Published in final edited form as:

Comput Stat Data Anal. 2012 June ; 56(6): 2041–2050. doi:10.1016/j.csda.2011.12.010.

Statistical analysis of bivariate failure time data with Marshall–Olkin Weibull models

Yang Li^{a,*}, Jianguo Sun^a, and Shuguang Song^b

^aDepartment of Statistics, University of Missouri, United States

^bSupport & Service Technology, The Boeing Company, United States

Abstract

This paper discusses parametric analysis of bivariate failure time data, which often occur in medical studies among others. For this, as in the case of univariate failure time data, exponential and Weibull models are probably the most commonly used ones. However, it is surprising that there seem no general estimation procedures available for fitting the bivariate Weibull model to bivariate right-censored failure time data except some methods for special situations. We present and investigate two general but simple estimation procedures, one being a graphical approach and the other being a marginal approach, for the problem. An extensive simulation study is conducted to assess the performances of the proposed approaches and shows that they work well for practical situations. An illustrative example is provided.

Keywords

Bivariate failure time data; Parametric estimation; Marginal approach; Weibull models

1. Introduction

The failure time data analysis is commonly performed in almost all medical studies as well as many other fields and for this, many statistical procedures have been developed for various situations and problems (Kalbfleisch and Prentice, 2002; Klein and Moeschberger, 2002). In particular, one simple and appealing type of procedure is a parametric approach, which is efficient and usually preferred if some parametric models are believed to be reasonable. Among them, exponential and Weibull models are probably the most commonly used ones. Let T denote a failure time variable. Then T is said to follow the exponential model if its density function has the form

$$f(t) = \lambda_0 e^{-\lambda_0 t} \quad t \geq 0,$$

where λ_0 is some parameter. In this case, the hazard and survival functions of T are given by $\lambda(t) = \lambda_0$ and $S(t) = \exp(-\lambda_0 t)$, respectively. That is, it has a constant hazard.

*Corresponding author. ylx33@mail.missouri.edu (Y. Li).

The Weibull model assumes that the density function of T has the form

$$f(t) = \lambda \beta t^{\beta-1} e^{-\lambda t^\beta}, \quad t \geq 0$$

and the corresponding hazard and survival functions are given by $\lambda(t) = \lambda \beta t^{\beta-1}$ and $S(t) = \exp(-\lambda t^\beta)$, respectively, where λ and β are some parameters. It is obvious that if $\beta = 1$, the Weibull model reduces to the exponential model. To fit either the exponential or the Weibull model to failure time data, the most commonly used approach, also the standard one, is to employ the maximum likelihood estimation.

Suppose that there exist two related failure time variables T_1 and T_2 . To describe them, various bivariate models or distributions have been proposed in the literature and this is especially the case when T_1 and T_2 represent the times of the two components of a system in a reliability study. The references for this include Balakrishnan and Basu (1995), Block (1977), Diawara and Carpenter (2010), Hanagal and Ahmadi (2009), Iyer et al. (2002), Kundu and Gupta (2009), Nair and Nair (1988), Regoli (2009), Sahu and Dey (2000) and Xie et al. (2011). In particular, Marshall and Olkin (1967) introduced the so-called Marshall–Olkin bivariate exponential (MOBE) model that assumes that the joint survival function of T_1 and T_2 has the form

$$S(t_1, t_2) = \exp\{-\lambda_1 t_1 - \lambda_2 t_2 - \lambda_{12} \max(t_1, t_2)\}, \quad (1)$$

where λ_1 , λ_2 and λ_{12} are some parameters. As with univariate failure time data, the model above can be generalized to

$$S(t_1, t_2) = \exp\{-\lambda_1 t_1^{\beta_1} - \lambda_2 t_2^{\beta_2} - \lambda_{12} \max(t_1^{\beta_1}, t_2^{\beta_2})\} \quad (2)$$

by introducing two more parameters β_1 and β_2 (Lu, 1989). This model with $\beta_1 = \beta_2$ was referred to as the Marshall–Olkin bivariate Weibull (MOBW) model by Kundu and Dey (2009) and others. In the following, we will use MOBW to denote the model defined in (2). It is easy to see that the parameter λ_{12} in models (1) and (2) represents the correlation between T_1 and T_2 and $\lambda_{12} = 0$ means that they are independent. To further see the relationship between T_1 and T_2 under model (1), let X_1 , X_2 and X_{12} be three independent exponential variables with the hazards λ_1 , λ_2 and λ_{12} , respectively. Then it can be shown that $T_1 = \min(X_1, X_{12})$ and $T_2 = \min(X_2, X_{12})$ follow the MOBE model. It is also straightforward to show that under model (1), we have

$$T_1 \sim \text{Exp}(\lambda_1 + \lambda_{12}), \quad T_2 \sim \text{Exp}(\lambda_2 + \lambda_{12}).$$

That is, both T_1 and T_2 follow exponential distributions marginally.

