

O MODELO LINDLEY-WEIBULL COM PROPORÇÃO DE CURA: UMA ABORDAGEM BAYESIANA

Patricia Gilavert FERNANDES¹
Adriano Kamimura SUZUKI¹
Erlandson Ferreira SARAIVA²

- RESUMO: Neste artigo propomos um novo modelo de sobrevivência de longa duração denominado Lindley-Weibull com proporção de cura (LWcr). Consideramos que o evento de interesse pode ser causado por M diferentes fatores de risco latente. Assumimos que M segue uma distribuição Lindley discreta e o tempo de ativação do j -ésimo fator segue uma distribuição Weibull. Para a estimação dos parâmetros de interesse, desenvolvemos uma abordagem Bayesiana com distribuições a priori não informativas. Para obter as estimativas dos parâmetros com base em suas distribuições a posteriori, utilizamos métodos de Monte Carlo em Cadeias de Markov (MCMC). Com o objetivo de identificar possíveis observações influentes, utilizamos o método Bayesiano de análise de influência de deleção de casos baseado na divergência ψ . Como ilustração, aplicamos o modelo proposto a conjuntos de dados simulados e reais.
- PALAVRAS-CHAVE: Análise de sobrevivência; distribuição Lindley-Weibull; modelo de sobrevivência de longa duração; proporção de cura; abordagem Bayesiana.

1 Introdução

No conjunto de dados de sobrevivência, a presença de uma grande quantidade de observações dos quais os maiores tempos são censurados pode indicar que existe uma fração (proporção) de indivíduos curados na população em estudo. Caso o gráfico do estimador de Kaplan-Meier da função de sobrevivência apresente cauda longa em um nível visivelmente acima de zero, as observações censuradas determinam uma possível presença de uma fração de curados.

¹Universidade de São Paulo - USP, Departamento de Matemática Aplicada e Estatística, CEP: 13566-590, São Carlos, SP, Brasil. E-mail: patriciagilavert@gmail.com; suzuki@icmc.usp.br

²Universidade Federal de Mato Grosso do Sul - UFMS, Instituto de Matemática, CEP: 79070-900, Campo Grande, MS, Brasil. E-mail: erlandson.saraiva@ufms.br

Os modelos de sobrevivência com fração de cura, também conhecidos como modelos de longa duração (MALLER e ZHOU, 1996; RODRIGUES *et al.*, 2008), têm sido amplamente desenvolvidos e utilizados em várias aplicações, como em dados de estudos clínicos (YAKOVLEV e TSODIKOV, 1996). Sob uma abordagem que leva em conta uma sequência estocástica de causas latentes, Cooner *et al.* (2007) induzem a ocorrência do evento de interesse por meio de mecanismos de ativação latentes. Nos últimos anos têm surgido vários trabalhos sobre distribuições que levam em consideração diferentes mecanismos de ativação com causas latentes competitivas (IBRAHIM *et al.*, 2001; TSODIKOV *et al.*, 2003). Por exemplo, Cancho *et al.* (2013), Suzuki *et al.* (2016) e Suzuki *et al.* (2017) consideraram que esse número de causas segue uma distribuição Binomial Negativa, Poisson e Logarítmica, respectivamente.

Neste artigo, propomos um novo modelo para análise de dados de sobrevivência de longa duração, denominado Lindley-Weibull com proporção de cura (LWcr). Assumimos que há diferentes mecanismos de ativação com causas latentes para o evento de interesse, cujo o número dessas causas é modelado por uma distribuição Lindley discreta (BAKOUCH *et al.*, 2014) e o tempo até a ocorrência do evento de interesse segue uma distribuição Weibull.

A principal motivação para a escolha da distribuição Lindley diz respeito às suas propriedades. De acordo com Bakouch *et al.* (2014), a distribuição uniparamétrica Lindley discreta tem expressões de forma fechada para sua função de distribuição cumulativa e função taxa de falha, o que não ocorre nos modelos Poisson e Gama discreta, por exemplo. Além disso, esta distribuição tem como caso particular a distribuição Geométrica, que é obtida quando o parâmetro da distribuição Lindley tende a zero.

Para estimação dos parâmetros, consideramos uma abordagem Bayesiana com distribuições *a priori* não informativas, *i.e.*, com “grandes” variâncias. Como as distribuições *a posteriori* condicionais dos parâmetros não possuem forma analítica “fechada”, obtemos as estimativas por meio do uso de métodos de Monte Carlo em Cadeias de Markov (MCMC).

Desenvolvemos um estudo de simulação para verificar o ajuste do modelo a conjuntos de dados artificiais, gerados sob três esquemas de ativação (primeiro, último e aleatório) do evento de interesse. No estudo de simulação, também consideramos conjuntos de dados com a presença de observações perturbadas. O interesse em conjuntos de dados com esta característica é que casos perturbados podem ser influentes e podem provocar distorções nas estimativas dos parâmetros. Para estes casos, utilizamos o diagnóstico de influência de deleção de casos com base na medida de divergência ψ (PENG e DEY, 1995; WEISS, 1996) para identificar as observações influentes.

A divergência ψ inclui várias medidas de divergência como casos particulares, tais como: a divergência Kullback-Leibler, a distância J , a distância L_1 e a medida de divergência χ^2 . Como ilustração, ajustamos o modelo proposto a um conjunto de dados reais referente a um ensaio clínico de melanoma cutâneo de fase III, baixado do website <http://merlot.stat.uconn.edu/~mhchen/survbook/> (IBRAHIM *et*

al., 2001). As implementações computacionais foram feitas utilizando o *software* R (R CORE TEAM, 2017).

O restante do artigo está organizado como segue. Nas Seções 2 e 3 apresentamos o modelo LWcr, a abordagem Bayesiana e a metodologia para identificação de casos influentes. Na Seção 4 desenvolvemos o estudo de simulação e a aplicação aos dados reais. Na Seção 5 fazemos as considerações finais.

2 Modelo Lindley-Weibull com proporção de cura (LWcr)

Seja M o número não observável de possíveis causas do evento de interesse para um indivíduo na população. Suponha que M segue uma distribuição de Lindley de parâmetro θ (BAKOUCH *et al.*, 2014), com função de probabilidade dada por:

$$P(M = m|\theta) = \frac{p^m}{1 + \theta} [\theta(1 - 2p) + (1 - p)(1 + \theta m)], \quad (1)$$

em que $p = \exp(-\theta)$, $\theta > 0$ e $m = 0, 1, 2, \dots$

A função geradora de probabilidade de M é dada por:

$$G_M(s) = \frac{\theta(1 - 2p) + 1 - p}{(1 + \theta)(1 - sp)} + \frac{\theta p(1 - p)s}{(1 + \theta)(1 - sp)^2}, \quad (2)$$

para $|s| < 1$.

Denote por Z_j o tempo para que a j -ésima causa produza o evento de interesse, para $j = 1, \dots, M$.

Assim como M , os tempos Z_j 's também são variáveis latentes (não observáveis). Assuma que os Z_j 's sejam independentes entre si e de M . Considere que Z_1, \dots, Z_M sejam variáveis aleatórias independentes e identicamente distribuídas com distribuição Weibull de parâmetros a e b , $Z_j \stackrel{i.i.d.}{\sim} Weibull(a, b)$, com função densidade de probabilidade dada por $f(z_j|a, b) = \frac{a}{b} \left(\frac{z_j}{b}\right)^{(a-1)} \exp\left\{-\left(\frac{z_j}{b}\right)^a\right\}$ e função de distribuição acumulada dada por $F(z_j|a, b) = 1 - \exp\left\{-\left(\frac{z_j}{b}\right)^a\right\}$, para $z_j > 0$, $a > 0$, $b > 0$ e $j = 1, \dots, M$.

