

UM MODELO DE REGRESSÃO WEIBULL BIVARIADO PARA DADOS DE SOBREVIVÊNCIA

Gladys Dorotea Cacsire BARRIGA¹
Francisco LOUZADA NETO¹
Vicente Garibay CANCHO²

- RESUMO: Neste artigo, propomos um novo modelo Weibull bivariado (BVW), que permite modelar dados de sobrevivência dependentes. O modelo proposto é motivado por uma aplicação de um resultado de Morgenstern (1956). Uma extensão deste novo modelo é também considerada adicionando covariáveis ao modelo BVW. Uma metodologia Bayesiana baseada em métodos de Monte Carlo via cadeias de Markov é desenvolvida para os modelos propostos. Ilustramos a metodologia com três conjuntos de dados reais.
- PALAVRAS-CHAVE: Distribuição bivariada; Dados de sobrevivência bivariado; inferência bayesiana; distribuição Weibull bivariado

1 Introdução

Na análise de sobrevivência e de confiabilidade existem situações em que observamos dois tempos de vida para um mesmo paciente ou equipamento (dados de sobrevivência bivariados). Por exemplo, na área médica podemos ter interesse em estudar os tempos de vida de órgãos humanos emparelhados, por exemplo rins, olhos e recorrência dupla de uma determinada doença. Em aplicações industriais este tipo de dados pode ser observado em sistemas cujo tempo de duração depende da durabilidade de dois componentes. Por exemplo, a avaria de geradores duais em uma planta de potência ou o tempo de vida de motores em um avião bimotor são ilustrações de dados de sobrevivência bivariada. Dados de sobrevivência bivariada em geral são correlacionadas e o estudo dessa dependência tem sido foco de muitas

¹Departamento de Estatística, Universidade Federal de São Carlos – UFSCar, CEP:13565-905, São Carlos, SP, Brasil. E-mail: dfln@power.ufscar.com.br

²Departamento de Matemática Aplicada e Estatística-ICMC, Universidade de São Paulo - USP, CEP:13560-970, São Carlos, SP, Brasil. E-mail: garibay@icmc.usp.br

pesquisas. A abordagem mais popular são os modelos de fragilidade em que um ou mais efeitos aleatórios são incluídos na função de taxa de falha, para levar em conta a dependência entre as observações (ver por exemplo, Clayton, 1978; Vaupel, *et al.* 1979; Hougaard, 1989; e Oakes, 1989). Neste caso, os tempos marginais são condicionalmente independentes dada a variável de fragilidade.

Uma abordagem paramétrica alternativa para a modelagem anterior é considerar modelos que diretamente modelem a dependência nos dados. A família de distribuição exponencial e Weibull bivariados são modelos muito importantes que permitem modelar diretamente a dependência dos dados bivariados, veja, por exemplo, Gumbel (1960) Freud (1961), Marshall & Olkin, (1967) e Block & Basu, (1974), Lee (1979), Achcar & Santander (1993) e Sahu & Dey (2000). Neste artigo propomos um novo modelo Weibull bivariado baseado na distribuição bivariada de Morgenstern (1956).

Na Seção 2, apresentamos o modelo Weibull bivariado (BVW) e estudamos algumas propriedades desse modelo. Na Seção 3, estendemos o modelo BVW ao modelo de regressão BVW. Nas Seções 4 apresentamos a metodologia Bayesiana baseada em métodos de Monte Carlo via Cadeia de Markov (MCMC) para o modelo BVW e o modelo de regressão BVW. Na seção 5, apresentamos três aplicações dos modelos propostos considerando dados reais. Na Seção 6 finalizamos o trabalho com algumas observações gerais.

2 Modelo Weibull bivariado

Considere T_1 e T_2 duas variáveis aleatórias não negativas que denotam os tempos de sobrevivência bivariado de um sistema com dois componentes. O modelo Weibull bivariado (BVW), obtido a partir da distribuição de Morgenstern (1956) tem função de densidade conjunta para (T_1, T_2) dado por,

$$f(t_1, t_2) = \alpha_1 \alpha_2 \lambda_1^{\alpha_1} \lambda_2^{\alpha_2} t_1^{\alpha_1 - 1} t_2^{\alpha_2 - 1} \exp \{ -(\lambda_1 t_1)^{\alpha_1} - (\lambda_2 t_2)^{\alpha_2} \} [1 + \tau h_1(t_1) h_2(t_2)] \quad (1)$$

onde $t_1 > 0$, $t_2 > 0$, $\alpha_1, \alpha_2, \lambda_1, \lambda_2 > 0$, $-1 \leq \tau \leq 1$ é o parâmetro de dependência no modelo e $h_j(t_j)$ é definida por,

$$h_j(t_j) = 2 \exp \{ -(\lambda_j t_j)^{\alpha_j} \} - 1, \quad j = 1, 2. \quad (2)$$

Note que quando $\tau = 0$ em (1) temos que T_1 e T_2 são independentes. Além disso quando $\alpha_1 = \alpha_2 = 1$ é a função de densidade de probabilidade conjunta da distribuição exponencial bivariado (BVE) tipo II de Gumbel (1960).

As densidades marginais da distribuição BVW são dadas por

$$f(t_j) = \alpha_j \lambda_j^{\alpha_j} t_j^{\alpha_j - 1} \exp \{ -(\lambda_j t_j)^{\alpha_j} \}, \quad (3)$$

a qual é uma distribuição Weibull com parâmetro de forma α_j e parâmetro de escala λ_j , de modo que, a média e variância de T_j ; $j = 1, 2$, são dados por

$$E(T_j) = \lambda_j^{-1} \Gamma\left(1 + \frac{1}{\alpha_j}\right) \text{ e } Var(T_j) = \lambda_j^{-2} \left(\Gamma\left(1 + \frac{2}{\alpha_j}\right) - \Gamma^2\left(1 + \frac{1}{\alpha_j}\right) \right). \quad (4)$$

O valor esperado do produto cruzado T_1T_2 é dado por,

$$E(T_1T_2) = \lambda_1^{-1}\lambda_2^{-1} \left(1 + \tau[1 - 2^{-\frac{1}{\alpha_1}}][1 - 2^{-\frac{1}{\alpha_2}}] \right) \Gamma\left(1 + \frac{1}{\alpha_1}\right)\Gamma\left(1 + \frac{1}{\alpha_2}\right).$$

Então, o coeficiente de correlação entre T_1 e T_2 , é dado por

$$\rho(T_1, T_2) = \frac{\tau \left(1 - 2^{-\frac{1}{\alpha_1}} \right) \left(1 - 2^{-\frac{1}{\alpha_2}} \right)}{\sqrt{\left(\frac{\Gamma(\frac{2}{\alpha_1}+1)}{\Gamma(\frac{1}{\alpha_1}+1)^2} - 1 \right) \left(\frac{\Gamma(\frac{2}{\alpha_2}+1)}{\Gamma(\frac{1}{\alpha_2}+1)^2} - 1 \right)}}, \quad (5)$$

observe quando $\tau = 0$ o coeficiente de correlação, $\rho = 0$.