In the following, we will study the question how to fit model (2) to bivariate failure time data. Among others, one of the early references on this was given by Bhattacharyya and

Johnson (1971), who considered the MOBE model and pointed out that the maximum likelihood estimate may not exist or be unique. It is apparent that this applies to the MOBW model too. More recently Kundu and Dey (2009) discussed the MOBW model for the case where $\beta_1 = \beta_2$ and when one observes exact failure time data. They developed an EM algorithm for the maximum likelihood estimate. The algorithm was later generalized to the situation of right-censored failure time data by Nandi and Dewan (2010) with the use of a pseudolikelihood function. But again they assumed that $\beta_1 = \beta_2$. Several recent references on the topic are given by in Li (2008, 2009), Li and Pellerey (2011), Mai and Scherer (2010), Navarro and Balakrishnan (2010) and Sarhan et al. (2011). However, it is surprising that there seem no general procedures available in the literature for fitting model (2) to right-censored failure time data.

In the remainder of this paper, we will present and investigate two general estimation procedures for fitting model (2) to right-censored failure time data. First we will describe in Section 2 a graphical approach and then a marginal approach in Section 3. Both methods are simple and can be easily implemented. Section 4 gives some results obtained from an extensive simulation study conducted to assess the performances of the proposed approaches, and they indicate that the procedures work well for practical situations. An illustrative example is provided in Section 5 and Section 6 contains some discussion and concluding remarks.

2. A Graphical Estimation Procedure (GEP)

Let T_1 and T_2 be defined as in the previous section and suppose that we observe bivariate right-censored failure time data $\{(t_{1i}, t_{2i}, \delta_{1i}, \delta_{2i}); i = 1, \dots, n\}$. Here δ_{ki} denotes the censoring indicator indicating if t_{ki} represents the true failure time (by $\delta_{ki} = 1$) or the right-censoring time (by $\delta_{ki} = 0$), $k = 1, 2$, $i = 1, \dots, n$. Assume that $(T_1, T_2) \sim \text{MOBW}(\beta_1, \beta_2, \lambda_1, \lambda_2, \lambda_{12})$, the model given in (2), and the main goal is to estimate the unknown parameters $\beta_1, \beta_2, \lambda_1, \lambda_2$ and λ_{12} .

To present the graphical estimation procedure, note that under the assumption, it is easy to show that the marginal distribution functions of T_1 and T_2 have the forms, respectively,

$$F_1(t_1) = 1 - \exp\{-(\lambda_1 + \lambda_{12}) t_1^{\beta_1}\} \quad (3)$$

and

$$F_2(t_2) = 1 - \exp\{-(\lambda_2 + \lambda_{12}) t_2^{\beta_2}\}. \quad (4)$$

That is, both T_1 and T_2 follow the univariate Weibull distribution. Let $S_1(t_1)$ and $S_2(t_2)$ denote the marginal survival functions of T_1 and T_2 , respectively. Then we have

$$\begin{aligned} \log(-\log(S_1(t_1))) &= \beta_1 \log(t_1) + \log(\lambda_1 + \lambda_{12}), \\ \log(-\log(S_2(t_2))) &= \beta_2 \log(t_2) + \log(\lambda_2 + \lambda_{12}). \end{aligned}$$

This suggests that if $S_1(t_1)$ and $S_2(t_2)$ are known, one can fit the linear models given above to estimate the parameters involved.

Specifically, let $\hat{S}_k$ denote the Kaplan–Meier estimate of S_k based on the right-censored data $\{(t_{ki}, \delta_{ki}); i = 1, \dots, n\}$ on T_k , $k = 1, 2$. Define

$$\mathbf{b}_k = \begin{pmatrix} \log(\lambda_k + \lambda_{12}) \\ \beta_k \end{pmatrix}, \quad \mathbf{y}_k = \begin{pmatrix} \log(-\log(\hat{S}_k(t_{k1}))) \\ \vdots \\ \log(-\log(\hat{S}_k(t_{kn}))) \end{pmatrix}, \quad X_k = \begin{pmatrix} 1 & \log(t_{k1}) \\ \vdots & \vdots \\ 1 & \log(t_{kn}) \end{pmatrix}.$$

Then it is natural to fit the data $\{(\hat{S}_k(t_{ki}), t_{ki}); i = 1, \dots, n\}$ to the linear model

$$y_k = X_k \mathbf{b}_k$$

and estimate β_k and $\lambda_k + \lambda_{12}$ by the least-square estimates, $k = 1, 2$.

Let $\hat{\beta}_1, \hat{\beta}_2, \hat{\lambda}_1 + \hat{\lambda}_{12}$ and $\hat{\lambda}_2 + \hat{\lambda}_{12}$ denote the estimates defined above. To estimate individual λ_1, λ_2 and λ_{12} , first assume that a consistent estimate, $\hat{S}(t_1, t_2)$, of the joint survival function $S(t_1, t_2)$ is available, which will be discussed below. Note that given $\hat{\beta}_1$ and $\hat{\beta}_2$, we can divide the observed data $\{(t_{1i}, t_{2i})\}$ into the following two parts

$$I_1 = \{(t_{1i}, t_{2i}); t_{1i}^{\hat{\beta}_1} > t_{2i}^{\hat{\beta}_2}\}, \quad I_2 = \{(t_{1i}, t_{2i}); t_{1i}^{\hat{\beta}_1} < t_{2i}^{\hat{\beta}_2}\}.$$