De forma similar à Cooner *et al.* (2006), Cooner *et al.* (2007) e Cancho *et al.* (2013), assumimos que são necessárias a ocorrência de R das M causas, $R \leq M$, para que o evento de interesse ocorra. Assim, considerando $Z_{(1)} \leq \dots \leq Z_{(R)} \leq \dots \leq Z_{(M)}$ como sendo as M estatísticas de ordem, com $Z_{(R)}$ sendo a R -ésima estatística de ordem, então o tempo até a ocorrência do evento de interesse é dado por $Y = Z_{(R)}$. O valor R pode ser uma constante fixa, uma função de M ou uma variável aleatória discreta com distribuição condicional em M (CANCHO *et al.*, 2013). Se $M = 0$, o tempo para evento é assumido como infinito ($Y = +\infty$) com $P(Y = +\infty|M = 0) = 1$.

Neste artigo, consideramos as seguintes três especificações para R :

- (1) *Esquema de primeira ativação*. Assumimos que o evento de interesse ocorre devido a qualquer uma das possíveis causas. Neste caso, temos que $R = 1$

e o tempo de vida observado é definido pela variável aleatória $Y = Z_{(1)} = \min\{Z_1, \dots, Z_M\}$.

A função de sobrevivência de Y é dada por:

$$S_{\text{pop}}(y) = G_M(S(y)) = \frac{\theta(1-2p) + 1 - p}{(1+\theta)(1-S(y)p)} + \frac{\theta p(1-p)S(y)}{(1+\theta)(1-S(y)p)^2}, \quad (3)$$

em que $G_M(\cdot)$ é dado na expressão (2), $p = \exp\{-\theta\}$ e $S(y) = \exp\left\{\left(-\frac{y}{b}\right)^a\right\}$ é a função de sobrevivência da distribuição Weibull de parâmetros a e b .

A função densidade de probabilidade associada a (3) é dada por:

$$f_{\text{pop}}(y) = \frac{\frac{\theta(1-2p)+1-p}{1+\theta}f(y)p}{(1-S(y)p)^2} + \frac{\theta p(1-p)f(y)[1+S(y)p]}{(1+\theta)(1-S(y)p)^3}. \quad (4)$$

- (2) *Esquema de última ativação.* Assumimos que o evento de interesse ocorre somente após a ocorrência de todas as M causas. Neste caso, temos $R = M$ e o tempo de vida observado é dado por $Y = Z_{(M)} = \max\{Z_1, \dots, Z_M\}$. A função de sobrevivência de Y é dada por:

$$S_{\text{pop}}(y) = 1 + G_M(0) - G_M(1 - S(y)). \quad (5)$$

Assim, da expressão (2), a função de sobrevivência é dada por:

$$S_{\text{pop}}(y) = 1 + \frac{\theta(1-2p) + 1 - p}{1+\theta} - \frac{\theta(1-2p) + 1 - p}{(1+\theta)(1-(1-S(y))p)} - \frac{\theta p(1-p)(1-S(y))}{(1+\theta)(1-(1-S(y))p)^2}. \quad (6)$$

A função de sobrevivência dada em (6) leva à seguinte função densidade de probabilidade:

$$f_{\text{pop}}(y) = \frac{(\theta(1-2p) + 1 - p)f(y)p}{(1+\theta)(1-(1-S(y))p)^2} + \frac{\theta p(1-p)f(y)[1+(1-S(y))p]}{(1+\theta)(1-(1-S(y))p)^3}. \quad (7)$$

- (3) *Esquema de ativação aleatória.* Assumindo $M \geq 1$, consideramos para R uma distribuição uniforme discreta sobre $\{1, 2, \dots, M\}$. A função de sobrevivência de Y é dada por:

$$\begin{aligned} S_{\text{pop}}(y) &= P(Y > y) = P(M = 0) + \{1 - P(M = 0)\}S(y) \\ &= \frac{\theta(1-2p) + 1 - p + p(2\theta + 1)S(y)}{1 + \theta}. \end{aligned} \quad (8)$$

A primeira derivada, em relação a y , da função de sobrevivência $S_{\text{pop}}(y)$ dada

em (8), fornece a função densidade de probabilidade:

$$f_{\text{pop}}(y) = -S'_{\text{pop}}(y) = \frac{p(2\theta + 1)f(y)}{1 + \theta}, \quad (9)$$

em que $f(y) = -S'(y)$ é a primeira derivada de $-S(y)$ em relação a y .

Note que temos a mesma fração de cura nos três casos considerados, que é dada por:

$$p_0 = P(M = 0) = \lim_{y \rightarrow \infty} S_{\text{pop}}(y) = \frac{\theta(1 - 2p) + 1 - p}{1 + \theta}. \quad (10)$$

Embora a fração de cura p_0 seja a mesma para os esquemas (1), (2) e (3) considerados, os modelos diferem pelas suas funções de sobrevivência, de densidade e de risco. Uma propriedade referente às funções de sobrevivência dadas em (3), (8) e (6) é que para qualquer função de distribuição $F(y)$, temos $S_{\text{pop}}^{(1)}(y) \leq S_{\text{pop}}^{(2)}(y) \leq S_{\text{pop}}^{(3)}(y)$, para $y > 0$, em que $S_{\text{pop}}^{(1)}(y)$, $S_{\text{pop}}^{(2)}(y)$ e $S_{\text{pop}}^{(3)}(y)$ são as funções de sobrevivência dadas em (3), (8) e (6), respectivamente. Uma demonstração desta propriedade pode ser obtida no Teorema 2.1 em Kim *et al.* (2011).

Como ilustração, apresentamos no Apêndice A o gráfico da função densidade de probabilidade e da função de sobrevivência para alguns valores fixos de a , b e θ do modelo LWcr sob esquema de primeira ativação, última ativação e ativação aleatória, respectivamente.

3 Abordagem Bayesiana e comparação de modelos

Considere que o tempo de vida denotado por Y está sujeito a censura à direita. Seja C_i o i -ésimo tempo de censura. Assim, em uma amostra de tamanho n , o i -ésimo tempo observado T_i é dado por $T_i = \min\{Y_i, C_i\}$. Além disso, há o interesse no indicador de censura $\delta_i = I(Y_i \leq C_i)$, em que $\delta_i = 1$ se T_i é o tempo de falha e $\delta_i = 0$ se T_i for o tempo de censura, para $i = 1, \dots, n$.

Assumindo que temos disponível um vetor com p covariáveis, seja $\mathbf{x}^\top = (\mathbf{1}, \mathbf{x}_1, \dots, \mathbf{x}_p)$. Relacionamos o parâmetro θ da distribuição Lindley dada em (1) com as covariáveis \mathbf{x}_i , por meio da seguinte função de ligação:

$$\log(\theta_i) = \mathbf{x}_i^\top \boldsymbol{\beta}, \quad (11)$$

para $i = 1, \dots, n$, em que $\boldsymbol{\beta}^\top = (\beta_0, \beta_1, \dots, \beta_p)$ denota o vetor de parâmetros, de modo que para cada grupo de indivíduos, representado por \mathbf{x}_i , temos uma fração de cura diferente.

A função de verossimilhança é dada por:

$$L(\boldsymbol{\vartheta}; \mathcal{D}) \propto \prod_{i=1}^n f_{\text{pop}}(t_i; \boldsymbol{\vartheta})^{\delta_i} S_{\text{pop}}(t_i; \boldsymbol{\vartheta})^{1-\delta_i}, \quad (12)$$

em que $\boldsymbol{\vartheta} = (\boldsymbol{\beta}, a, b)^\top$, $\mathcal{D} = (\mathbf{t}, \boldsymbol{\delta}, \mathbf{x})$, $\mathbf{t} = (t_1, \dots, t_n)^\top$ e $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)^\top$, enquanto que $f_{\text{pop}}(\cdot; \boldsymbol{\vartheta})$ e $S_{\text{pop}}(\cdot; \boldsymbol{\vartheta})$ são as funções de densidade e de sobrevivência impróprias, respectivamente, descritas na Seção 2.