Assumindo um sistema composto por dois componentes, ligados em série, tem-se que o sistema funciona se os dois componentes funcionam simultaneamente; assumindo-se ainda que eles têm distribuição de sobrevivência com função de densidade de probabilidade dada por (1), a função de confiabilidade de um sistema no tempo t_0 é dada por

$$R_s(t_0) = \exp \{ -(\lambda_1 t_0)^{\alpha_1} - (\lambda_2 t_0)^{\alpha_2} \} [1 + \tau(e^{-(\lambda_1 t_0)^{\alpha_1}} - 1)(e^{-(\lambda_2 t_0)^{\alpha_2}} - 1)]. \quad (6)$$

Considerando agora um sistema onde os componentes estão ligados em paralelo e possuem distribuição de sobrevivência com função de densidade de probabilidade dada por (1), a função de confiabilidade do sistema no tempo t_0 é dada por

$$R_p(t_0) = e^{-(\lambda_1 t_0)^{\alpha_1}} + (1 - e^{-(\lambda_1 t_0)^{\alpha_1}})e^{-(\lambda_2 t_0)^{\alpha_2}} \left[1 - \tau(1 - e^{-(\lambda_2 t_0)^{\alpha_2}})e^{-(\lambda_1 t_0)^{\alpha_1}} \right]. \quad (7)$$

3 Modelo de regressão Weibull bivariado

Na maioria de situações práticas, o tempo de sobrevivência é influenciada por covariáveis. Por exemplo, a sobrevivência de portadores de alguma doença pode depender da idade, pressão sanguínea, tipo de tratamento utilizado, dentre outros fatores. Considerando a transformação $Y_1 = \log(T_1)$, e $Y_2 = \log(T_2)$ e a reparametrização $\alpha_j = 1/\sigma_j$ e $\lambda_j = \exp(-\mu_j)$ $j = 1, 2$ em (1), temos a distribuição valor extremo bivariado (BVEV) com a seguinte função de densidade

$$f(y_1, y_2) = \frac{1}{\sigma_1\sigma_2} \exp \left\{ \frac{y_1 - \mu_1}{\sigma_1} + \frac{y_2 - \mu_2}{\sigma_2} - \exp \left\{ \frac{y_1 - \mu_1}{\sigma_1} \right\} - \exp \left\{ \frac{y_2 - \mu_2}{\sigma_2} \right\} \right\} \left[1 + \tau \left(2 \exp \left\{ - \exp \left\{ \frac{y_1 - \mu_1}{\sigma_1} \right\} \right\} - 1 \right) \left(2 \exp \left\{ - \exp \left\{ \frac{y_2 - \mu_2}{\sigma_2} \right\} \right\} - 1 \right) \right], \quad (8)$$

onde $(y_1, y_2) \in R^2$, $\mu_1, \mu_2 \in R$, $\sigma_1, \sigma_2 > 0$ e $-1 \leq \tau \leq 1$. Note que no caso $\tau = 0$ temos que Y_1 e Y_2 são variáveis aleatórias independentes com distribuição valor

extremo com parâmetros de locação μ_j e parâmetros de escala σ_j , $j = 1, 2$ (ver por exemplo, Lawless, 2003).

De (8), pode-se representar o modelo log-linear dado por

$$Y_j = \mu_j + \sigma_j U_j, \quad j = 1, 2. \quad (9)$$

onde (U_1, U_2) tem densidade dada por

$$f(u_1, u_2) = \exp(u_1 - e^{u_1} + u_2 - e^{u_2}) [1 + \tau(2 \exp\{-e^{u_1}\} - 1)(2 \exp\{-e^{u_2}\} - 1)], \quad (10)$$

que denominaremos distribuição valor extremo bivariado padrão (BVEP). Note que no caso $\tau = 0$, temos U_1 e U_2 são variáveis aleatórias independentes com distribuição valor extremo padrão.

Em (9), assumindo uma relação linear entre a variável Y_j e vetor de covariáveis $\mathbf{X}_j = (x_{1j}, \dots, x_{q_jj})^\top$, $j = 1, 2$, temos o modelo de regressão Weibull bivariado dado por,

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1^\top \boldsymbol{\beta}_1 \\ \mathbf{X}_2^\top \boldsymbol{\beta}_2 \end{pmatrix} + \begin{pmatrix} \sigma_1 U_1 \\ \sigma_2 U_2 \end{pmatrix}, \quad (11)$$

onde $\boldsymbol{\beta}_j = (\beta_{1j}, \dots, \beta_{q_jj})$ é o vetor de parâmetros q_j -dimensional associado ao vetor de covariadas \mathbf{X}_j , $j = 1, 2$ e (U_1, U_2) tem distribuição valor extremo bivariado dado por (10). Note que o modelo (11), permite considerar diferentes covariáveis.

Da relação $Y_1 = \log(T_1)$ e $Y_2 = \log(T_2)$, obtemos

$$\begin{pmatrix} T_1 \\ T_2 \end{pmatrix} = \begin{pmatrix} \exp(\mathbf{X}_1^\top \boldsymbol{\beta}_1) V_1^{\sigma_1} \\ \exp(\mathbf{X}_2^\top \boldsymbol{\beta}_2) V_2^{\sigma_2} \end{pmatrix} \quad (12)$$

onde $(V_1 = e^{U_1}, V_2 = e^{U_2})$ é uma variável aleatória com distribuição de Gumbel tipo II. Procedimentos de inferência para modelos de sobrevivência bivariado sob a perspectiva frequentista, podem ser encontradas em, Hougaard (1989), Joe (1997), e Shih e Louis (1995). Sahu & Dey (2000) considerando uma extensão do modelo exponencial bivariado de Marshall & Olkin (1967) desenvolveram um procedimento de inferência bayesiana baseado em métodos de simulação de Monte Carlo via cadeias de Markov (MCMC).

4 Inferência Bayesiana

Nesta seção apresentamos o procedimento de inferência Bayesiana baseado em métodos MCMC, para os modelos propostos nas seções anteriores.

4.1 O modelo Weibull bivariado

Sejam $(t_{11}, t_{21}), \dots, (t_{1n}, t_{2n})$, n pares de observações independentes de (T_1, T_2) cuja função de densidade de probabilidade conjunta é dado por (1), a função de verossimilhança para $\boldsymbol{\theta} = (\alpha_1, \alpha_2, \lambda_1, \lambda_2, \tau)$, dado o conjunto de observações, é dada por

$$L(\boldsymbol{\theta}|D) \propto \alpha_1^n \alpha_2^n \lambda_1^{n\alpha_1} \lambda_2^{n\alpha_2} D_1^{\alpha_1-1} D_2^{\alpha_2-1} \exp(-\lambda_1^{\alpha_1} S_1 - \lambda_2^{\alpha_2} S_2) \prod_{i=1}^n [1 + \tau h_2(t_{i1}) h_2(t_{i2})] \quad (13)$$

sendo D o conjunto de observações, $D_j = \prod_{i=1}^n t_{ij}$, $S_j = \sum_{i=1}^n t_{ij}^{\alpha_j}$ e $h_j(t_{ij})$ dada em (2), $j = 1, 2$, $i = 1, \dots, n$.

Para a inferência Bayesiana é assumida independência a priori entre os parâmetros com as seguintes distribuições a priori: $\alpha_j \sim G(a_j, b_j)$, $j = 1, 2$, $\lambda_j \sim G(p_j, q_j)$ $j = 1, 2$ e $(1 - \tau)/2 \sim Beta(r_1, r_2)$, a qual implica que $\pi(\tau) \propto (1 - \tau)^{r_1-1} (1 + \tau)^{r_2-1}$, onde $G(a_j, b_j)$ denota a distribuição gama com média a_j/b_j e variância a_j/b_j^2 e $Beta(r_1, r_2)$ a distribuição beta. Assim, a densidade a priori conjunta para $\boldsymbol{\theta}$ é dada por

$$\pi(\boldsymbol{\theta}) \propto (1 - \tau)^{r_1-1} (1 + \tau)^{r_2-1} \prod_{j=1}^2 \alpha_j^{a_j-1} \lambda_j^{p_j-1} \exp\{-b_j \alpha_j - q_j \lambda_j\}. \quad (14)$$