Also under model (2), we have

$$\log\{-\log S(t_{1i}, t_{2i}) - (\lambda_1 + \lambda_{12}) * t_{1i}^{\beta_1}\} = \beta_2 * \log(t_{2i}) + \log(\lambda_2) \quad (5)$$

for the data points in I_1 and

$$\log\{-\log S(t_{1i}, t_{2i}) - (\lambda_2 + \lambda_{12}) * t_{2i}^{\beta_2}\} = \beta_1 * \log(t_{1i}) + \log(\lambda_1) \quad (6)$$

for the data points in I_2 . Then as before, by replacing β_1, β_2 and $\lambda_1 + \lambda_{12}$ with $\hat{\beta}_1, \hat{\beta}_2$ and $\hat{\lambda}_1 + \hat{\lambda}_{12}$, respectively, as well as $S(t_1, t_2)$ with $\hat{S}(t_1, t_2)$ in Eq. (5), it is natural to fit the data I_1 to model (5) and estimate λ_2 by the least-square estimate. Let $\hat{\lambda}_2$ denote the resulting estimate of λ_2 . It follows that one can estimate λ_{12} and λ_1 by $\hat{\lambda}_{12} = \hat{\lambda}_2 + \hat{\lambda}_{12} - \hat{\lambda}_2$ and $\hat{\lambda}_1 = \hat{\lambda}_1 + \hat{\lambda}_{12} - \hat{\lambda}_{12}$, respectively. Furthermore, one can perform the goodness of fit test by

fitting the data I_2 to model (6) with substituting $\beta_1, \beta_2, \lambda_1, \lambda_2, \lambda_{12}$ and $S(t_1, t_2)$ by $\hat{\beta}_1, \hat{\beta}_2, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_{12}$ and $\hat{S}(t_1, t_2)$, respectively.

The estimation procedure proposed above requires an estimate of the joint survival function $S(t_1, t_2)$. For this, if there does not exist any censoring, it is apparent that one can use the empirical joint survival function

$$\hat{S}(t_1, t_2) = \frac{\sum_{i=1}^n I(t_{1i} \geq t_1, t_{2i} \geq t_2)}{n}.$$

In the case of right-censored data, one choice is to use the estimate given in Dabrowska (1988) and Hougaard (2000) and described below. Define

$$R(t_1, t_2) = \sum_i I(T_{1i} \geq t_1, T_{2i} \geq t_2)$$

and

$$K_{11}(t_1, t_2) = \sum_i I(T_{1i} = t_1, \delta_{1i} = 1, T_{2i} = t_2, \delta_{2i} = 1),$$

which can be regarded as the bivariate risk and failure numbers, respectively. Also define

$$\begin{aligned} K_{10}(t_1, t_2) &= \sum I(T_{1i} = t_1, \delta_{1i} = 1, T_{2i} \geq t_2), & K_{01}(t_1, t_2) &= \sum I(T_{2i} = t_2, \delta_{2i} = 1, T_{1i} \geq t_1), \\ L_{11}(t_1, t_2) &= \frac{K_{11}(t_1, t_2)}{R(t_1, t_2)}, & L_{10}(t_1, t_2) &= \frac{K_{10}(t_1, t_2)}{R(t_1, t_2)}, & L_{01}(t_1, t_2) &= \frac{K_{01}(t_1, t_2)}{R(t_1, t_2)}. \end{aligned}$$

Then the Dabrowska bivariate survival estimate is given by

$$\hat{S}(t_1, t_2) = \hat{S}_1(t_1) \hat{S}_2(t_2) \prod_{0 \leq u \leq t_1, 0 \leq v \leq t_2} \{1 - H(u, v)\},$$

where

$$\hat{S}_1(t_1) = \prod_{u \leq t_1} \{1 - L_{10}(u, 0)\}, \quad \hat{S}_2(t_2) = \prod_{u \leq t_2} \{1 - L_{01}(0, u)\},$$

and

$$H(t_1, t_2) = \frac{L_{10}(t_1, t_2)L_{01}(t_1, t_2) - L_{11}(t_1, t_2)}{\{1 - L_{10}(t_1, t_2)\}\{1 - L_{01}(t_1, t_2)\}}.$$

Note that $\hat{S}$ jumps only at the observed time points where either $\delta_1 = 1$ or $\delta_2 = 1$.