3.1 Distribuições a *priori* e a *posteriori*

Para estimação dos parâmetros, consideramos uma abordagem Bayesiana. Para isto, devemos especificar uma distribuição a *priori* para os parâmetros $\boldsymbol{\beta}$, a e b . Esta distribuição deve representar probabilisticamente o conhecimento que se têm sobre os parâmetros antes da realização do experimento.

Assumimos que $\boldsymbol{\beta}$, a e b são independentes a *priori*, isto é,

$$\pi(\boldsymbol{\vartheta}) = \pi(a)\pi(b) \prod_{j=0}^p \pi(\beta_j), \quad (13)$$

em que $\boldsymbol{\vartheta} = (\boldsymbol{\theta}, a, b)$.

Como os parâmetros a e b são definidos em $(0, +\infty)$, consideramos a transformação $\log(a)$ e $\log(b)$ e assumimos as seguintes distribuições a *priori*:

$$\log(a) \sim N(0, \sigma_a^2), \log(b) \sim N(0, \sigma_b^2) \text{ e } \beta_j \sim N(0, \sigma_j^2),$$

em que $N(0, \sigma^2)$ denota a distribuição normal com média 0 e variância σ^2 . Fixamos $\sigma_a^2 = \sigma_b^2 = \sigma_j^2 = 10^3$ para obtermos distribuições a *priori* pouco informativas, *i.e.*, com grandes variâncias, para $j = 0, \dots, p$.

Atualizando a distribuição a *priori* conjunta em (13), via função de verossimilhança $L(\boldsymbol{\vartheta}; \mathcal{D})$ dada como em (12), obtemos a distribuição a *posteriori* conjunta para $\boldsymbol{\vartheta}$ dada por:

$$\pi(\boldsymbol{\vartheta}|\mathcal{D}) \propto L(\boldsymbol{\vartheta}; \mathcal{D}) \exp \left\{ -\frac{1}{2} \left[\frac{(\log(a))^2}{\sigma_a^2} + \frac{(\log(b))^2}{\sigma_b^2} + \sum_{j=0}^p \frac{\beta_j^2}{\sigma_{\beta_j}^2} \right] \right\}. \quad (14)$$

Como a densidade a *posteriori* conjunta é analiticamente intratável, a inferência sobre os parâmetros é baseada nos métodos de simulação de Monte Carlo em cadeias de Markov (MCMC). Similarmente à Suzuki *et al.* (2016), utilizamos o algoritmo Metropolis-Hastings. Todas as implementações computacionais foram realizadas utilizando o *software* R (R CORE TEAM, 2017).

3.2 Critérios de comparação de modelos

Na literatura, diversas metodologias se propõem a analisar a adequabilidade de um modelo a um certo conjunto de dados, além de selecionar o melhor dentre uma coleção de modelos.

Neste artigo, utilizamos os critérios de seleção de modelos: LPML (*Logarithm of the Pseudo Marginal Likelihood*), EAIC (*Expected Akaike Information Criterion*),

EBIC (*Expected Bayesian*, ou *Schwarz, Information Criterion*) e DIC (*Deviance Information Criterion*).

O LPML é calculado com base nas ordenadas da densidade preditiva condicional (CPO): considere \mathcal{D} os dados completos, $\mathcal{D}^{(-i)}$ os dados com a i -ésima observação excluída e $\pi(\boldsymbol{\vartheta}|\mathcal{D}^{(-i)})$ a densidade a *posteriori* de $\boldsymbol{\vartheta}$ dado $\mathcal{D}^{(-i)}$, em que $i = 1, \dots, n$.

Para a i -ésima observação, a CPO_i é dada por:

$$CPO_i = \int_{\Theta} g(y_i|\boldsymbol{\vartheta})\pi(\boldsymbol{\vartheta}|\mathcal{D}^{(-i)})d\boldsymbol{\vartheta} = \left\{ \int_{\Theta} \frac{\pi(\boldsymbol{\vartheta}|\mathcal{D})}{g(y_i|\boldsymbol{\vartheta})} d\boldsymbol{\vartheta} \right\}^{-1}, \quad (15)$$

em que $g(y_i|\boldsymbol{\vartheta})$ é a função densidade de probabilidade correspondente ao modelo estudado.

Na comparação de vários modelos, valores altos de CPO_i indicam um melhor ajuste do modelo. Para uma discussão mais detalhada da estatística CPO e suas aplicações para seleção de modelos, ver os trabalhos de Gelfand *et al.* (1992) e Geisser *et al.* (1992) e Eddy (1979).

Para o modelo proposto não é possível obter uma forma analítica fechada de CPO_i . No entanto, uma estimativa de Monte Carlo de CPO_i pode ser obtida por meio de uma simples amostra MCMC da distribuição a *posteriori* $\pi(\boldsymbol{\vartheta}|\mathcal{D})$. Assim, considere $\boldsymbol{\vartheta}^{(1)}, \boldsymbol{\vartheta}^{(2)}, \dots, \boldsymbol{\vartheta}^{(M)}$ uma amostra de tamanho M de $\pi(\boldsymbol{\vartheta}|\mathcal{D})$ após o *burn-in* (aquecimento). Uma aproximação de Monte Carlo de CPO_i (CHEN *et al.*, 2000) é dada por:

$$\widehat{CPO}_i = \left\{ \frac{1}{M} \sum_{m=1}^M \frac{1}{g(y_i|\boldsymbol{\vartheta}^{(m)})} \right\}^{-1}. \quad (16)$$

A estatística LPML é dada por $LPML = \sum_{i=1}^n \log(\widehat{CPO}_i)$. O melhor modelo é aquele que corresponde ao maior valor da LPML.

Os critérios EAIC, EBIC e DIC são critérios calculados com base na média a *posteriori* da *deviance*, $E\{D(\boldsymbol{\vartheta})\}$, que é uma medida de ajuste e que pode ser aproximada por:

$$\bar{D} = \frac{1}{M} \sum_{m=1}^M D(\boldsymbol{\vartheta}^{(m)}),$$

em que $D(\boldsymbol{\vartheta}) = -2 \sum_{i=1}^n \log(g(y_i|\boldsymbol{\vartheta}))$ é a *deviance* e $g(\cdot)$ é a função densidade de probabilidade correspondente ao modelo estudado.

Os critérios EAIC, EBIC e DIC podem ser calculados, respectivamente, por:

$$\widehat{EAIC} = \bar{D} + 2q, \quad \widehat{EBIC} = \bar{D} + q \log(n) \quad \text{e} \quad \widehat{DIC} = 2\bar{D} - \hat{D},$$

em que q é o número de parâmetros do modelo e $\widehat{D} = D\left(\frac{1}{M} \sum_{m=1}^M \boldsymbol{\vartheta}^{(m)}\right)$.

Comparando modelos alternativos, o modelo preferido é aquele com os menores valores para estes três critérios.

3.3 Método Bayesiano de análise de influência de deleção de casos

O método da deleção de casos (COOK e WEISBERG, 1982) é uma ferramenta muito utilizada quando o objetivo é avaliar a influência de uma observação no ajuste de um modelo. As técnicas de verificação de influência local têm sido amplamente utilizadas, ver por exemplo Louzada *et al.* (2013) e Suzuki *et al.* (2016).

Neste trabalho, consideramos a análise de influência de deleção de casos baseada na divergência ψ que é calculada da seguinte maneira: seja $D_\psi(P; P_{(-i)})$ a divergência ψ entre P e $P_{(-i)}$, em que P indica a distribuição a *posteriori* de $\boldsymbol{\vartheta}$ para os dados completos e $P_{(-i)}$ a distribuição a *posteriori* sem a i -ésima observação. A medida de divergência ψ é definida como:

$$D_\psi(P, P_{(-i)}) = \int_{\boldsymbol{\vartheta} \in \Theta} \psi \left(\frac{\pi(\boldsymbol{\vartheta} | \mathcal{D}^{(-i)})}{\pi(\boldsymbol{\vartheta} | \mathcal{D})} \right) \pi(\boldsymbol{\vartheta} | \mathcal{D}) d\boldsymbol{\vartheta}, \quad (17)$$

em que $\psi(\cdot)$ é uma função convexa com $\psi(1) = 0$. Várias escolhas de ψ são dadas em Dey e Birimiwal (1994). Por exemplo, $\psi(z) = -\log(z)$ define a divergência de Kullback-Leibler (K-L), $\psi(z) = (z - 1) \log(z)$ a distância J (ou a versão simétrica da divergência de K-L), $\psi(z) = 0,5|z - 1|$ a distância variacional ou norma L_1 e $\psi(z) = (z - 1)^2$ define a divergência χ^2 .