Combinando a função de verossimilhança (13), com a distribuição a priori dada em (14), a distribuição a posteriori conjunta para $\boldsymbol{\theta}$ é dada por,

$$\begin{aligned} \pi(\boldsymbol{\theta}|D) \propto & \alpha_1^{n+a_1-1} \alpha_2^{n+a_2-1} \lambda_1^{n\alpha_1+p_1-1} \lambda_2^{n\alpha_2+p_2-1} D_1^{\alpha_1-1} D_2^{\alpha_2-1} \\ & \exp(-b_1 \alpha_1 - b_2 \alpha_2 - \lambda_1^{\alpha_1} S_1 + \lambda_1 q_1 - \lambda_2^{\alpha_2} S_2 + \lambda_2 q_2) \\ & (1 - \tau)^{r_1-1} (1 + \tau)^{r_2-1} \prod_{i=1}^n [1 + \tau h_1(t_{i1}) h_2(t_{i2})]. \end{aligned} \quad (15)$$

Observe que a densidade a posteriori conjunta(15) não é uma densidade padrão, portanto só podemos avaliar as densidades a posteriori marginais por meio de métodos de aproximação, tais como o método de Laplace (Tierney & Kadane, 1986) ou usando métodos de simulação de Monte Carlo via Cadeias de Markov, tais como Gibbs Sampling (Casella & George, 1992) e Metropolis-Hastings (Chib & Greenberg, 1995). Nesta análise consideramos os métodos de simulação por serem de simples implementação computacional. Para a obtenção de amostras da distribuição a posteriori conjunta $\boldsymbol{\theta}$, são utilizados métodos de simulação de Monte Carlo via Cadeias de Markov (MCMC). Para isso são gerados amostras das distribuições condicionais $\pi(\theta_k | \boldsymbol{\theta}_{(k)}, D)$ para $k = 1, \dots, p + 2$, onde $\boldsymbol{\theta}_{(k)} = (\theta_1, \dots, \theta_{k-1}, \theta_{k+1}, \dots, \theta_{p+2})$, utilizando o algoritmo de Metropolis-Hasting quando as distribuições condicionais não são facilmente identificadas como distribuições condicionais. As distribuições condicionais para o algoritmo são dadas

por

$$\begin{aligned}
 \pi(\alpha_1|\alpha_2, \lambda_1, \lambda_2, \delta, D) &\propto \alpha_1^{n+a_1-1} \varphi(\boldsymbol{\alpha}, \boldsymbol{\lambda}, \tau) \exp\{-(b_1 + \log(D_1))\alpha_1 - S_1\lambda^{\alpha_1}\} \\
 \pi(\alpha_2|\alpha_1, \lambda_1, \lambda_2, \delta, D) &\propto \alpha_2^{n+a_2-1} \varphi(\boldsymbol{\alpha}, \boldsymbol{\lambda}, \tau) \exp\{-(b_2 + \log(D_2))\alpha_2 - S_2\lambda^{\alpha_2}\} \\
 \pi(\lambda_1|\alpha_1, \alpha_2, \lambda_2, \delta, D) &\propto \lambda_1^{n\alpha_1+p_1-1} \varphi(\boldsymbol{\alpha}, \boldsymbol{\lambda}, \tau) \exp\{-S_1\lambda^{\alpha_1} - q_1\lambda_1\} \\
 \pi(\lambda_2|\alpha_1, \alpha_2, \lambda_1, \delta, D) &\propto \lambda_2^{n\alpha_2+p_2-1} \varphi(\boldsymbol{\alpha}, \boldsymbol{\lambda}, \tau) \exp\{-S_2\lambda_2 - q_2\lambda_2\} \\
 \pi(\tau|\alpha_1, \alpha_2, \lambda_1, \lambda_2, D) &\propto (1-\tau)^{r_1-1}(1+\tau)^{r_2-1} \varphi(\boldsymbol{\alpha}, \boldsymbol{\lambda}, \tau),
 \end{aligned} \tag{16}$$

onde $\varphi(\boldsymbol{\alpha}, \boldsymbol{\lambda}, \tau) = \prod_{i=1}^n [1 + \tau h_1(t_{i2})h_2(t_{i2})]$ sendo $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ e $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$. Observe que as densidades dadas em (16) não são conhecidas, sendo necessário a utilização do algoritmo de Metropolis-Hasting para gerar amostras de $\alpha_1, \alpha_2, \lambda_1, \lambda_2$, e τ .

4.2 Modelo de regressão Weibull bivariado

Sejam (T_1, T_2) os tempos de sobrevivência bivariados com função de densidade conjunta dado por (1). Considere que o logaritmo de tempos de sobrevivência $Y_j = \log(T_j)$, $j = 1, 2$, se relaciona com o vetor de covariáveis X_j através da equação dada em (11). Considere também (C_1, C_2) tempos de censura bivariados. Para $i = 1, \dots, n$ suponha que (T_{i1}, T_{i2}) e (C_{i1}, C_{i2}) são independentes. Para cada i as quantidades observadas são representadas pelas variáveis aleatórias $Z_{ij} = \min(Y_{ij}, \log(C_{ij}))$ e $\delta_{ij} = I(Z_{ij} = Y_{ij})$, indicador de censura, $j = 1, 2$. A função de verossimilhança (veja, Lawless, 2003) de $\boldsymbol{\theta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \sigma_1, \sigma_2, \tau)$, dado o conjunto de observações, é dada por

$$\begin{aligned}
 L(\boldsymbol{\theta}|D) &\propto \prod_{i=1}^n [f(y_{i1}, y_{i2})]^{\delta_{i1}\delta_{i2}} [f_1(y_{i1}, y_{i2})]^{\delta_{i1}(1-\delta_{i2})} [f_2(y_{i1}, y_{i2})]^{(1-\delta_{i1})\delta_{i2}} \\
 &\quad [S(y_{i1}, y_{i2})]^{(1-\delta_{i1})(1-\delta_{i2})},
 \end{aligned} \tag{17}$$

onde

$$\begin{aligned}
 f(y_{i1}, y_{i2}) &= \sigma_1^{-1} \sigma_2^{-1} \exp(z_{i1} - e^{z_{i1}} + z_{i2} - e^{z_{i2}}) \\
 &\quad [1 + \tau(2 \exp(-e^{z_{i1}}) - 1)(2 \exp(-e^{z_{i2}}) - 1)], \\
 f_1(y_{i1}, y_{i2}) &= \sigma_1^{-1} \exp(z_{i1} - e^{z_{i1}} - e^{z_{i2}}) [1 + \tau(2 \exp(-e^{z_{i1}}) - 1)(\exp(-e^{z_{i2}}) - 1)] \\
 f_2(y_{i1}, y_{i2}) &= \sigma_2^{-1} \exp(z_{i2} - e^{z_{i1}} - e^{z_{i2}}) [1 + \tau(\exp(-e^{z_{i1}}) - 1)(2 \exp(-e^{z_{i2}}) - 1)] \\
 S(y_{i1}, y_{i2}) &= \exp(-e^{z_{i1}} - e^{z_{i2}}) [1 + \tau(\exp(-e^{z_{i1}}) - 1)(\exp(-e^{z_{i2}}) - 1)]
 \end{aligned}$$

sendo $z_{ij} = \frac{y_{ij} - \mathbf{x}_j^\top \boldsymbol{\beta}_j}{\sigma_j}$, $j = 1, 2$, $i = 1, \dots, n$.