Note that all parameter estimates $\hat{\beta}_1, \hat{\beta}_2, \hat{\lambda}_1, \hat{\lambda}_2$ and $\hat{\lambda}_{12}$ are functions of the Kaplan–Meier estimates $\hat{S}_k(t_k)$ and $\hat{S}(t_1, t_2)$. It is easy to show that they are asymptotically unbiased and consistent. Also their distributions can be approximated by the normal distributions. With respect to the variance estimation, for $\hat{\beta}_k$ and $\hat{\lambda}_k + \hat{\lambda}_{12}$, it is natural to use

$$\text{cov}\{\hat{\mathbf{b}}_k\} = (\mathbf{X}'_k \mathbf{X}_k)^{-1} \mathbf{X}'_k \text{cov}(\mathbf{y}_k) \mathbf{X}_k (\mathbf{X}'_k \mathbf{X}_k)^{-1}$$

with $\text{cov}(\mathbf{y}_k)$ approximated by the sandwich estimate $H \text{cov}(\hat{S}_k(t_k)) H'$ and $\text{cov}(\hat{S}_k(t_k))$ by the modified Greenwood's formula given in Kang and Koehler (1997), where

$$H = H(S_k) = \begin{pmatrix} \frac{1}{S_k(t_{k1}) \log(S_k(t_{k1}))} & 0 & \cdots & 0 \\ 0 & \frac{1}{S_k(t_{k2}) \log(S_k(t_{k2}))} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \frac{1}{S_1(t_{kn}) \log(S_k(t_{kn}))} \end{pmatrix}$$

with S_k replaced by $\hat{S}_k$. To estimate the variances of $\hat{\lambda}_1, \hat{\lambda}_2$ and $\hat{\lambda}_{12}$, the same approach does not work. As in general, it is difficult to estimate the covariance matrix of $\hat{S}(t_1, t_2)$ (Dabrowska, 1989). Corresponding to this, we propose to apply the simple bootstrap procedure. The simulation results below show that these variance estimates seem to work well for practical situations.

3. A Marginal Estimation Procedure (MEP)

It is easy to see that the graphical estimation procedure proposed in the previous section is actually a two-step procedure. In this section, we will employ the same idea but make use of the likelihood approach instead of fitting the linear model.

To motivate the marginal estimation procedure, note that instead of considering the observed data on both T_1 and T_2 together, we can first consider the observed right-censored failure time data on T_1 and T_2 separately. Based on (3) and (4), this yields the following two marginal likelihood functions

$$L_1(\beta_1, (\lambda_1 + \lambda_{12})) = \prod_{i=1}^n \{(\lambda_1 + \lambda_{12}) \beta_1 t_{1i}^{\beta_1 - 1}\}^{\delta_{1i}} \exp\{-(\lambda_1 + \lambda_{12}) t_{1i}^{\beta_1}\}$$

and

$$L_2(\beta_2, (\lambda_2 + \lambda_{12})) = \prod_{i=1}^n \{(\lambda_2 + \lambda_{12}) \beta_2 t_{2i}^{\beta_2 - 1}\}^{\delta_{2i}} \exp\{-(\lambda_2 + \lambda_{12}) t_{2i}^{\beta_2}\}$$

with respect to parameters $(\beta_1, \lambda_1 + \lambda_{12})$ and $(\beta_2, \lambda_2 + \lambda_{12})$, respectively. Let $(\tilde{\beta}_1, \widetilde{\lambda_1 + \lambda_{12}})$ and $(\tilde{\beta}_2, \widetilde{\lambda_2 + \lambda_{12}})$ denote the maximum likelihood estimates obtained from the above two marginal likelihood functions. Then they can be determined by solving the following two sets of equations

$$\frac{\partial \log L_1(\beta_1, (\lambda_1 + \lambda_{12}))}{\partial \beta_1} = 0, \quad \frac{\partial \log L_1(\beta_1, (\lambda_1 + \lambda_{12}))}{\partial (\lambda_1 + \lambda_{12})} = 0$$

and

$$\frac{\partial \log L_2(\beta_2, (\lambda_2 + \lambda_{12}))}{\partial \beta_2} = 0, \quad \frac{\partial \log L_2(\beta_2, (\lambda_2 + \lambda_{12}))}{\partial (\lambda_2 + \lambda_{12})} = 0,$$

respectively, by applying, for example, the Newton–Raphson algorithm.

To estimate the individual parameters λ_1 , λ_2 and λ_{12} , note the fact that under model (2) and given β_1 and β_2 , we have

$$\min(T_1^{\beta_1}, T_2^{\beta_2}) \sim \text{Exp}(\lambda_1 + \lambda_2 + \lambda_{12}). \quad (7)$$

Define $\lambda = \lambda_1 + \lambda_2 + \lambda_{12}$, $t_i^* = \min(t_{1i}^{\beta_1}, t_{2i}^{\beta_2})$ and

$$\delta_i = \begin{cases} 1 & \text{if } \delta_{1i} = \delta_{2i} = 1 \text{ or } \delta_{1i} = 0, \delta_{2i} = 1, T_1^{\beta_1} \geq T_2^{\beta_2} \text{ or } \delta_{1i} = 1, \delta_{2i} = 0, T_1^{\beta_1} \leq T_2^{\beta_2} \\ 0 & \text{otherwise.} \end{cases}$$