Utilizando (15), a medida de divergência ψ pode ser escrita como:

$$D_\psi(P, P_{(-i)}) = E_{\boldsymbol{\vartheta} | \mathcal{D}} \left[\psi \left(\frac{CPO_i}{g(y_i | \boldsymbol{\vartheta})} \right) \right]. \quad (18)$$

A partir da expressão (18) podemos obter, por exemplo, a divergência K-L:

$$\begin{aligned} D_{K-L}(P, P_{(-i)}) &= -E_{\boldsymbol{\vartheta} | \mathcal{D}} \{ \log(CPO_i) \} + E_{\boldsymbol{\vartheta} | \mathcal{D}} \{ \log [g(y_i | \boldsymbol{\vartheta})] \} \\ &= -\log(CPO_i) + E_{\boldsymbol{\vartheta} | \mathcal{D}} \{ \log [g(y_i | \boldsymbol{\vartheta})] \}. \end{aligned} \quad (19)$$

As estimativas de Monte Carlo de $D_\psi(P, P_{(-i)})$ são dadas por:

$$\widehat{D}_\psi(P, P_{(-i)}) = \frac{1}{Q} \sum_{q=1}^Q \psi \left(\frac{\widehat{CPO}_i}{g(y_i | \boldsymbol{\vartheta}^{(q)})} \right). \quad (20)$$

Assim, obtemos a estimativa da divergência $D_{K-L}(P, P_{(-i)})$ da seguinte forma:

$$\widehat{D}_{K-L}(P, P_{(-i)}) = -\log(\widehat{CPO}_i) + \frac{1}{Q} \sum_{q=1}^Q \log [g(y_i | \boldsymbol{\vartheta}^{(q)})]. \quad (21)$$

A divergência $D_\psi(P, P_{(-i)})$ pode ser interpretada como sendo o efeito de excluir o caso i dos dados completos sobre a distribuição a *posteriori* de ϑ . Como discutido por Peng e Dey (1995) e Weiss (1996), pode ser difícil para um pesquisador definir um ponto de corte para o valor da medida de divergência, de modo a determinar se um pequeno subconjunto de observações é influente ou não. Neste trabalho utilizamos a proposta de Peng e Dey (1995) e Weiss (1996) que consideram uma moeda viesada com probabilidade de sucesso p . Assim, a divergência ψ entre as moedas viesadas é dada por:

$$D_\psi(f_0, f_1) = \int \psi \left(\frac{f_0(x)}{f_1(x)} \right) f_1(x) dx, \quad (22)$$

em que $f_0(x) = p^x(1-p)^{1-x}$ e $f_1(x) = 0,5$, $x = 0$ ou 1 . Se $D_\psi(f_0, f_1) = d_\psi(p)$, pode ser verificado que d_ψ satisfaz a seguinte equação:

$$d_\psi(p) = \frac{\psi(2p) + \psi(2(1-p))}{2}. \quad (23)$$

É possível ver que d_ψ aumenta à medida que p se afasta de $0,50$. Além disso, $d_\psi(p)$ é simétrico em torno de $p = 0,50$ e d_ψ atinge seu mínimo em $p = 0,50$. Neste ponto, $d_\psi(0,5) = 0$ e $f_0 = f_1$. Se considerarmos $p > 0,80$ (ou $p < 0,20$) um forte viés em uma moeda, então $d_{L_1}(0,80) = 0,30$. Esta equação implica que o caso i é considerado influente quando $d_{L_1} > 0,30$. Assim, se a divergência de Kullback-Leibler for usada, pode-se considerar uma observação influente quando $d_{K-L} > 0,2231$. Da mesma forma, se a distância J e a divergência χ^2 forem usadas, uma observação pode ser considerada influente quando $d_J > 0,4159$ e $d_{\chi^2} > 0,5625$, respectivamente. Neste trabalho utilizamos estes valores de corte para determinar se uma observação é influente.

4 Resultados

Nesta seção apresentamos um estudo de simulação, o diagnóstico de observações influentes e uma aplicação a dados reais.

4.1 Estudo de simulação

Realizamos um estudo de simulação com dois objetivos: *i*) avaliar as estimativas dos parâmetros para os modelos propostos bem como um estudo para verificar se é possível distinguir entre os modelos ajustados a um conjunto de dados específico com base nos critérios de seleção de modelos adotados; *ii*) examinar o desempenho da abordagem diagnóstica proposta, considerando conjuntos de dados simulados sem nenhum, um ou dois casos perturbados.

Neste estudo consideramos o modelo LWcr sob os três esquemas de ativação (primeira, última e aleatória) com a distribuição Weibull com parâmetros $a = 2$ e $b = 5$ para os tempos Z_j , $j = 1, 2, \dots, M$.

Para cada indivíduo i , $i = 1, \dots, n$, o número de causas do evento de interesse, M_i é gerado a partir de uma distribuição Lindley com parâmetro θ_i . Nas simulações, tem-se uma covariável binária x com valores obtidos por meio de uma distribuição Bernoulli com parâmetro 0,50. Portanto, $\theta_{0i} = \exp(\beta_0 + \beta_1 x_i)$. Escolhemos $\beta_0 = -0,50$ e $\beta_1 = 0,20$ de modo que a proporção de cura para os dois níveis de x é $p_0^{(0)} = 0,19$ e $p_0^{(1)} = 0,25$, respectivamente. Os tempos de censura foram amostrados a partir da distribuição Uniforme no intervalo $(0, \tau)$, em que τ controla a proporção de censura da população não curada. No estudo, a proporção de observações censuradas é de 55%, aproximadamente.

4.1.1 Propriedades frequentistas

Este estudo de simulação objetiva demonstrar o bom comportamento das estimativas Bayesianas, com base no erro quadrático médio (EQM) e na média a *posteriori*. Também, investigar se os critérios de comparação de modelos discriminam o verdadeiro modelo. Geramos amostras de tamanho $n = 300$ para as três diferentes estruturas de ativação. Portanto, construímos três diferentes configurações de simulação, cada uma com 600 conjuntos de dados gerados. Uma vez simulados os dados, fizemos o ajuste dos três modelos. As seguintes distribuições a *priori* independentes foram consideradas para a execução do algoritmo Metropolis-Hastings: $\beta_j \sim N(0, 10^3)$, $j = 0, 1$, $\log(a) \sim N(0, 10^3)$ e $\log(b) \sim N(0, 10^3)$. Assim, assumimos uma distribuição a *priori* própria e pouco informativa, implicando que a distribuição a *posteriori* será sempre própria. Para eliminar o efeito dos valores iniciais, desconsideramos as primeiras 10.000 iterações.

A convergência do algoritmo Metropolis-Hastings foi monitorada usando o método proposto por Geweke (1992) e o gráfico de traços.

Para cada amostra registramos a média a *posteriori* do parâmetro e os valores DIC, EAIC, EBIC e LPML. As estatísticas de resumo da simulação para os parâmetros, assumindo os três modelos cruzados com os três esquemas de ativação simulados são dadas na Tabela 1, em que Média denota a média aritmética das 600 estimativas dadas por $\sum_{j=1}^{600} \hat{\theta}_{kj} / 600$, EQM é o erro quadrático médio empírico

dado por $\sum_{j=1}^{600} (\hat{\theta}_{kj} - \theta_k)^2 / 600$, em que θ_k é o valor do parâmetro conhecido e $\hat{\theta}_k$ é o valor estimado. Observe que as médias são muito próximas do valor real e EQM foi pequeno quando o ajuste foi realizado para o verdadeiro esquema de ativação.