Como na Seção 4.1 para a inferência Bayesiana é assumido independência a priori entre os parâmetros e assumimos a mesma densidade a priori para o parâmetro de dependência, τ e para $1/\sigma_j \sim G(a_j, b_j)$, $j = 1, 2$ e $\boldsymbol{\beta}_j \sim N_{q_j}(\boldsymbol{\mu}_j; \Sigma_j)$ onde $N_k(\boldsymbol{\mu}, \Sigma)$ denota a distribuição normal k -variada. Daí pode-se demonstrar que a

distribuição a priori conjunta de $\theta = (\beta_1, \beta_2, \sigma_1, \sigma_2, \tau)$ é dado por,

$$\pi(\theta) \propto (1 - \tau)^{r_1 - 1} (1 + \tau)^{r_2 - 1} \prod_{j=1}^2 \sigma_j^{-(a_j + 1)} \exp\left\{-\frac{b_j}{\sigma_j}\right\} \phi_{q_j}(\beta_j | \mu_j, \Sigma_j) \quad (18)$$

onde $\phi_{q_j}(\beta_j | \mu_j, \Sigma_j)$ é a função de densidade de probabilidade da distribuição normal q_j -variada com vetor de médias μ_j e matriz de covariância $\Sigma_j = \text{diagonal}(\sigma_{1j}^2, \dots, \sigma_{q_j j}^2)$. Como na Seção 4.1, todos os hiperparâmetros das distribuições a priori consideradas são conhecidos.

Combinando a função de verossimilhança (17), com a distribuição a priori dada em (18), a distribuição a posteriori conjunta para θ é dada por

$$\begin{aligned} \pi(\theta | D) &\propto (1 - \tau)^{r_1 - 1} (1 + \tau)^{r_2 - 1} \prod_{j=1}^2 \sigma_j^{-(2n + a_j + 1)} \exp\left\{-\frac{b_j}{\sigma_j}\right\} \phi_{q_j}(\beta_j | \mu_j, \Sigma_j) \\ &\times \exp\left\{\sum_{i=1}^n [\delta_{i1} \delta_{i2} (z_{i1} + z_{i2} - e^{z_{i1}} - e^{z_{i2}}) \delta_{i1} (1 - \delta_{i2}) (z_{i1} - e^{z_{i1}} - e^{z_{i2}}) + \right. \\ &\left. + (1 - \delta_{i1}) \delta_{i2} (z_{i2} - e^{z_{i1}} - e^{z_{i2}}) - (1 - \delta_{i1}) (1 - \delta_{i2}) (e^{z_{i1}} + e^{z_{i2}})]\right\} \\ &\times \prod_{i=1}^n [1 + \tau g_1(z_{i1}) g_2(z_{i2})]^{\delta_{i1} \delta_{i2}} [1 + \tau g_1(z_{i1}) \rho_2(z_{i2})]^{\delta_{i1} (1 - \delta_{i2})} \\ &\times [1 + \tau \rho_1(z_{i1}) g_2(z_{i2})]^{\delta_{i2} (1 - \delta_{i1})} [1 + \tau \rho_1(z_{i1}) \rho_2(z_{i2})]^{(1 - \delta_{i1}) (1 - \delta_{i2})}, \end{aligned} \quad (19)$$

onde $\rho_j = \exp(-\exp\{\frac{y_{ij} - \mathbf{x}_j^\top \beta_j}{\sigma_j}\}) - 1$.

A densidade a posteriori de $\theta = (\beta_1, \beta_2, \sigma_1, \sigma_2)$ dada em (19) é intratável analiticamente, porque a integração da densidade a posteriori conjunta não é simples. Uma alternativa é utilizar MCMC tais como amostrador de Gibbs ou algoritmo de Metropolis-Hasting. Nesta direção obtemos inicialmente as condicionais completas para os parâmetros do modelo

$$\begin{aligned} \pi(\beta_1 | \beta_2, \sigma_1, \sigma_2, \tau, D) &\propto \exp\left\{\sum_{i=1}^n [\delta_{i1} \delta_{i2} (z_{i1} - e^{z_{i1}}) - \delta_{i1} (1 - \delta_{i2}) (z_{i1} - e^{z_{i1}}) \right. \\ &\left. - [(1 - \delta_{i1}) \delta_{i2} + (1 - \delta_{i1}) (1 - \delta_{i2})] e^{z_{i1}}]\right\} \phi_{q_1}(\beta_1 | \mu_1, \Sigma_1) \Psi(\beta_1, \beta_2, \sigma_1, \sigma_2, \tau), \\ \pi(\beta_2 | \beta_1, \sigma_1, \sigma_2, \tau, D) &\propto \exp\left\{\sum_{i=1}^n [\delta_{i1} \delta_{i2} (z_{i2} - e^{z_{i2}}) - \delta_{i1} (1 - \delta_{i2}) (z_{i2} - e^{z_{i2}}) \right. \\ &\left. - [(1 - \delta_{i1}) \delta_{i2} + (1 - \delta_{i1}) (1 - \delta_{i2})] e^{z_{i2}}]\right\} \phi_{q_2}(\beta_2 | \mu_2, \Sigma_2) \Psi(\beta_1, \beta_2, \sigma_1, \sigma_2, \tau), \\ \pi(\sigma_1 | \beta_1, \beta_2, \sigma_2, \tau, D) &\propto \exp\left\{-\frac{b_1}{\sigma_1} + \sum_{i=1}^n [\delta_{i1} \delta_{i2} (z_{i1} - e^{z_{i1}}) - \delta_{i1} (1 - \delta_{i2}) (z_{i1} - e^{z_{i1}}) \right. \\ &\left. - [(1 - \delta_{i1}) \delta_{i2} + (1 - \delta_{i1}) (1 - \delta_{i2})] e^{z_{i1}}]\right\} \Psi(\beta_1, \beta_2, \sigma_1, \sigma_2, \tau) \sigma_1^{-(2n + a_1 + 1)}, \end{aligned} \quad (20)$$

$$\begin{aligned} \pi(\sigma_2|\beta_1, \beta_2, \sigma_2, \tau, D) &\propto \exp\left\{-\frac{b_2}{\sigma_2} + \sum_{i=1}^n [\delta_{i1}\delta_{i2}(z_{i2} - e^{z_{i2}}) - \delta_{i1}(1 - \delta_{i2})(z_{i2} - e^{z_{i2}}) \right. \\ &\quad \left. - [(1 - \delta_{i1})\delta_{i2} + (1 - \delta_{i1})(1 - \delta_{i2})]e^{z_{i2}}]\right\} \Psi(\beta_1, \beta_2, \sigma_1, \sigma_2, \tau) \sigma_2^{-(2n+a_2+1)}, \\ \pi(\tau|\beta_1, \beta_2, \sigma_2, \tau, D) &\propto (1 - \tau)^{r_1-1} (1 + \tau)^{r_2-1} \Psi(\beta_1, \beta_2, \sigma_1, \sigma_2, \tau) \end{aligned}$$

onde

$$\begin{aligned} \Psi(\beta_1, \beta_2, \sigma_1, \sigma_2, \tau) &= \prod_{i=1}^n [1 + \tau g_1(z_{i1}) g_2(z_{i2})]^{\delta_{i1}\delta_{i2}} [1 + \tau g_1(z_{i1}) \rho_2(z_{i2})]^{\delta_{i1}(1-\delta_{i2})} \\ &\quad \times [1 + \tau \rho_1(z_{i1}) g_2(z_{i2})]^{(1-\delta_{i1})\delta_{i2}} [1 + \tau \rho_1(z_{i1}) \rho_2(z_{i2})]^{(1-\delta_{i1})(1-\delta_{i2})}. \end{aligned}$$

Observe que as condicionais completas dadas em (4.2) não são conhecidas, sendo necessário a utilização do algoritmo de Metropolis-Hasting para gerar amostras de $\beta_1, \beta_2, \sigma_1, \sigma_2$ e τ .

5 Aplicação

Nesta seção a metodologia desenvolvida é aplicada a três conjuntos de dados reais.

5.1 Modelo Weibull bivariado

Para ilustrar a metodologia apresentada anteriormente consideramos o conjunto de dados de tempos de falhas de cabos elétricos (em minutos) introduzidas por Lawless (2003, p.492). O mesmo conjunto de dados foi utilizado por Draper *et al.* (1989), em sua análise Bayesiana considerando a Distribuição BVE-tipo II de Gumbel (1960). As variáveis T_1 e T_2 são, respectivamente, o tempo inicial de falha no material e o tempo adicional subsequente para uma falha eventual. Para a análise Bayesiana os dados da Tabela 1 foram transformados em horas.