Then the data $\{(t_i^*, \delta_i); i = 1, \dots, n\}$ and (7) together give the marginal likelihood function

$$L_m(\lambda) = \prod_{i=1}^n (\lambda)^{\delta_i} \exp(-\lambda t_i^*)$$

with respect to the parameter λ . Let $\tilde{\lambda}$ denote the maximum likelihood estimate of λ given by L_m . It can be easily shown that $\tilde{\lambda} = \sum_{i=1}^n \delta_i / \sum_{i=1}^n t_i^*$ with β_1 and β_2 replaced by $\tilde{\beta}_1$ and $\tilde{\beta}_2$, the marginal maximum likelihood estimates from L_1 and L_2 , respectively. Thus it is natural to estimate λ_1 , λ_2 and λ_{12} by

$$\tilde{\lambda}_1 = \tilde{\lambda} - \lambda_2 + \widetilde{\lambda_{12}}, \quad \tilde{\lambda}_2 = \tilde{\lambda} - \lambda_1 + \widetilde{\lambda_{12}},$$

and

$$\tilde{\lambda}_{12} = \lambda_1 + \widetilde{\lambda_{12}} + \lambda_2 + \widetilde{\lambda_{12}} - \tilde{\lambda},$$

respectively.

It is easy to show that all estimates $\tilde{\beta}_1$, $\tilde{\beta}_2$, $\tilde{\lambda}_1$, $\tilde{\lambda}_2$ and $\tilde{\lambda}_{12}$ are asymptotically unbiased and consistent and their distributions can be approximated by normal distributions. To estimate their variances, for $\tilde{\beta}_1$ and $\tilde{\beta}_2$, it is natural and simple to use the inverses of the observed Fisher information matrices resulting from the marginal likelihood functions L_1 and L_2 . They can be expressed as

$$I_1(\beta_1, \lambda_1 + \lambda_{12}) = \begin{bmatrix} \sum_{i=1}^n \frac{\delta_{1i}}{\beta_1^2} + (\lambda_1 + \lambda_{12}) \sum_{i=1}^n t_{1i}^{\beta_1} \{\log(t_{1i})\}^2, & \sum_{i=1}^n t_{1i}^{\beta_1} \log(t_{1i}) \\ \sum_{i=1}^n t_{1i}^{\beta_1} \log(t_{1i}), & \sum_{i=1}^n \frac{\delta_{1i}}{(\lambda_1 + \lambda_{12})^2} \end{bmatrix}$$

and

$$I_2(\beta_2, \lambda_2 + \lambda_{12}) = \begin{bmatrix} \sum_{i=1}^n \frac{\delta_{2i}}{\beta_2^2} + (\lambda_2 + \lambda_{12}) \sum_{i=1}^n t_{2i}^{\beta_2} \{\log(t_{2i})\}^2, & \sum_{i=1}^n t_{2i}^{\beta_2} \log(t_{2i}) \\ \sum_{i=1}^n t_{2i}^{\beta_2} \log(t_{2i}), & \sum_{i=1}^n \frac{\delta_{2i}}{(\lambda_2 + \lambda_{12})^2} \end{bmatrix}$$

with replacing β_1 , β_2 , $\lambda_1 + \lambda_{12}$ and $\lambda_2 + \lambda_{12}$ by $\tilde{\beta}_1$, $\tilde{\beta}_2$, $\tilde{\lambda}_1 + \widetilde{\lambda_{12}}$ and $\tilde{\lambda}_2 + \widetilde{\lambda_{12}}$, respectively. To estimate the variances of λ_1 , λ_2 and λ_{12} , as with the GEP, it is again very difficult and we propose to use the simple bootstrap procedure. As for the GEP, the simulation results below show that the procedure works well for practical situations.

4. A simulation study

We conducted an extensive simulation study to assess the performances of the two estimation procedures proposed in the previous sections. In the study, to generate T_1 and T_2 from the MOBW model, we first generated independent X_1 , X_2 and X_{12} from the exponential distributions with the hazards λ_1 , λ_2 and λ_{12} , respectively. Then we defined $T_1 = T_1^{*1/\beta_1}$ and $T_2 = T_2^{*1/\beta_2}$, where $T_1^* = \min(X_1, X_{12})$ and $T_2^* = \min(X_2, X_{12})$, which gives $(T_1, T_2) \sim \text{MOBW}(\beta_1, \beta_2, \lambda_1, \lambda_2, \lambda_{12})$. The right-censoring times for both T_1 and T_2 were generated from the same bivariate distribution with the same β_1 and β_2 but different λ_1 , λ_2 and λ_{12} ,

which were chosen to give appropriate censoring percentages. In the simulation study, in addition to the two estimation procedures given in the previous sections, we also considered the estimation procedure proposed by Nandi and Dewan (2010) based on the maximum likelihood approach for the situation where $\beta_1 = \beta_2$. It should be noted that to maximize their pseudo-likelihood function, instead of following the procedure in the M-step of the EM algorithm given there, we employed the Newton–Raphson algorithm for simplicity. The results given below are based on 1000 replications.