A Tabela 2 mostra a porcentagem de amostras cuja distribuição gerada foi considerada a melhor de acordo com os critérios DIC, EAIC, EBIC e LPML. O verdadeiro modelo, a partir do qual a amostra foi gerada, mostra uma porcentagem mais elevada.

4.2 Diagnóstico de observações influentes

Consideramos conjuntos de dados simulados com nenhum, um ou dois casos perturbados para examinar o desempenho das medidas de diagnóstico propostas.

Uma amostra de tamanho 400 foi gerada do modelo LWcr sob o esquema de primeira ativação.

Tabela 1 - Resultados das amostras simuladas cruzadas com os três esquemas de ativação

Verdadeiro modelo	Parâmetro	Modelo ajustado					
		Primeira Ativação		Última Ativação		Ativação Aleatória	
		Média	EQM	Média	EQM	Média	EQM
Primeira	a	2,0111	0,0205	1,3480	0,4358	1,7581	0,0708
	b	5,1381	0,7718	2,0223	8,8863	3,1923	3,2972
	β_0	-0,6892	0,0276	-0,4990	0,0551	-0,6122	0,0289
	β_1	0,1713	0,0209	-0,0781	0,1073	0,1197	0,0455
Última	a	3,3984	2,0057	2,2584	0,0954	2,9451	0,9278
	b	8,0202	9,2767	4,9394	0,0485	6,4041	2,0071
	β_0	-0,4877	0,0595	-0,5099	0,0513	-0,5201	0,0503
	β_1	-0,0153	0,0669	-0,0072	0,0702	-0,0133	0,0833
Aleatória	a	2,3649	0,1604	1,6331	0,1508	2,0661	0,0233
	b	6,4623	2,5789	3,2876	2,9682	4,6415	0,1784
	β_0	-0,5471	0,0438	-0,4697	0,0703	-0,5738	0,0405
	β_1	0,1947	0,0199	0,1229	0,0358	0,2706	0,0451

Tabela 2 - Porcentagens de amostras cujo modelo ajustado foi considerado o melhor de acordo com os critérios DIC, EAIC, EBIC e LPML

Verdadeiro modelo	Modelo ajustado											
	Primeira Ativação				Última Ativação				Ativação aleatória			
	DIC	EAIC	EBIC	LPML	DIC	EAIC	EBIC	LPML	DIC	EAIC	EBIC	LPML
Primeira	74,54	76,04	76,04	76,21	4,99	2,66	2,66	3,16	20,47	21,30	21,30	20,63
Última	4,16	4,33	4,33	4,49	71,21	70,88	70,88	71,71	24,63	24,79	24,79	23,79
Aleatória	27,95	26,79	26,79	27,45	23,29	23,63	23,63	24,96	48,76	49,58	49,58	47,59

Nos dados simulados, y_i variou de 4×10^{-6} à 10,87 com mediana = 2,54, média = 3,18 e desvio padrão = 2,41. Os casos 34, 109 e 398 foram selecionados para a perturbação.

Para criar uma observação influente no conjunto de dados, escolhemos um e dois casos selecionados de forma aleatória, e perturbamos a variável resposta da seguinte forma: $\tilde{y}_i = y_i + 4S_y$, $i = 34, 109$ e 398 , em que S_y é o desvio padrão dos y_i 's. Realizamos as simulações via MCMC de forma semelhante à Seção 4.1.1.

O modelo LWcr foi ajustado sob o esquema de primeira ativação. Na Tabela 3 observamos que as inferências a *posteriori* são sensíveis após a perturbação do(s) caso(s) selecionado(s), exceto para o parâmetro β_1 . O conjunto de dados (a) denota o conjunto de dados original simulado sem perturbação e os conjuntos de dados (b)-(f) denotam conjuntos de dados com casos perturbados.

Tabela 3 - Média e Desvio Padrão (DP) das estimativas dos parâmetros obtidas a partir do ajuste do modelo LWcr sob o esquema de primeira ativação para cada conjunto de dados

Conjunto de dados	Casos perturbados	Parâmetros							
		a		b		β_0		β_1	
		Média [DP]	Média [DP]	Média [DP]	Média [DP]	Média [DP]	Média [DP]	Média [DP]	Média [DP]
(a)	Nenhum	2,0913 [0,1242]	5,0283 [0,5129]	-0,5590 [0,1272]	0,0491 [0,1232]				
(b)	34	1,8352 [0,1079]	6,9790 [1,6122]	-0,8020 [0,2094]	0,0478 [0,1310]				
(c)	109	1,8327 [0,1087]	7,0116 [1,5751]	-0,8064 [0,2100]	0,0494 [0,1326]				
(d)	398	1,8148 [0,1053]	7,2685 [1,9027]	-0,8391 [0,2259]	0,0535 [0,1338]				
(e)	{34, 109}	1,7244 [0,1037]	9,9071 [9,2291]	-1,0501 [0,3861]	0,0380 [0,1378]				
(f)	{109, 398}	1,7161 [0,1012]	10,2359 [9,3767]	-1,0831 [0,4091]	0,0494 [0,1387]				

Para monitorar a convergência das amostras, utilizamos os métodos recomendados por Cowless e Carlin (1996) e também o gráfico de traços, como o apresentado na Figura 1 para o ajuste ao conjunto de dados (a).

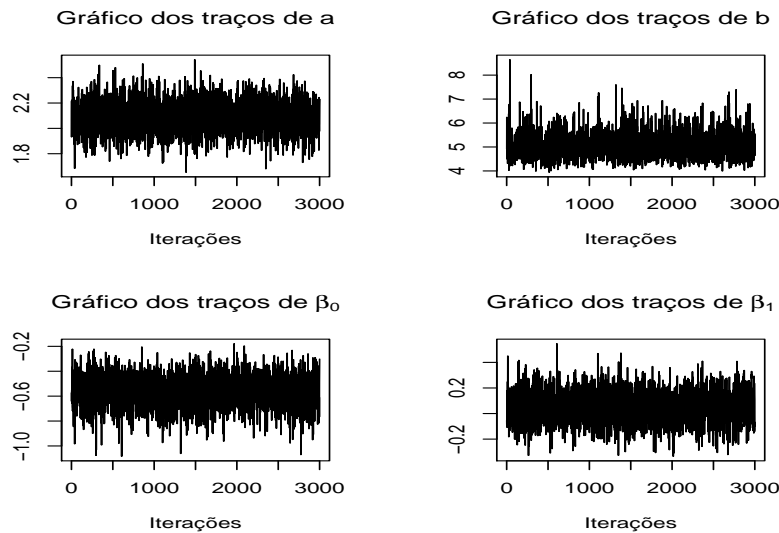


Figura 1 - Gráfico de traços.

A Tabela 4 mostra os valores obtidos dos critérios DIC, EAIC, EBIC e LPML para cada versão perturbada dos dados originais. Como esperado, o conjunto de dados (a) foi o que melhor se ajustou ao modelo LWcr sob o esquema de primeira ativação.

Tabela 4 - Critérios Bayesianos para cada conjunto de dados

Conjunto de dados	Critérios Bayesianos			
	DIC	EAIC	EBIC	LPML
(a)	1.132,7770	1.136,8180	1.152,7840	-566,3869
(b)	1.151,2690	1.155,2790	1.171,2450	-578,7066
(c)	1.151,4720	1.155,4370	1.171,4030	-578,3573
(d)	1.151,9110	1.155,9420	1.171,9080	-578,8259
(e)	1.162,2540	1.166,2260	1.182,1920	-582,5929
(f)	1.162,3210	1.166,3370	1.182,3030	-582,6900

Consideramos agora a amostra das distribuições *a posteriori* dos parâmetros do modelo LWcr sob o primeiro esquema de ativação para calcular as medidas de divergência ψ em (17) descritas na Seção 3.3. Os resultados na Tabela 5 mostram que antes da perturbação (conjunto de dados (a)) os casos selecionados não são influentes de acordo com todas as medidas de divergência ψ . No entanto, após a perturbação (conjunto de dados (b)-(f)) as medidas aumentam, o que indica que os casos perturbados são influentes.

