & \frac{n}{\theta_3}+\frac{n}{1-\theta_2-\theta_3}\end{pmatrix}$, $H_2=\begin{pmatrix} \frac{n}{\theta_2}+\frac{n}{1-\theta_2-\theta_3} & \frac{n}{1-\theta_2-\theta_3} & 0\\ \frac{n}{1-\theta_2-\theta_3} & \frac{n}{\theta_3}+\frac{n}{1-\theta_2-\theta_3} & 0\\ 0 & 0 & \frac{n(r_2-r_1^2)}{\alpha^2}\end{pmatrix}$, $H_3=\Sigma_1$. Then $h_1=H_1$, $h_2=\frac{n(r_2-r_1^2)}{\alpha^2}$, $h_3=\frac{n}{\theta_1^2}$.

Choose $\Theta_k=\{((\theta_2,\theta_3),\alpha,\theta_1)|a_{1k}<\alpha<b_{1k}, a_{2k}<\theta_1<b_{2k}, a_{3k}<\theta_2, a_{4k}<\theta_3, \theta_2+\theta_3<d_k\}$, such that $a_{1k},a_{2k},a_{3k},a_{4k}\longrightarrow 0$, $b_{1k},b_{2k}\longrightarrow\infty$, $d_k\longrightarrow 1$. According to Lemma A.1, it is not difficult to obtain that the reference prior for $\{(\theta_2,\theta_3),\alpha,\theta_1\}$ is also

$$\omega_1(\alpha,\theta_1,\theta_2,\theta_3)=\frac{1}{\sqrt{\alpha^2\theta_1^2\theta_2\theta_3(1-\theta_2-\theta_3)}}.$$

(III) The reference prior of $\{\alpha,\theta_1,\theta_2,\theta_3\}$ is $\omega_2(\alpha,\theta_1,\theta_2,\theta_3)$.

□

Proof: Due to $S=\Sigma^{-1}$, we can obtain

$$S_1=\frac{\alpha^2}{n(r_2-r_1^2)},\ S_2=\begin{pmatrix} \frac{\alpha^2}{n(r_2-r_1^2)} & \frac{-\alpha\theta_1(1+r_1-\ln\theta_1)}{n(r_2-r_1^2)}\\ \frac{-\alpha\theta_1(1+r_1-\ln\theta_1)}{n(r_2-r_1^2)} & \frac{\theta_1^2k(\theta_1)}{n(r_2-r_1^2)}\end{pmatrix},$$

$$S_3=\begin{pmatrix} \frac{\alpha^2}{n(r_2-r_1^2)} & \frac{-\alpha\theta_1(1+r_1-\ln\theta_1)}{n(r_2-r_1^2)} & 0\\ \frac{-\alpha\theta_1(1+r_1-\ln\theta_1)}{n(r_2-r_1^2)} & \frac{\theta_1^2k(\theta_1)}{n(r_2-r_1^2)} & 0\\ 0 & 0 & \frac{\theta_2(1-\theta_2)}{n}\end{pmatrix},\ S_4=S.$$

Thus

$$H_1 = \frac{n(r_2-r_1^2)}{\alpha^2}, H_2 = \begin{pmatrix} \frac{nk(\theta_1)}{\alpha_2} & \frac{n(1+r_1-\ln\theta_1)}{\alpha\theta_1} \\ \frac{n(1+r_1-\ln\theta_1)}{\alpha\theta_1} & \frac{n}{\theta_1^2} \end{pmatrix},$$

$$H_3 = \begin{pmatrix} \frac{nk(\theta_1)}{\alpha_2} & \frac{n(1+r_1-\ln\theta_1)}{\alpha\theta_1} & 0 \\ \frac{n(1+r_1-\ln\theta_1)}{\alpha\theta_1} & \frac{n}{\theta_1^2} & 0 \\ 0 & 0 & \frac{n}{\theta_2(1-\theta_2)}, \end{pmatrix},$$

$$H_4 = \Sigma.$$

Then $h_1 = \frac{n(r_2-r_1^2)}{\alpha^2}$, $h_2 = \frac{n}{\theta_1^2}$, $h_3 = \frac{n}{\theta_2(1-\theta_2)}$, $h_4 = \frac{n(1-\theta_2)}{\theta_3(1-\theta_2-\theta_3)}$.