Table 1 presents the results on the estimation of all five parameters obtained based on the simulated data with $\beta_1 = \beta_2 = 0.5$, $\lambda_1 = \lambda_2 = \lambda_{12} = 1.0$, $n = 100$ and 20% right-censoring for both failure time variables. For each of the three estimation procedures, the results include the averages of the parameter estimates, the averages of the estimated standard errors (ESD), and the sample standard deviations of the parameter estimates (SSD). Here as before, we use GEP to denote the graphical estimation procedure and MEP for the marginal estimation procedure. Table 2 gives the results obtained under the same setup as that for Table 1 except $n = 200$. In addition to the situation where $\beta_1 = \beta_2$, we also investigated the case where $\beta_1 \neq \beta_2$ and the obtained results are presented in Tables 3 and 4. Note that in Tables 3 and 4, we only considered the GEP and MEP given in Sections 2 and 3 since no maximum likelihood procedure is available for the situation. Also for Tables 3 and 4, the simulated data were generated in the same way as with Tables 1 and 2, respectively, except $\beta_1 = 0.5$ and $\beta_2 = 0.2$. The results given above suggest that both of the proposed estimation procedures seem to work reasonably well, especially on the estimation of β_1 and β_2 . As expected, for the case of $\beta_1 = \beta_2$, the maximum likelihood approach gave more accurate estimates than the other two approaches, and the variance estimates became smaller when the sample size increased for all three approaches.

To assess the approximation of the normal distributions to the distributions of the parameter estimates, we studied the quantile plots of the standardized estimates against the standard normal variable. Fig. 1 displays the quantile plots of the standardized $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\lambda}_1$, $\hat{\lambda}_2$ and $\hat{\lambda}_{12}$, the estimates given by the GEP, based on the simulated data used in Table 4. That is, we have $\beta_1 = 0.5$, $\beta_2 = 0.2$, $\lambda_1 = \lambda_2 = \lambda_{12} = 1.0$ and $n = 200$ with 20% right-censoring. Fig. 2 gives the quantile plots of the standardized $\tilde{\beta}_1$, $\tilde{\beta}_2$, $\tilde{\lambda}_1$, $\tilde{\lambda}_2$ and $\tilde{\lambda}_{12}$, the estimates given by the MEP, based on the same simulated data. It can be seen that the normal distribution approximation seems to work well except for heavier tails on $\tilde{\lambda}_1$, $\tilde{\lambda}_2$ and $\tilde{\lambda}_{12}$. We also tried several other settings such as those with one or both of β_1 and β_2 being larger than 1 and with smaller or larger sample sizes. The performances of the both proposed procedures were similar to those given above. However, the simulation results did suggest that the MEP actually gave better results than the GEP for smaller sample sizes such as $n = 50$.

5. An application

In this section we apply the estimation procedures proposed in the previous sections to a set of bivariate right-censored failure time data arising from the Diabetic Retinopathy Study (DRS) (Huster et al., 1989). The study was conducted by the National Eye Institute to assess the effect of laser photocoagulation in delaying the onset of severe visual loss such as blindness in the patients with diabetic retinopathy. It consisted of 197 high risk patients. At

the beginning of the study, for each patient, one eye was randomly selected for laser photocoagulation and the other was given no treatment, serving as the control. The times to blindness in both eyes were recorded in months and the censoring was caused by death, dropout, or the end of the study.

Let T_1 denote the time to the blindness of the untreated or control eye and T_2 the time to the blindness of the treated eye. Assume that T_1 and T_2 follow the MOBW model $\text{MOBW}(\beta_1, \beta_2, \lambda_1, \lambda_2, \lambda_{12})$. For the analysis, we divided all observed values in the original data by 100 so that the estimates of the shape and scale parameters are of the same order. Table 5 presents the estimation results given by the three procedures investigated in the simulation study and includes the parameter estimates, the estimated standard errors (ESD) and the 95% confidence intervals (CI). It can be seen that all three methods gave similar results and indicated that the laser photocoagulation treatment did seem to significantly delay the onset of severe visual loss in the patients with diabetic retinopathy. In other words, the treated eye seems to have had longer health time than the untreated eye.

To check the appropriateness of the three estimation procedures for the data, we plotted the estimated marginal survival functions of T_1 and T_2 based on each of the three procedures in Fig. 3. For comparison, we also obtained the Kaplan–Meier estimates (KMEs) of the two marginal survival functions separately based on the univariate right-censored data on T_1 and T_2 , respectively, and included them in Fig. 3. They suggest that the two estimation procedures proposed in the previous sections seem to fit the data well, especially the marginal estimation procedure. The maximum likelihood procedure given by Nandi and Dewan (2010) seems to be appropriate for the data too.

6. Concluding remarks

This paper discussed the fitting of the Marshall–Olkin bivariate Weibull model to bivariate right-censored failure time data and for the problem, two simple estimation procedures were proposed. As pointed before, although the maximum likelihood approach would be preferred in general, it may not exist for the problem except for some special cases. The two proposed procedures have the advantages that they apply to more general situations than that given in Nandi and Dewan (2010) and can be easily implemented. A drawback is that they may be less efficient than the maximum likelihood approach if the latter is available.