Tabela 5 - Divergência ψ para os dados simulados do modelo LWcr sob o esquema de primeira ativação

Conjunto de dados	Número do caso	d_{K-L}	d_J	d_{L_1}	d_{χ^2}
(a)	34	0,0223	0,0452	0,0843	0,0481
	109	0,0265	0,0539	0,0919	0,0580
	398	0,0644	0,1327	0,1433	0,1565
(b)	34	3,3321	8,0500	0,8628	651,0800
(c)	109	2,8780	6,3652	0,8149	238,9346
(d)	398	3,1346	6,9207	0,8367	289,3834
(e)	34	0,7827	1,8648	0,5012	12,4008
	109	0,9988	2,4106	0,5597	23,0884
(f)	109	0,8730	2,0899	0,5287	15,3656
	398	0,9861	2,3600	0,5576	20,1190

As Figuras 2, 3 e 4 mostram as medidas de divergência ψ para os casos (a), (c) e (f), respectivamente. Na Figura 2, observamos que não foram detectados pontos influentes como era esperado, pois nenhuma observação foi perturbada no conjunto de dados (a). Nas Figuras 3 e 4 as medidas de divergências detectaram as observações perturbadas como pontos influentes.

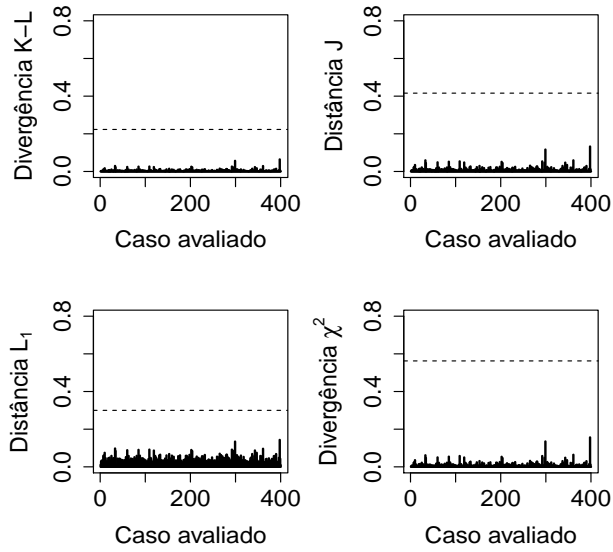


Figura 2 - Divergência ψ para o conjunto de dados (a).

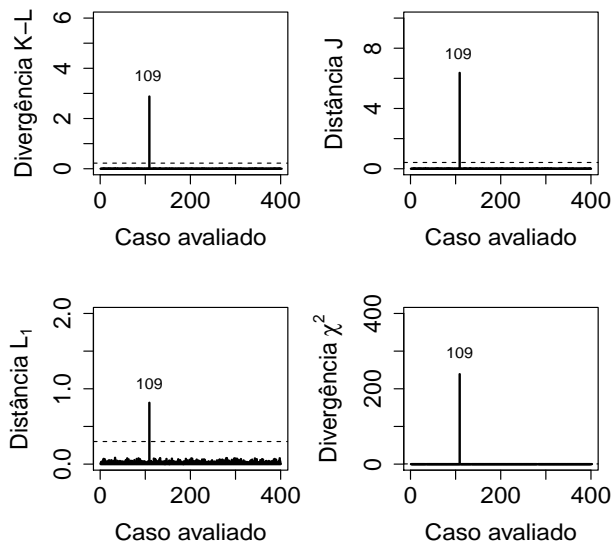


Figura 3 - Divergência ψ para o conjunto de dados (c).

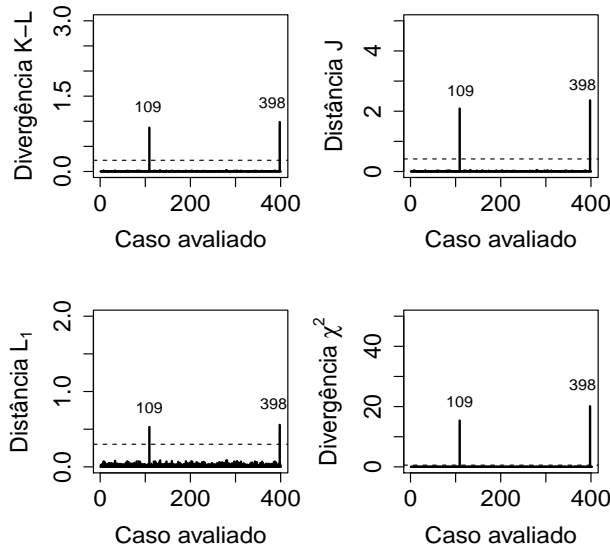


Figura 4 - Gráfico da divergência ψ para o conjunto de dados (f).

4.3 Aplicação a dados reais

Esta seção relata a aplicação do modelo proposto a um conjunto de dados reais de um ensaio clínico de melanoma cutâneo de fase III (rotulado como dados E1690) conduzido pelo Eastern Cooperative Oncology Group e disponível em <http://merlot.stat.uconn.edu/~mhchen/survbook/> (IBRAHIM *et al.*, 2001).

O conjunto de dados é o resultado de um ensaio para a avaliação do desempenho do tratamento pós-operatório com uma dose elevada de interferon alfa-2b para prevenir a recorrência. Os pacientes foram incluídos no estudo entre os anos de 1991 a 1995 e o acompanhamento foi até o ano de 1998.

O tempo de sobrevivência, sem recidiva em anos, foi considerado como variável resposta. Após excluir indivíduos com dados incompletos e tempos de observação faltantes, obtivemos um subconjunto de $n = 205$ pacientes com aproximadamente 72% de censura. As seguintes variáveis foram coletadas de cada paciente: tempo observado (em anos) e sexo (covariável). Primeiro, foram ajustados os três modelos sob a primeira, última e aleatória ativações. Para todos os modelos as seguintes distribuições *a priori* independentes foram adotadas: $\beta_j \sim N(0, 10^3)$, $j = 0, 1$, $\log(a) \sim N(0, 10^3)$ e $\log(b) \sim N(0, 10^3)$. Após desconsideradas as primeiras 20.000 iterações para eliminar o efeito dos valores iniciais e um salto de tamanho 10 para evitar problemas de autocorrelação, obtivemos uma amostra efetiva de tamanho 40.000 sobre a qual a inferência *a posteriori* foi baseada. A convergência

do algoritmo Metropolis-Hastings foi monitorada usando o método proposto por Geweke (1992) e os gráficos de traços.

Os valores dos critérios DIC, EAIC, EBIC e LPML foram obtidos para comparar o modelo LWcr sob os esquemas de ativação primeira, última e aleatória. Esses critérios de informação forneceram os valores da Tabela 6.

Tabela 6 - Critérios Bayesianos para os modelos ajustados

Ativação	Critério			
	DIC	EAIC	EBIC	LMPL
Primeira	453,9805	458,0596	471,3517	-226,9593
Última	457,7821	462,3170	475,6090	-229,3377
Aleatória	455,2088	459,4560	472,7481	-227,6106

O modelo LWcr sob a primeira ativação foi o que apresentou o melhor ajuste de acordo com todos os critérios apresentados na Tabela 6. Assim, o modelo LWcr sob o esquema de primeira ativação foi selecionado como o melhor modelo. O resumo a *posteriori* de seus parâmetros é apresentado na Tabela 7. As amostras

Tabela 7 - Resumo a *posteriori* dos parâmetros do modelo LWcr sob o esquema de primeira ativação

Parâmetro	Média	Desvio padrão	Intervalo HPD (95%)	
			LI	LS
a	1,6737	0,2104	1,2837	2,1017
b	6,1817	2,1622	4,3102	11,2741
β_0	0,1136	0,1901	-0,2837	0,4070
β_1	0,4029	0,1640	0,0935	0,7315

consideradas para as distribuições a *posteriori* dos parâmetros do modelo LWcr sob a primeira ativação e as medidas divergência ψ em (17), descritas na Seção 3.3, foram computadas. As medidas de divergências não detectaram observações influentes, como podemos observar no gráfico de índices das quatro medidas de divergência ψ apresentado na Figura 5.