Tabela 1 - Dados de tempos de falha de cabos elétricos

t_1	288	106	246	700	473	155	414	1374	128
t_2	30	8	66	72	25	7	30	90	4
t_1	1227	254	435	1155	195	117	724	300	
t_2	39	46	85	85	27	27	21	96	

Assumindo a distribuição a priori (14) com valores de hiperparâmetros dados por $r_1 = r_2 = 1$, $a_1 = a_2 = b_1 = b_2 = 0.001$ e $p_1 = p_2 = q_1 = q_2 = 0.001$, geramos duas cadeias paralelas cada uma com 25.000 iterações; monitorou-se a convergência da amostras a partir do CODA do R-plus. Para cada parâmetro as 5.000 primeiras iterações foram descartadas para eliminar o efeito dos valores iniciais, em seguida foram tomadas amostras de 10 em 10 totalizando uma amostra final de tamanho

4.000. Na Tabela 2, reportamos o resumo a posteriori dos parâmetros do modelo juntamente com os resultados da estimativa dos fatores de redução de escala potencial \hat{R} (veja, Gelman & Rubin, 1992), para todos os parâmetros. Observamos valores bastante próximos de um, o que indica que as cadeias convergiram. Na Figura 1, mostramos os histogramas das distribuições marginais a posteriori dos parâmetros do modelo BVW, conjuntamente com os valores esperados estimados de T_1 e T_2 .

Tabela 2 - Sumário a posteriori dos parâmetros do modelo Weibull bivariado

Parâmetros	Média	Mediana	D.P	Int. Cred. (95%)	\hat{R}
α_1	1,252	1,241	0,231	(0,833 ; 1,735)	1,017
α_2	1,385	1,372	0,274	(0,891 ; 1,969)	1,041
λ_1	0,117	0,114	0,025	(0,076 ; 0,176)	1,021
λ_2	1,247	1,221	0,245	(0,840 ; 1,810)	1,001
τ	0,612	0,698	0,327	(-0,221 ; 0,988)	1,002
$E(T_1)$	8,565	8,271	1,879	(5,710 ;13,400)	1,009
$E(T_2)$	0,774	0,762	0,147	(0,522 ; 1,108)	1,001

Na Tabela 3, apresentamos os intervalos de credibilidade para o parâmetro τ para vários níveis de credibilidade. Os resultados desta tabela mostram que os dados indicam uma moderada dependência entre as variáveis T_1 e T_2 , isto é $\tau \neq 0$, embora que $\tau = 0$ esteja contido no intervalo de credibilidade de 90%.

Tabela 3 - Intervalos de credibilidade para τ

Intervalo de Credibilidade	LI	LS
50%	0,4518	0,8836
75%	0,2211	0,9442
80%	0,1464	0,9539
85%	0,0601	0,9661
90%	-0,0502	0,9656

Como ilustração considerando os dados da Tabela 1, obtemos a estimativa Bayesiana da confiabilidade de sistemas com dois componentes ligados em série e paralelo no tempo t_0 dado em (6) e (7), respectivamente os quais são apresentados na Tabela 4.

Análise de sensibilidade

Em nossa análise dos dados da Tabela 1 conduzimos uma análise de sensibilidade, considerando diferentes valores para os hiperparâmetros da distribuição a priori dos parâmetro modelo. A Tabela 5 mostra que as médias

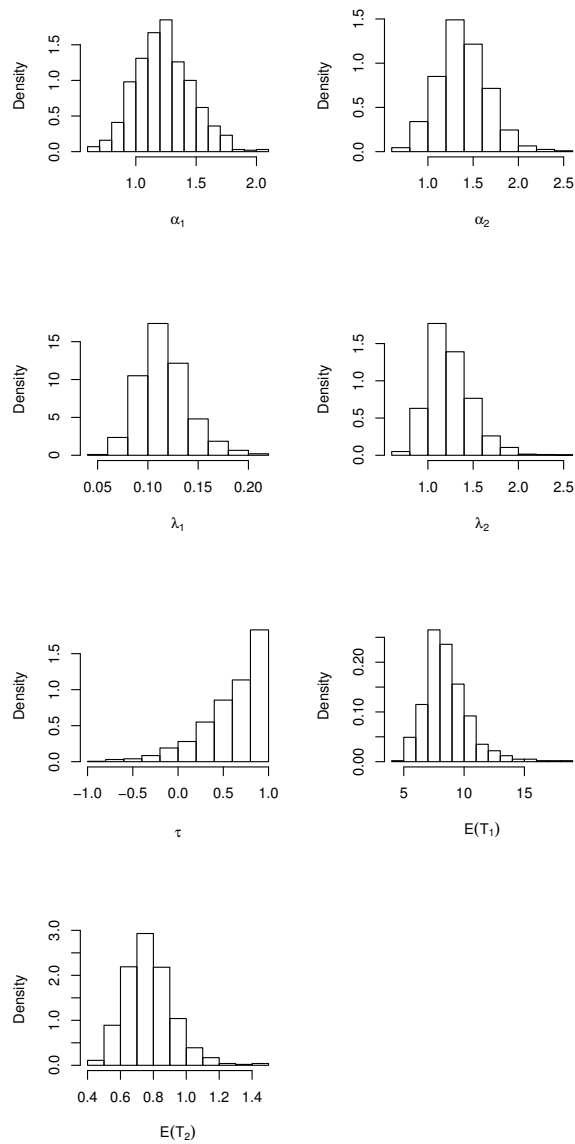


Figura 1 - Modelo Weibull bivariado dependente: Histograma da distribuição marginal a posteriori de, α_1 , α_2 , λ_1 , λ_2 , τ , $E(T_1)$ e $E(T_2)$.

Tabela 4 - Estimativa Bayesiana da confiabilidade de um sistema com dois componentes ligados em série e paralelo

t_0	$R_s(t_0)$	$D.P$	$R_p(t_0)$	$D.P$
0,00	1,00000	0,000000	1,0000	0,0000000
0,01	0,994700	0,006363	1,0000	0,0000158
0,25	0,800500	0,074040	0,9950	0,0055200
0,50	0,583600	0,091700	0,9802	0,0161900
0,75	0,400500	0,088440	0,9593	0,0274900
1,00	0,262600	0,079240	0,9362	0,0376300
1,25	0,167000	0,068730	0,9126	0,0462700
1,50	0,104500	0,057710	0,8894	0,0535700
1,75	0,065290	0,046960	0,8667	0,0597400
2,00	0,041170	0,037340	0,8445	0,0649800
2,25	0,026380	0,029320	0,8228	0,0694600
2,50	0,017250	0,022940	0,8014	0,0733000
3,00	0,007857	0,014270	0,7596	0,0794600
3,50	0,003874	0,009245	0,7191	0,0840000
4,00	0,002050	0,006251	0,6797	0,0872800
5,00	0,000686	0,003140	0,6048	0,0910100

a posteriori de cada parâmetro do modelo são próximas e não comprometem os resultados das Tabelas 2-4.

5.2 Modelo de regressão Weibull bivariado

Para ilustrar a metodologia apresentada na Seção 4.2, consideramos dois conjuntos de dados de retinopatia diabética. O primeiro é apresentado por Huster *et al.* (1989) e o segundo é apresentado por Achcar *et al.* (2005).