A practical question that one may be interested in is which of the two proposed estimation procedures should be used for a given data set. If one is only interested in the estimation of β_1 and β_2 , the MEP may be preferred although both procedures should be reasonable in general. This is because the MEP is likelihood-based and may perform a little better than the GEP for small sample size situations. On the other hand, for the estimation of λ_1 , λ_2 and λ_{12} , one may prefer the GEP as it seems to give better normal approximations to the distributions of the proposed estimates than the MEP. Based on the simulation results, the approximation seems to be quite reasonable for $n = 200$ or large sample sizes.

There exist several directions for future research. One is that as mentioned before, it would be useful to develop a better and simple variance estimation method for the proposed estimates of λ_1 , λ_2 and λ_{12} instead of using the bootstrap procedure. Another direction is that

as for any parametric model, it would be helpful to develop some goodness-of-fit or model-checking procedures. Although one could employ the marginal model-checking procedure used in Section 5, one needs to be careful as it only makes use of the marginal information. The focus of this paper has been on right-censored failure time data. In many situations, interval-censored failure time data may arise (Sun, 2006) and thus it is useful to generalize the proposed estimation procedures to such data.

Acknowledgments

The authors wish to thank the Associate Editor and two reviewers for their valuable comments and suggestions that greatly improved the paper.

References

- Balakrishnan, N.; Basu, AP. Exponential Distribution: Theory, Methods and Applications. CRC Press; The Netherlands: 1995.
- Bhattacharyya, GK.; Johnson, RA. Maximum likelihood estimation and hypothesis testing in the bivariate exponential model of Marshall and Olkin. 1971. <http://handle.dtic.mil/100.2/AD737530>
- Block HW. A characterization of a bivariate exponential distribution. The Annals of Statistics. 1977; 5 (4):808–812.
- Dabrowska DM. Kaplan–Meier estimate on the plane. Annals of Statistics. 1988; 16 (4):1475–1489.
- Dabrowska DM. Kaplan–Meier estimate on the plane. A Journal of Multivariate Analysis. 1989; 29 (2):308–325.
- Diawara N, Carpenter M. Mixture of bivariate exponential distributions. Communications in Statistics —Theory and Methods. 2010; 39 (15):2711–2720.
- Hanagal DD, Ahmadi KA. Bayesian estimation of the parameters of bivariate exponential distributions. Communications in Statistics—Simulation and Computation. 2009; 38 (7):1391–1413.
- Hougaard, P. Analysis of Multivariate Survival Data. Springer; New York: 2000.
- Huster WJ, Brookmeyer R, Self SG. Modeling paired survival data with covariates. Biometrics. 1989; 45:145–156. [PubMed: 2655727]
- Iyer SK, Manjunath D, Manivasakan R. Bivariate exponential distributions using linear structures. The Indian Journal of Statistics, Series A. 2002; 64 (1):156–166.
- Kalbfleisch, JD.; Prentice, RL. The Statistical Analysis of Failure Time Data. 2. John Wiley & Sons, Inc; New Jersey: 2002.
- Kang SS, Koehler KJ. Modification of the Greenwood formula for correlated response times. Biometrics. 1997; 53 (3):885–899. [PubMed: 9290220]
- Klein, JP.; Moeschberger, ML. Survival Analysis. 2. Springer; New York: 2002.
- Kundu D, Dey AK. Estimating the parameters of the Marshall–Olkin bivariate Weibull distribution by EM algorithm. Computational Statistics and Data Analysis. 2009; 53 (4):956–965.
- Kundu D, Gupta RD. Bivariate generalized exponential distribution. Journal of Multivariate Analysis. 2009; 100:581–593.
- Li H. Tail dependence comparison of survival Marshall–Olkin copulas. Methodology and Computing in Applied Probability. 2008; 10 (1):39–54.
- Li H. Orthant tail dependence of multivariate extreme value distributions. Journal of Multivariate Analysis. 2009; 100:243–256.
- Li X, Pellerey F. Generalized Marshall–Olkin distributions and related bivariate aging properties. Journal of Multivariate Analysis. 2011; 102:1399–1409.
- Lu JC. Weibull extensions of the Freund and Marshall–Olkin bivariate exponential. IEEE Transactions on Reliability. 1989; 38 (5):615–619.
- Mai JF, Scherer M. The Pickands representation of survival Marshall–Olkin copulas. Statistics and Probability Letters. 2010; 80:357–360.

- Marshall AW, Olkin I. A multivariate exponential distribution. *Journal of American Statistical Association*. 1967; 62:30–44.
- Nair NU, Nair VKR. A characterization of the bivariate exponential distribution. *Biometrical Journal*. 1988; 30:107–112.
- Nandi S, Dewan I. An EM algorithm for estimating the parameters of bivariate Weibull distribution under random censoring. *Computational Statistics and Data Analysis*. 2010; 54 (6):1559–1569.
- Navarro J, Balakrishnan N. Study of some measures of dependence between order statistics and systems. *Journal of Multivariate Analysis*. 2010; 101:52–67.
- Regoli G. A class of bivariate exponential distributions. *Journal of Multivariate Analysis*. 2009; 100:1261–1269.
- Sahu SK, Dey DK. A comparison of frailty and other models for bivariate survival data. *Lifetime Data Analysis*. 2000; 6:207–228. [PubMed: 10949859]
- Sarhan AM, Hamilton DC, Smith B, Kundu D. The bivariate generalized linear failure rate distribution and its multivariate extension. *Computational Statistics and Data Analysis*. 2011; 55:644–654.
- Sun, J. *The Statistical Analysis of Interval-censored Failure Time Data*. Springer; New York: 2006.
- Xie YJ, Xie M, Goh TN. A MEWMA chart for a bivariate exponential distribution. *Journal of Quality Technology*. 2011; 43 (1):50–65.