Choose $\Theta_k = \{(\alpha, \theta_1, \theta_2, \theta_3)|a_{1k} < \alpha < b_{1k}, a_{2k} < \theta_1 < b_{2k}, a_{3k} < \theta_2, a_{4k} < \theta_3, \theta_2+\theta_3 < d_k\}$, such that $a_{1k}, a_{2k}, a_{3k}, a_{4k} \longrightarrow 0$, $b_{1k}, b_{2k} \longrightarrow \infty$, $d_k \longrightarrow 1$. According to Lemma A.1, we can obtain the reference prior for $\{\alpha, \theta_1, \theta_2, \theta_3\}$ is

$$\omega_2(\alpha, \theta_1, \theta_2, \theta_3) = \frac{1}{\sqrt{\alpha^2\theta_1^2\theta_2\theta_3(1-\theta_2)(1-\theta_2-\theta_3)}}.$$

The reference prior for $\pi_u(\alpha, \lambda_0, \lambda_1, \lambda_2)$ can be obtained from $\omega_u(\alpha, \theta_1, \theta_2, \theta_3)$, $u = 1, 2, 3, 4$, according to the one-to-one transformation from $(\alpha, \lambda_0, \lambda_1, \lambda_2)$ to $(\alpha, \theta_1, \theta_2, \theta_3)$. □

Appendix 2

Proving Theorem 5.1 needs the following two results (Guan et al., 2013):

$$\int_{0<x+y<1} x^\zeta y^\beta (1-x-y)^\mu dxdy = B(\zeta+1, \beta+\mu+2)B(\beta+1, \mu+1), \quad \text{(B.1)}$$

$$\int_{0<x+y<1} x^\zeta y^\beta (1-x)^\kappa (1-x-y)^\mu dxdy = B(\zeta+1, \beta+\kappa+\mu+2)B(\beta+1, \mu+1), \text{(B.2)}$$

where $B(\cdot, \cdot)$ is a beta function.

(I) The likelihood function under parameters $(\alpha, \theta_1, \theta_2, \theta_3)$ is

$$L_2 = \alpha^n\theta_1^n\theta_2^{n_1}\theta_3^{n_2}(1-\theta_2-\theta_3)^{n_0} \times \exp\left\{-\theta_1\sum_{i=1}^n T_i^\alpha\right\}\prod_{i=1}^n T_i^{\alpha-1},$$

then the joint posterior distribution of $(\alpha, \theta_1, \theta_2, \theta_3)$ under the prior ω_1 can be written as

$$p_1(\alpha, \theta_1, \theta_2, \theta_3|Data) \propto \alpha^{n-1}\theta_1^{n-1}\theta_2^{n_1-\frac{1}{2}}\theta_3^{n_2-\frac{1}{2}} \times (1-\theta_2-\theta_3)^{n_0-\frac{1}{2}} \exp\left\{-\theta_1\sum_{i=1}^n T_i^\alpha\right\}\prod_{i=1}^n T_i^{\alpha-1}. \text{ (B.3)}$$

Denote the right side of (B.3) as R, using (B.1), we have

$$R_1 = \int_{0<\theta_2+\theta_3<1} R d\theta_2 d\theta_3 = B\left(n_1+\frac{1}{2}, n_0+n_2+1\right)B\left(n_0+\frac{1}{2}, n_2+\frac{1}{2}\right) \times \alpha^{n-1}\theta_1^{n-1}\exp\left\{-\theta_1\sum_{i=1}^n T_i^\alpha\right\}\prod_{i=1}^n T_i^{\alpha-1},$$

$$R_2 = \int_0^\infty R_1 d\theta_1 = B\left(n_1+\frac{1}{2}, n_0+n_2+1\right) \times B\left(n_0+\frac{1}{2}, n_2+\frac{1}{2}\right)\Gamma(n)\frac{\alpha^{n-1}\prod_{i=1}^n T_i^{\alpha-1}}{\left(\sum_{i=1}^n T_i^\alpha\right)^n}.$$

Denote $c = B\left(n_1+\frac{1}{2}, n_0+n_2+1\right)B\left(n_0+\frac{1}{2}, n_2+\frac{1}{2}\right)\Gamma(n)\prod_{i=1}^n T_i^{-1}$, then c is a constant. For any $0 < \varepsilon < 1$, we have

$$R_3 = \int_0^\infty R_2 d\alpha = c\int_0^\infty \frac{\alpha^{n-1}\prod_{i=1}^n T_i^\alpha}{\left(\sum_{i=1}^n T_i^\alpha\right)^n}d\alpha = c\int_0^\varepsilon \frac{\alpha^{n-1}\prod_{i=1}^n T_i^\alpha}{\left(\sum_{i=1}^n T_i^\alpha\right)^n}d\alpha + c\int_\varepsilon^\infty \frac{\alpha^{n-1}\prod_{i=1}^n T_i^\alpha}{\left(\sum_{i=1}^n T_i^\alpha\right)^n}d\alpha.$$

Obviously, $c\int_0^\varepsilon \frac{\alpha^{n-1}\prod_{i=1}^n T_i^\alpha}{(\sum_{i=1}^n T_i^\alpha)^n}d\alpha < \infty$.