5.2.1 Dados de Huster *et al.* (1989)

Os dados consistem dos tempos de acompanhamento de 197 pacientes diabéticos com retinopatia com menos de 60 anos de idade nos Estados Unidos. O evento de interesse é a perda da acuidade visual. O objetivo principal do estudo é avaliar a eficácia do tratamento de fotocoagulação retardando o começo da perda da acuidade visual. O tratamento foi atribuído aleatoriamente a um olho de cada paciente. O outro olho foi considerado como controle. Então, o primeiro componente no vetor de tempos é o tempo até perda da acuidade visual no olho tratado (T_1) e o segundo componente é o tempo até perda da acuidade visual no olho sem tratamento (T_2). Os pacientes podem ser censurados, o que aconteceu para 73% dos olhos tratados e 49% dos olhos sem tratamento. Como covariável considerou-se a idade no diagnóstico da diabete (x), categorizada em 2 grupos, com ponto de corte em 20 anos (58% dos sujeitos foram considerados com menos de 20

Tabela 5 - Análise de sensibilidade dos hiperparâmetros da distribuição a priori dos parâmetros do modelo Weibull bivariado

Hiperparâmetros	Parâmetros	Média	D.P	Int. Cred. (95%)
$a_1 = b_1 = 0.1$	α_1	1,248	0,231	(0,822 ; 1,689)
$a_2 = b_2 = 0.1$	α_2	1,368	0,256	(0,903 ; 1,954)
$p_1 = q_1 = 0.1$	λ_1	0,120	0,022	(0,081 ; 0,182)
$p_2 = q_2 = 0.1$	λ_2	1,252	0,238	(0,832 ; 1,831)
$r_1 = r_2 = 0.5$	τ	0,614	0,382	(-0,217 ; 0,997)
$a_1 = b_1 = 0.01$	α_1	1,234	0,242	(0,804 ; 1,801)
$a_2 = b_2 = 0.01$	α_2	1,392	0,261	(0,901 ; 1,987)
$p_1 = q_1 = 0.01$	λ_1	0,115	0,024	(0,075 ; 0,181)
$p_2 = q_2 = 0.01$	λ_2	1,253	0,235	(0,819 ; 1,856)
$r_1 = r_2 = 2$	τ	0,615	0,289	(-0,231 ; 0,976)
$a_1 = b_1 = 0.00001$	α_1	1,227	0,305	(0,786 ; 1,912)
$a_2 = b_2 = 0.00001$	α_2	1,391	0,301	(0,865 ; 2,091)
$p_1 = q_1 = 0.00001$	λ_1	0,119	0,031	(0,056 ; 0,197)
$p_2 = q_2 = 0.00001$	λ_2	1,233	0,287	(0,812 ; 1,867)
$r_1 = r_2 = 3$	τ	0,622	0,338	(-0,242 ; 0,983)

anos de idade). Para essa covariável, 1 representa a categoria adulto e 0 a categoria jovem.

Esses dados foram analisados de uma perspectiva clássica por Huster *et al.* (1989) considerando a cópula de Clayton com distribuições marginais Weibull e por Therneau & Grambsch (2000) usando modelo de Cox com efeitos aleatórios . Desde uma perspectiva Bayesiana, Sahu & Dey (2000) consideraram esses dados em seu estudo comparativo entre os modelos de fragilidade e os modelos bivariados exponencial e Weibull.

Na Figura 2 temos as curvas de sobrevivência de Kaplan-Meier para cada olho, separados por faixa etária. Observa-se que, aparentemente, o tratamento é mais efetivo para pacientes diagnosticados com diabetes já adultos.

Em nossa análise consideramos o modelo dado em (11) com

$$Y_j = \log(T_j) = \beta_{0j} + \beta_{1j}x_j + \sigma_j U_j, \quad j = 1, 2.$$

onde U_1 e U_2 tem distribuição valor extremo bivariado dado em (10).

Assumindo que as distribuições a priori dada em (18) com valores dos hiperparâmetros dados por $r_1 = r_2 = 2$, $a_1 = a_2 = b_1 = b_2 = 0.001$, $\mu_{\beta_{k,j}} = 0$, $\sigma_{\beta_{k,j}}^2 = 100$ $k = 0, 1$ $j = 1, 2$, geramos duas cadeias paralelas cada uma com 25.000 iterações; monitorou-se a convergência da amostras a partir do CODA do software R-plus. Para cada parâmetro as 5.000 primeiras iterações foram descartadas para eliminar o efeito dos valores iniciais, em seguida foram tomadas amostras de 10 em 10 totalizando uma amostra final de tamanho 4.000 observações.

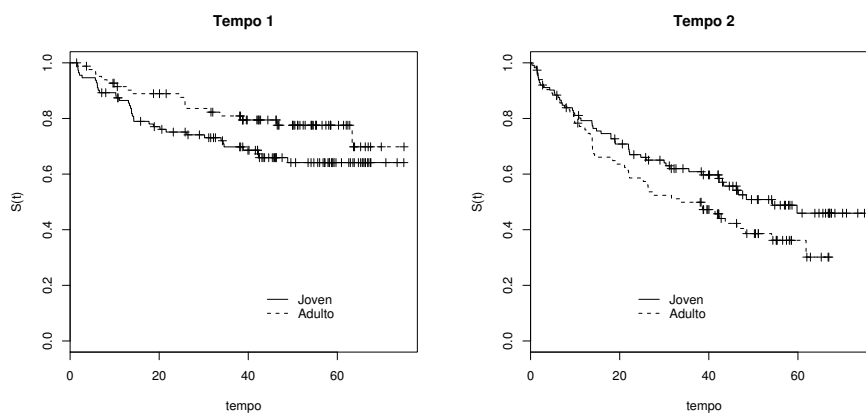


Figura 2 - Estimativa da função de sobrevivência para dados de retinopatia diabética, Tempo 1: grupo com tratamento , Tempo 2: grupo controls.

Tabela 6 - Sumário a posteriori dos parâmetros do modelo de regressão Weibull bivariado para os dados de Huster, *et al.* (1989)

Parâmetros				
Tempo T_1	Média	Desvio Padrão	Int. Cred.(95%)	\hat{R}
β_{01}	5,077	0,284	(4,595 ; 5,703)	0,999
β_{11}	0,642	0,392	(0,025 ; 1,341)	1,002
σ_1	1,311	0,173	(1,020 ; 1,704)	1,020
Tempo T_2				
β_{02}	4,456	0,183	(4,114 ; 4,839)	1,006
β_{12}	-0,453	0,247	(-0,9437 ; 0,019)	1,000
σ_2	0,842	0,484	(-0,944 ;0,019)	1,020
τ	0,700	0,161	(0,343 ; 0,952)	1,000

Na Tabela 6, apresentamos o resumo a posteriori dos parâmetros do modelo Weibull bivariado para os dados de Huster, *et al.* (1989), a partir dos resultados da tabela observamos que a covariável x (idade) é significativa nos dois tempos (T_1 e T_2). Os resultados desta tabela mostram ainda que os dados indicam uma dependência significativa entre as variáveis Y_1 e Y_2 . Esses resultados são compatíveis ao resultado obtido por Sahu & Dey (2000) e ao Romeo *et al.* (2007).

5.2.2 Dados de Achcar *et al* (2005)

Consideramos agora os dados de sobrevivência bivariados apresentados por Achcar *et al* (2005) referentes aos tempos até a ocorrência de retinopatia leve e grave em 71 pacientes portadores de *diabetes mellitus* tipo 1. Esses dados foram coletados no Hospital das Clínicas da Faculdade de Medicina de Riberão Preto, da universidade de São Paulo (HCFMRP-SP). Além dessas variáveis esses dados incluem as covariáveis sexo, cor, hipertensão arterial, nefropatia, neuropatia e controle glicêmico. O início da medição dos tempos (em anos) até a ocorrência da retinopatia foi contado a partir do diagnóstico de diabetes no HCFMRP-SP. Achcar *et al.* (2005) apresenta uma análise Bayesiana para esses dados considerando o modelo de riscos proporcionais, assumindo a distribuição Weibull para os tempos de sobrevivência com presença de efeitos aleatórios que captura a correlação entre as variáveis.