Fig. 1.
The quantile plots of the standardized parameter estimates by the graphical estimation procedure.

Fig. 2.
The quantile plots of the standardized parameter estimates by the marginal estimation procedure.

Fig. 3.
Estimates of marginal survival functions for the DRS data.

Table 1

Estimation of parameters based on simulated data with $\beta_1 = \beta_2$ and $n = 100$.

True value		β_1	β_2	λ_1	λ_2	λ_{12}
		0.5	0.5	1	1	1
GEP	Estimate	0.4997	0.5024	1.0991	1.1007	0.9253
	ESD	0.0523	0.0521	0.3650	0.3596	0.3209
	SSD	0.0556	0.0547	0.3969	0.3489	0.3108
MEP	Estimate	0.5067	0.5065	1.0778	1.0694	0.9628
	ESD	0.0448	0.0448	0.3592	0.3486	0.3063
	SSD	0.0454	0.0467	0.3525	0.3431	0.2978
MLE	Estimate	0.5002	0.5002	0.9904	0.9875	1.0020
	ESD	0.0282	0.0282	0.1410	0.1422	0.1346
	SSD	0.0335	0.0335	0.1955	0.1896	0.1624

Table 2

Estimation of parameters based on simulated data with $\beta_1 = \beta_2$ and $n = 200$.

True value		β_1	β_2	λ_1	λ_2	λ_{12}
		0.5	0.5	1	1	1
GEP	Estimate	0.4987	0.4979	0.9466	0.9326	1.0617
	ESD	0.0382	0.0382	0.2653	0.2305	0.2211
	SSD	0.0395	0.0416	0.2683	0.2349	0.2280
MEP	Estimate	0.5016	0.5021	1.0437	1.0431	0.9753
	ESD	0.0312	0.0313	0.2850	0.2769	0.2541
	SSD	0.0305	0.0313	0.2825	0.2641	0.2484
MLE	Estimate	0.4955	0.4955	0.9603	0.9634	1.0143
	ESD	0.0252	0.0252	0.1004	0.1003	0.0971
	SSD	0.0234	0.0234	0.1397	0.1355	0.1151

Table 3

Estimation of parameters based on simulated data with β_1 β_2 and $n = 100$.

True value		β_1	β_2	λ_1	λ_2	λ_{12}
		0.5	0.2	1	1	1
GEP	Estimate	0.4979	0.1998	1.0673	1.0866	0.9351
	ESD	0.0544	0.0218	0.4076	0.3401	0.3085
	SSD	0.0574	0.0228	0.3814	0.3337	0.3061
MEP	Estimate	0.5056	0.2037	1.0704	1.0793	0.9683
	ESD	0.0446	0.0181	0.3543	0.3504	0.3017
	SSD	0.0446	0.0179	0.3511	0.3300	0.2963

Table 4

Estimation of parameters based on simulated data with β_1 β_2 and $n = 200$.

True value		β_1	β_2	λ_1	λ_2	λ_{12}
		0.5	0.2	1	1	1
GEP	Estimate	0.4991	0.1998	0.9416	0.9522	1.0570
	ESD	0.0392	0.0157	0.2791	0.2385	0.2254
	SSD	0.0396	0.0161	0.2658	0.2439	0.2367
MEP	Estimate	0.5037	0.2017	1.0409	1.0506	0.9825
	ESD	0.0310	0.0124	0.2843	0.2760	0.2536
	SSD	0.0319	0.0125	0.2904	0.2694	0.2535

Table 5

Estimation of parameters for the DRS data.

		β_1	β_2	λ_1	λ_2	λ_{12}
GEP	Estimate	0.9342	0.9960	1.3713	0.5920	0.3256
	ESD	0.0838	0.1195	0.4284	0.3047	0.2141
	CI	(0.7700, 1.0983)	(0.7618, 1.2301)	(0.5317, 2.2109)	(0, 1.1891)	(0, 0.7452)
MEP	Estimate	0.8231	0.7853	1.1239	0.3650	0.2295
	ESD	0.0730	0.0988	0.4701	0.4187	0.4265
	CI	(0.6800, 0.9662)	(0.5918, 0.9789)	(0.2024, 2.0453)	(0, 1.1856)	(0, 1.0654)
MLE	Estimate	0.8070	0.8070	1.2411	0.4976	0.1093
	ESD	0.0598	0.0598	0.1458	0.0785	0.0340
	CI	(0.6898, 0.9243)	(0.6898, 0.9243)	(0.9554, 1.5269)	(0.3437, 0.6515)	(0.0427, 0.1760)