$$\int_\varepsilon^\infty \frac{\alpha^{n-1}\prod_{i=1}^n T_i^\alpha}{\left(\sum_{i=1}^n T_i^\alpha\right)^n}d\alpha = \int_\varepsilon^\infty \frac{\alpha^{n-1}}{\prod_{i=1}^n \frac{\sum_{i=1}^n T_i^\alpha}{T_i^\alpha}}d\alpha \le \int_\varepsilon^\infty \frac{\alpha^{n-1}}{\prod_{i=1}^n \frac{T_{(n)}^\alpha}{T_i^\alpha}}d\alpha = \int_\varepsilon^\infty \alpha^{n-1}e^{-\alpha\sum_{i=1}^n \ln\left(\frac{T_{(n)}}{T_i}\right)}d\alpha < \infty,$$

where $T_{(n)}$ is the largest order statistic of T_i, $i = 1, 2, \ldots, n$, then $R_3 < \infty$. Thus

$$\int_0^\infty\int_0^\infty\int_0^1\int_0^1 R d\alpha d\theta_1 d\theta_2 d\theta_3 < \infty.$$

Therefore, the posterior distribution of $(\alpha, \theta_1, \theta_2, \theta_3)$ under the prior ω_1 is proper. Similarly, using (B.2), we can prove that the posterior distribution under the prior ω_2 is also proper.

(II) The posterior distribution of $(\alpha, \theta_1, \theta_2, \theta_3)$ under the prior ω_3 iss

$$\omega_3(\alpha, \theta_1, \theta_2, \theta_3|\text{data}) \propto \alpha^{n-1}\theta_1^{n-1}\theta_2^{n_1-\frac{1}{2}}\theta_3^{n_2-\frac{1}{2}}(1-\theta_3)^{-\frac{1}{2}} \times (1-\theta_2-\theta_3)^{n_0-\frac{1}{2}} \times (k(\theta_1))^{-\frac{1}{2}}\exp\{-\theta_1\sum_{i=1}^n T_i^\alpha\}\prod_{i=1}^n T_i^{\alpha-1}. \text{ (B.4)}$$

We denote the right side of (B.4) as P, using (B.2), we have

$$P_1 = \int_{0<\theta_2+\theta_3<1} P \mathrm{d}\theta_2 \mathrm{d}\theta_3$$
$$= B\left(n_2 + \frac{1}{2}, n_0 + n_1 + \frac{1}{2}\right) B\left(n_0 + \frac{1}{2}, n_1 + \frac{1}{2}\right)$$
$$\alpha^{n-1}\theta_1^{n-1}(k(\theta_1))^{-\frac{1}{2}} \exp\{-\theta_1 \sum_{i=1}^{n} T_i^{\alpha}\} \prod_{i=1}^{n} T_i^{\alpha-1},$$

$$P_2 = \int_0^{\infty} P_1 \mathrm{d}\theta_1$$
$$= B\left(n_2 + \frac{1}{2}, n_0 + n_1 + \frac{1}{2}\right) B\left(n_0 + \frac{1}{2}, n_1 + \frac{1}{2}\right)$$
$$\alpha^{n-1} \prod_{i=1}^{n} T_i^{\alpha-1} \int_0^{\infty} \theta_1^{n-1}(k(\theta_1))^{-\frac{1}{2}}$$
$$\times \exp\left\{-\theta_1 \sum_{i=1}^{n} T_i^{\alpha}\right\} \mathrm{d}\theta_1,$$

denote $Q = B\left(n_2 + \frac{1}{2}, n_0 + n_1 + \frac{1}{2}\right) B\left(n_0 + \frac{1}{2}, n_1 + \frac{1}{2}\right) \prod_{i=1}^{n} T_i^{-1}$, then Q is a constant. For any $\theta_1 > 0$, there exists a constant M_0 satisfying $|(k(\theta_1))^{-\frac{1}{2}}| < M_0$. Then

$$P_2 < QM\alpha^{n-1} \prod_{i=1}^{n} T_i^{\alpha} \int_0^{\infty} \theta_1^{n-1} \exp\left\{-\theta_1 \sum_{i=1}^{n} T_i^{\alpha}\right\} \mathrm{d}\theta_1$$
$$= \frac{QM\Gamma(n)\alpha^{n-1} \prod_{i=1}^{n} T_i^{\alpha}}{(T_i^{\alpha})^n},$$

according to (I),

$$P_3 = \int_0^{\infty} P_2 \mathrm{d}\alpha < \infty.$$

Therefore,

$$\int_0^{\infty}\int_0^{\infty}\int_0^{1}\int_0^{1} P \mathrm{d}\alpha \mathrm{d}\theta_1 \mathrm{d}\theta_2 \mathrm{d}\theta_3 < \infty.$$

Thus the posterior distribution of $(\alpha, \theta_1, \theta_2, \theta_3)$ under the prior ω_3 is proper. Similarly, we can prove that the posterior distribution under the prior ω_4 is also proper.