Considerando os dados de Achcar *et al.* (2005) é realizada uma análise Bayesiana assumindo que as variáveis T_1 (retinopatia leve) e T_2 (retinopatia grave) são variáveis aleatórias dependentes. Em nossa análise consideramos o modelo de regressão para as variáveis $Y_{i1} = \log(T_{i1})$ e o logaritmo do tempo até ocorrência de retinopatia leve, $Y_{i2} = \log(T_{i2})$ o logaritmo do tempo até ocorrência de retinopatia grave, e as covariáveis $\mathbf{X}^T = (x_1, x_2, \dots, x_7)$, sendo x_1 : sexo (0: masculino, 1: feminino), x_2 : cor de pele (0:branca, 1: não branca), x_3 presença de hipertensão arterial (0: não, 1: sim), x_4 nefropatia (0: não 1: sim) e x_5 neuropatia (0: não, 1 sim). O controle glicêmico, assumindo três possíveis categorias (bom, regular ou ruim), é representado por duas variáveis binárias, sendo x_6 : controle glicêmico bom (0: não, 1: sim) e x_7 : controle glicêmico regular (0: não, 1: sim). Desta forma, o modelo de regressão considerado é dado por,

$$Y_j = \beta_{0j} + \beta_{1j}x_{1j} + \beta_{2j}x_{2j} + \dots + \beta_{7j}x_{7j} + \sigma_j U_j, \quad j = 1, 2, \quad (21)$$

onde U_1 e U_2 tem distribuição valor extremo bivariado dado em (10).

Assumindo que as distribuições a priori (18) com valores dos hiperparâmetros dados por $r_1 = r_2 = 1$, $a_1 = a_2 = b_1 = b_2 = 0.001$, $\mu_{\beta_{kj}} = 0$, $\sigma_{\beta_{kj}}^2 = 100$ $k = 0, \dots, 7$ $j = 1, 2$, geramos duas cadeias paralelas cada uma com 35.000 iterações; monitorou-se a convergência das amostras a partir do CODA do software R-plus. Para cada parâmetro as 5.000 primeiras iterações foram descartadas para eliminar o efeito dos valores iniciais, em seguida foram tomadas amostras de 10 em 10 totalizando uma amostra final de tamanho 6.000 observações.

Na Tabela 7 são apresentados os resumos a posteriori dos parâmetros do modelo de regressão Weibull bivariado. Observamos que a covariável controle glicêmico regular apresenta um efeito significativo no tempo até a ocorrência da retinopatia leve. Além disso, observamos algum efeito das covariáveis sexo, cor e hipertensão arterial para o tempo até a ocorrência de retinopatia leve. Nenhuma covariável é significativa no tempo até a ocorrência de retinopatia grave.

Na Tabela 8, apresentamos os intervalos de credibilidade para o parâmetro τ para vários níveis de credibilidade. Os resultados desta tabela mostram que os

Tabela 7 - Sumário a posteriori dos parâmetros do modelo de regressão Weibull bivariado

Parâmetros				
Retinopatia leve	Média	Desvio Padrão	Int. Cred.(95%)	\hat{R}
β_{01}	3,022	0,151	(2,754 ; 3,349)	1,001
β_{11} (Sexo)	0,241	0,135	(-0,029 ; 0,504)	1,001
β_{21} (Cor)	-0,274	0,183	(-0,638 ; 0,090)	1,023
β_{31} (Hipertensão arterial)	-0,221	0,159	(-0,538; 0,091)	1,033
β_{41} (Nefropatia)	-0,231	0,156	(-0,547 ; 0,074)	1,001
β_{51} (Neuropatia)	0,131	0,162	(-0,182 ; 0,454)	1,015
β_{61} (CG-Bom)	0,532	0,489	(-0,204 ; 1,713)	1,011
β_{71} (CG-Regular)	0,292	0,151	(0,0183; 0,607)	1,046
σ_1	0,352	0,058	(0,260 ; 0,487)	1,020
Retinopatia grave				
β_{02}	5,582	1,811	(3,549 ; 10,43)	1,001
β_{12} (Sexo)	-0,411	0,855	(-2,652 ; 0,780)	1,017
β_{22} (Cor)	0,465	1,330	(-1,726 ; 3,802)	1,001
β_{32} (Hipertensão arterial)	-0,208	0,759	(-1,880 ; 1,307)	1,022
β_{42} (Nefropatia)	-0,941	0,952	(-3,381 ; 0,378)	1,001
β_{52} (Neuropatia)	-0,197	0,770	(-2,035 ; 1,208)	1,071
β_{62} (CG-Bom)	7,963	6,246	(-0,292 ; 23,330)	1,043
β_{72} (CG-Regular)	0,544	0,805	(-0,694 ; 2,622)	1,023
σ_2	0,842	0,484	(0,260 ; 0,487)	1,010
τ	0,683	0,287	(-0,042 ; 0,991)	1,002

dados indicam uma dependência significativa entre as variáveis Y_1 e Y_2 , isto é $\tau \neq 0$, embora, que $\tau = 0$ esteja contido no intervalo de credibilidade de 95%.

Tabela 8 - Intervalos de credibilidade para τ

Intervalo de Credibilidade	LI	LS
50%	0,545	0,897
75%	0,343	0,951
80%	0,272	0,961
85%	0,196	0,970
90%	0,085	0,981
92%	0,039	0,984

Com a finalidade de avaliar influência do parâmetro τ , ajustamos o modelo de regressão dada em (21) assumindo que T_1 e T_2 são independentes, isto é, $\tau = 0$. Nesse modelo consideramos as mesmas distribuições a priori do modelo com $\tau \neq 0$ com essa

escolha foram geradas duas cadeias cada uma com 35.000 iterações e monitorou-se a convergência das amostras a partir do CODA do R. Para cada parâmetro as 5.000 primeiras iterações foram descartadas e foram tomadas amostras de 10 em 10 totalizando uma amostra final de tamanho 6.000 observações. Na Tabela 9 apresentamos os resumos a posteriori dos parâmetros do modelo com $\tau = 0$.

Tabela 9 - Sumário a posteriori dos parâmetros do modelo de regressão Weibull bivariado considerando $\tau = 0$

Parâmetros	Média	Desvio Padrão	Int. Cred.(95%)	\hat{R}
Retinopatia leve				
β_{01}	3,027	0,1648	(2,750 ; 3,400)	1,001
β_{11} (Sexo)	0,242	0,142	(-0,055 ; 0,512)	1,000
β_{21} (Cor)	-0,284	0,193	(-0,657 ; 0,108)	1,023
β_{31} (Hipertensão arterial)	-0,209	0,168	(-0,554 ; 0,121)	1,012
β_{41} (Nefropatia)	-0,228	0,164	(-0,562 ; 0,084)	1,000
β_{51} (Neuropatia)	0,117	0,168	(-0,211 ; 0,459)	1,021
β_{61} (CG-Bom)	0,526	0,498	(-0,237 ; 1,729)	1,016
β_{71} (CG-Regular)	0,300	0,156	(0,004 ; 0,628)	1,016
σ_1	0,360	0,061	(0,260 ; 0,498)	1,030
Retinopatia grave				
β_{02}	6,155	2,046	(3,640 ; 11,560)	1,001
β_{12} (Sexo)	-0,533	0,974	(-2,992 ; 0,892)	1,003
β_{22} (Cor)	0,488	1,456	(-2,074 ; 4,045)	1,001
β_{32} (Hipertensão arterial)	-0,218	0,882	(-2,225 ; 1,519)	1,001
β_{42} (Nefropatia)	-1,120	1,046	(-3,740 ; 0,401)	1,004
β_{52} (Neuropatia)	-0,371	0,875	(-2,466 ; 1,222)	1,000
β_{62} (CG-Bom)	8,059	6,272	(-0,621 ; 22,440)	1,013
β_{72} (CG-Regular)	0,465	0,871	(-1,006 ; 2,558)	1,003
σ_2	0,979	0,538	(0,335 ; 2,375)	1,001

A partir da Tabela 9, observamos que a covariável controle glicêmico regular apresenta efeito significativo no tempo até a ocorrência de retinopatia leve (Y_1 .) Mas as covariáveis sexo, cor, hipertensão arterial, nefropatia, neuropatia e controle glicêmico bom não apresentam efeito significativo na variável Y_1 . Nenhuma das covariáveis apresenta um efeito significativo no logaritmo do tempo até a ocorrência de retinopatia grave.

Seguindo Raftery (1996), usamos a estimativa da log-verossimilhança marginal sob os modelos especificados, dados por, $\log f(\mathbf{t}_{obs}|M_1)$ e $\log f(\mathbf{t}_{obs}|M_2)$, onde M_1 e M_2 denotam respectivamente, o modelo de regressão Weibull dado em (11) e o modelo de regressão Weibull com $\tau = 0$, para decidir pelo melhor modelo. Na Tabela 10, mostramos a média e seu respectivo desvio padrão (D.P.) das log-verossimilhança marginais. Também, a razão $2 \log(f(\mathbf{t}_{obs}|M_1)/f(\mathbf{t}_{obs}|M_2)) = 7,12$, a qual direciona a evidência a favor do modelo de regressão Weibull dependente.

Tabela 10 - Log-verossimilhança marginal

	Modelo de Regressão Weibull Bivariado	
	Dependente	Independente
Média	- 62,100	-65,660
D.P.	4,973	4,662

6 Comentários finais

Neste artigo, propomos um modelo de distribuição Weibull bivariado (BVW), que permitem modelar dados de sobrevivência bivariados dependentes. Além disso, a partir do modelo BVW, construímos um modelo de regressão log-Weibull bivariado, que não precisa da suposição de riscos proporcionais como os modelos tradicionais de fragilidade.

O uso de técnicas Bayesianas, especialmente considerando técnicas de simulação de Monte Carlo em cadeias de Markov tem sido muito útil para a determinação de sumários a posteriori de interesse em modelos BVW com ou sem covariáveis.

Os códigos utilizados para a implementação da metodologia considerado os conjunto de dados apresentados na Seção 5 podem ser obtidos via email do primeiro autor.

BARRIGA, G. D. C.; LOUZADA NETO, F.; CANCHO, V. G. A bivariate Weibull regression model for survival data. *Rev. Bras. Biom.*, São Paulo, v.27, n.1, p.51-69, 2009.

■ **ABSTRACT:** *In this paper, we propose a new bivariate Weibull model (BVW), that it allows to model dependent survival data. The proposed model is motivated by an application of a result of Morgenstern (1956). An extension of this new model is also considered adding covariates to the model BVW. A Bayesian methodology based on methods of Monte Carlo Chains of Markov is developed for the proposed models. We illustrated the methodology with three set of real data.*

■ **KEYWORDS:** *Distribution bivariate; bivariate failure time; bayesian inference; bivariate Weibull distribution.*

Referências

ACHCAR, J. A.; GARCIA, C. A.; LAMAS, S. R.; MARTINEZ, E. Z.; PERDONÁ, G.C. Uma abordagem Bayesiana para dados multivariados de sobrevivência e binários na presença de covariáveis. *Rev. Mat. Estat.*, São Paulo, v.23, n.2 p.79-94, 2005.

- ACHCAR, J. A.; SANTANDER, L. A. M. Use of Markov chain Monte Carlo methods in a bayesian analysis of the Block & Basu bivariate exponential distribution. *J. Ital. Stat. Soc.*, Roma, v.3, p.233-250, 1993.
- BLOCK, H.; BASU, A.P. A continuous bivariate exponential extension. *J. Am. Stat. Assoc.*, Washington, v.69 p.1031-1037, 1974.
- CASELLA, G.; GEORGE, E. I. Explaining the gibbs sampler. *Am. Stat.*, Washington, v.46, p.167-174, 1992.
- CLATYON, D. G. A model for association in bivariate life-tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, Oxford, v.65, p.141-151, 1978.
- CHIB, S.; GREENBERG, E. Understanding the Metropolis-Hasting algorithm. *Am. Stat.*, Washington, v.49, n.4, p.327-335, 1995.
- DRAPER, N. R.; EVANS, M.; GUTTMAN, I. A Bayesian approach to system reliability when two component dependent. *Comp. Stat. Dat. Anal.*, Amsterdam, v.7, p.39-49, 1989.
- DOORNIK, J. *Ox: an object-oriented matrix programming language*. London: International Thomson Bussiness Press, 1996. 230p.
- FREUD, J. E. A bivariate extension of the exponential distribution. *J. Am. Stat. Assoc.*, Washington, v.56 p.971-977, 1961.
- GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences (with discussion) *Stat. Sci.*, Hayward, v.7, p.457-511, 1992.
- GUMBEL, E. J. Bivariate exponential distribution, *J. Am. Stat. Assoc.*, Washington, v.55 p.698-707, 1960.
- HOUGAARD, P. Fitting a multivariate failure time distribution. *IEEE Trans. Relia.* New York, v.38, p.444-448, 1989.
- HUSTER, W. J.; BROOKMEYER, R; SELF, S. G. Modeling paired survival data with covariates. *Biometrics*, Washington, v.45, p.145-156, 1989.
- JOE, H. (1997). *Multivariate models and dependence concepts*. London: Chapman & Hall, 1997. 349p.
- LAWLESS, J. F. *Statistical Models and Methods for lifetime data*. New York: John Wiley & Sons, 2003. 580p.
- LEE, L. Multivariate distributions having Weibull properties. *J. Multiv. Anal.*, Amsterdam, v.9, p.267-277, 1979.
- LU, J. C. Weibull extensions of the Freund and Marshall-Olkin bivariate exponential models. *IEEE Trans. Relia.*, New York, v.38, p.615-19, 1989.
- MARSHAL, A. W.; OLKIN, I. A. A multivariate exponential distribution, *J. Am. Stat. Assoc.*, Washington, v.62 p.30-44, 1967.
- MORGENSTERN, D. (1956). Einfache beispiele swedimensionaler vertielung. *Mitt. Math. Stat.*, Berlin, v.8, p.234-235, 1956.

OAKES, D. Bivariate survival models induced by frailties. *J. Am. Stat. Assoc.*, Washington, v.84, p.487-493, 1989.

RAFTERY, A. E. Hypothesis testing and selection models. In: *GILKS, W. R.; RICHARDSON, S.; SPIEGELHATER, D. J.(ed.)*. Markov chain Monte Carlo in practice. London: Chapman & Hall, 1996. p.163-187.

SAHU, S. K.; DEY, D. K. A comparison of frailty and other models for bivariate survival data. *Lifetime Data Anal.*, New York, v.6, p.207-228, 2000.

SHIH, J. H.; LOUIS, T. A. Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, Washington, v.51, p.1384-1399, 1995.

THERNEAU, T. M.; GRAMBSCH P. M. *Modeling survival data: extending the Cox Model*. New York: Springer-Verlag, 2000, 350p.

TIERNEY, L.; KADANE, J. B. Accurate approximations for posterior moments and marginal densities. *J. Am. Stat. Assoc.*, Washington, v.81 p. 82-86, 1986.

VAUPEL, J. W.; MANTON, K. G.; STALLARD, E. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, New York, v.16, p.439-454, 1979.

Recebido em 07.07.2008.

Aprovado após revisão em 13.05.2009.