A Figura 6 mostra as curvas de Kaplan-Meier e as curvas de sobrevivência LWcr estimadas, estratificadas pelo sexo, onde observou-se um patamar acima de zero para os últimos tempos observados, sendo um indicativo de que modelos que ignoram a possibilidade de cura não são adequados para estes dados. Esta figura também apresenta a curva de sobrevivência do modelo LWcr sob o esquema de primeira ativação estimada, na qual é possível observar o bom ajuste do modelo aos dados.

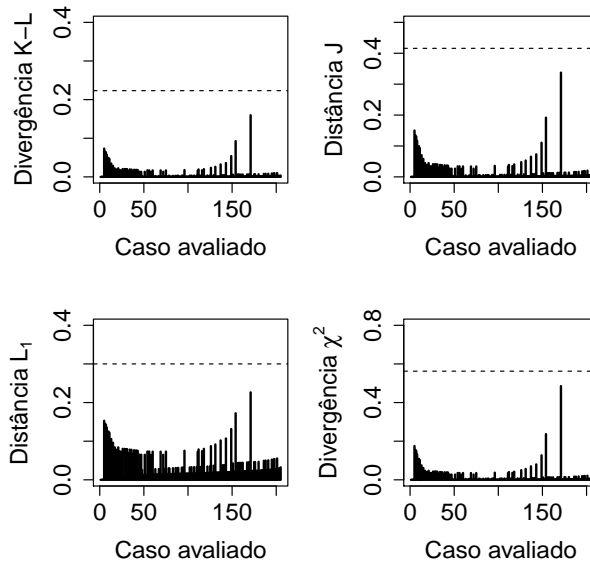


Figura 5 - Divergência ψ para os dados E1690.

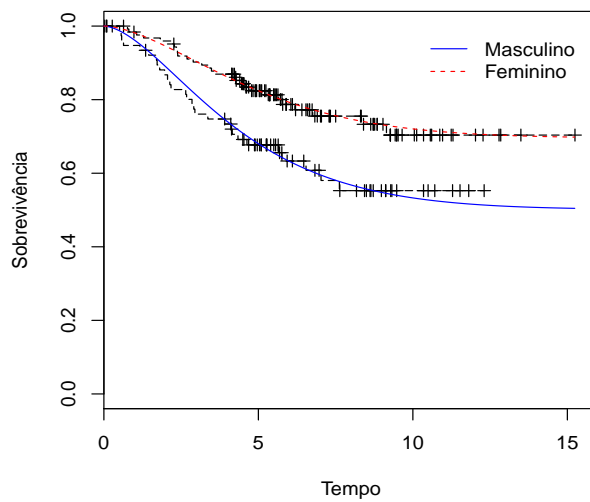


Figura 6 - Curvas de Kaplan-Meier e curvas de sobrevivência LWcr estimadas para os dados E1690.

5 Considerações finais

Neste artigo apresentamos o modelo LWcr para a modelagem de dados de sobrevivência com proporção de cura considerando que o evento de interesse pode ser causado por três esquemas de ativação latentes. Assumimos que os tempos de ativação latentes são provenientes de um distribuição Weibull de parâmetros a e b .

Para a estimação dos parâmetros de interesse assumimos uma abordagem Bayesiana com distribuições *a priori* não informativas. O procedimento de estimação dos parâmetros foi feito utilizando métodos de Monte Carlo em Cadeias de Markov. Além disso, apresentamos uma discussão sobre a seleção do modelo sob primeira, última e ativação aleatória, por meio do uso de critérios de seleção de modelos; e uma discussão sobre a análise de deleção de casos utilizando a medida de divergência ψ .

Ilustramos a performance do procedimento de estimação dos parâmetros do modelo proposto por meio de dados simulados e um conjunto de dados real. Para os dados simulados, os critérios de seleção de modelos indicaram o modelo sob o qual os dados foram gerados como sendo o melhor modelo na maioria dos casos simulados. Para o conjunto de dados real, o modelo LWcr sob o esquema de primeira ativação forneceu o melhor ajuste.

Pelo menos a princípio, outras distribuições para os tempos de ativação latente podem ser consideradas, além da extensão do modelo proposto para o caso multivariado.

Agradecimentos

Aos revisores e editores pelas sugestões que propiciaram um melhor entendimento do artigo.

FERNANDES, P. G.; SUZUKI, A. K.; SARAIVA, E. F. The Lindley-Weibull model with cure rate: a Bayesian approach. *Rev. Bras. Biom.*, Lavras, v.36, n.4, p.998-1022, 2018.

- **ABSTRACT:** *In this work we propose a new long-term survival model called Lindley-Weibull with a cure rate (LWcr). We consider that the event of interest can be caused by M different latent risk factors. We assume that M follows a discrete Lindley distribution and the activation time of j -th factor follows a Weibull distribution. For the estimation of the parameters of interest we developed a Bayesian approach with non-informative prior distributions. To obtain the parameter estimates based on their posterior distributions we used Monte Carlo Markov Chain (MCMC) methods. In order to detect influential observations it was used the Bayesian method of influence analysis of case deletion based on ψ -divergence. We apply the proposed model to simulated and an actual dataset as an illustration.*
- **KEYWORDS:** *Survival analysis; Lindley-Weibull distribution; long-term survival model; cure rate; Bayesian approach.*

Referências

BAKOUCH, H. S., JAZI, M. A.; NADARAJAH, S. A new discrete distribution. *Statistics*, v.48, n.1, p.200–240, 2014.

CANCHO, V. G., BANDYOPADHYAY, D., LOUZADA, F.; YIQI, B. The destructive negative binomial cure rate model with a latent activation scheme. *Statistical Methodology*, v.13, p.48–68, 2013.

CANCHO, V. G., SUZUKI, A. K., BARRIGA, G. D. C.; LOUZADA, F. A non-default fraction bivariate regression model for credit scoring: An application to Brazilian customer data. *Communications in Statistics: Case Studies, Data Analysis and Applications*, v.2, p.1–12, 2016.

CHEN, M. H., SHAO, Q. M.; IBRAHIM, J. G. *Monte Carlo Methods in Bayesian Computation*. New York: Springer-Verlag, 2000. 387p.

CHEN, M. H., IBRAHIM, J. G.; SINHA, D. Bayesian inference for multivariate survival data with a cure fraction. *Journal of Multivariate Analysis*, v.80, n.1, p.101–126, 2002.

COOK, R. D.; WEISBERG, S. *Residuals and Influence in Regression*. New York: Chapman and Hall/CRC Press, 1982. 240p.

COONER, F., BANERJEE, S.; McBEAN, A. M. Modelling geographically referenced survival data with a cure fraction. *Statistical Methods in Medical Research*, v.15, n.4, p.307–324, 2006.

COONER, F., BANERJEE, S., CARLIN, B. P.; SINHA, D. Flexible cure rate modeling under latent activation schemes. *Journal of the American Statistical Association*, v.102, n.478, p.560–572, 2007.

COWLESS, M. K.; CARLIN, B. P. Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, v.91, n.434, p.883–904, 1996.

DEY, D.; BIRMIWAL, L. Robust Bayesian analysis using divergence measures. *Statistics and Probability Letters*, v.20, n.4, p.287–294, 1994.

GEISSER, S.; EDDY, W. F. A predictive approach to model selection. *Journal of the American Statistical Association*, v.74, n.365, p.153–160, 1979.

GELFAND, A. E., DEY, D. K.; CHANG, H. Model determination using predictive distributions with implementation via sampling-based methods. In Bernardo, J. M. (Ed.). *Bayesian Statistics 4*. UK: Oxford University Press, p.147–167, 1992.

GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In Bernardo, J. M., Berger, J. O., Dawid, A. P. and Smith A. F. M. (Eds.). *Bayesian Statistics 4*. Oxford: Clarendon Press, p.169–193, 1992.

IBRAHIM, J. G., CHEN, M.; SINHA, D. *Bayesian Survival Analysis*. New York: Springer-Verlag, 2001. 480p.

KIM, S., CHEN, M.; DEY, D. A new threshold regression model for survival data with a cure fraction. *Lifetime data analysis*, v.17, n.1, p.101–122, 2011.

LOUZADA, F., SUZUKI, A. K.; CANCHO, V. G. The FGM long-term bivariate survival copula model: model, Bayesian estimation, and case influence diagnostics. *Communications in Statistics - Theory and Methods*, v.42, n.4, p.673–691, 2013.

MALLER, R. A.; ZHOU, X. *Survival Analysis with Long-Term Survivors*. New York: Wiley, 1996. 308p.

PENG, F.; DEY, D. Bayesian analysis of outlier problems using divergence measures. *The Canadian Journal of Statistics - La Revue Canadienne de Statistique*, v.23, n.2, p.199–213, 1995.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>, 2017.

RODRIGUES, J., CANCHO, V. G.; CASTRO, M. *Teoria Unificada de Análise de Sobrevivência*. São Pedro: 18^o Sinape, 2008. 94p.

SPIEGELHALTER, D. J.; THOMAS, A.; BEST, N. G. *WinBUGS Manual [version 1.4.1]*. MRC Biostatistics Unit, Cambridge University, UK, 2003.

SUZUKI, A. K., BARRIGA, G. D. C., LOUZADA, F.; CANCHO, V. G. A general long-term aging model with different underlying activation mechanisms: Modeling, Bayesian Estimation and Case Influence Diagnostics. *Communications in Statistics Theory and Methods*, v.46, n.6, p.3080–3098, 2017.

SUZUKI, A. K., CANCHO, V. G.; LOUZADA, F. The Poisson-Inverse-Gaussian regression model with cure rate: a Bayesian approach and its case influence diagnostics. *Statistical Papers*, v.57, n.1, p.133–159, 2016.

TSODIKOV, A. D., IBRAHIM, J. G.; YAKOVLEV, A. Y. Estimating cure rates from survival data: an alternative to two-component mixture models. *Journal of the American Statistical Association*, v.98, n.464, p.1063–1078, 2003.

WEISS, R. An approach to Bayesian sensitivity analysis. *Journal of the Royal Statistical Society Series B*, v.58, n.4, p.739–750, 1996.

YAKOVLEV, A. Y.; TSODIKOV, A. D. *Stochastic Models of Tumor Latency and Their Biostatistical Applications*. World Scientific: Singapore, 1996. 288p.

Recebido em 20.09.2017.

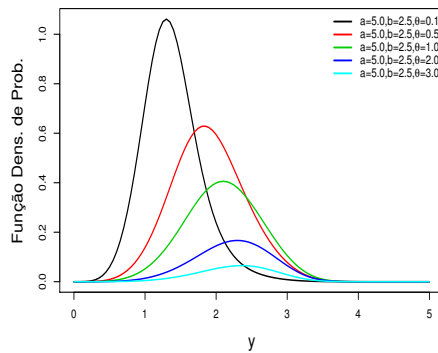
Aprovado após revisão em 11.04.2018.

APÊNDICE A

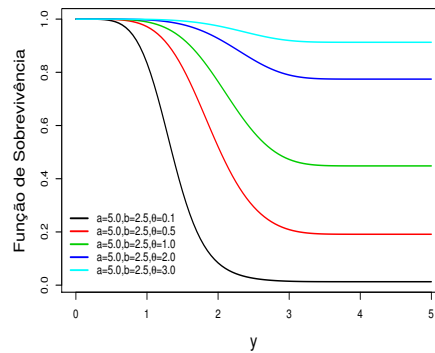
Neste apêndice apresentamos os gráficos da função densidade de probabilidade e da função de sobrevivência do modelo LWcr para alguns valores fixos de a , b e θ .

As Figuras 7 e 8 mostram os gráficos da função densidade de probabilidade e da função de sobrevivência do modelo LWcr sob o esquema de primeira ativação. As Figuras 7(a, b) mostram os gráficos para alguns valores de θ com $a = 5,00$ e $b = 2,50$; As Figuras 7(c, d) mostram os gráficos para alguns valores de b com $a = 2,00$ e $\theta = 0,75$.

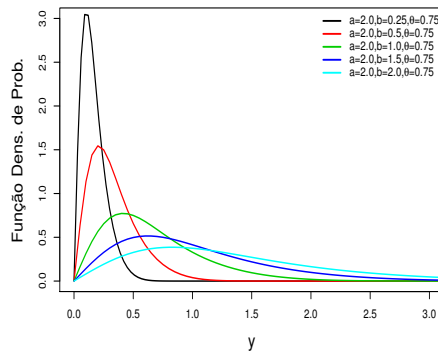
As Figuras 8(a, b) mostram os gráficos da função densidade de probabilidade e da função de sobrevivência do modelo LWcr sob o esquema de primeira ativação para alguns valores de a com $b = 1,50$ e $\theta = 0,50$.



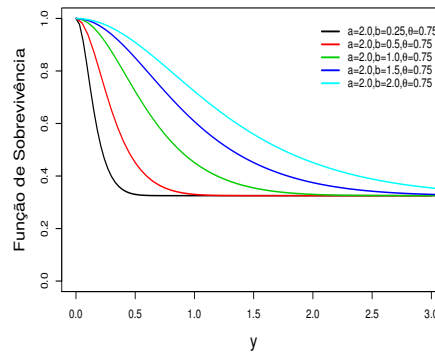
(a) Função densidade.



(b) Função de sobrevivência.



(c) Função densidade.



(d) Função de sobrevivência.

Figura 7 - Função densidade de probabilidade e função de sobrevivência do LWcr sob o esquema de primeira ativação.

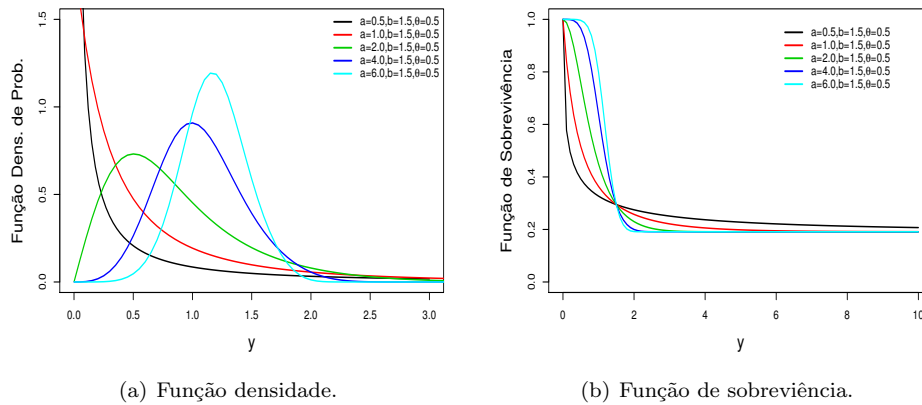


Figura 8 - Função densidade de probabilidade e função de sobrevivência do LWcr sob o esquema de primeira ativação.

As Figuras 9 e 10 mostram os gráficos da função densidade de probabilidade e da função de sobrevivência do modelo LWcr sob o esquema de última ativação. A Figuras 9(a,b) mostram os gráficos para alguns valores de θ com $a = 3,00$ e $b = 2,00$; as Figuras 10(a,b) mostram os gráficos para alguns valores de b com $a = 4,00$ e $\theta = 0,50$; e as Figuras 10(c,d) mostram os gráficos para alguns valores de a com $b = 3,00$ e $\theta = 0,75$.

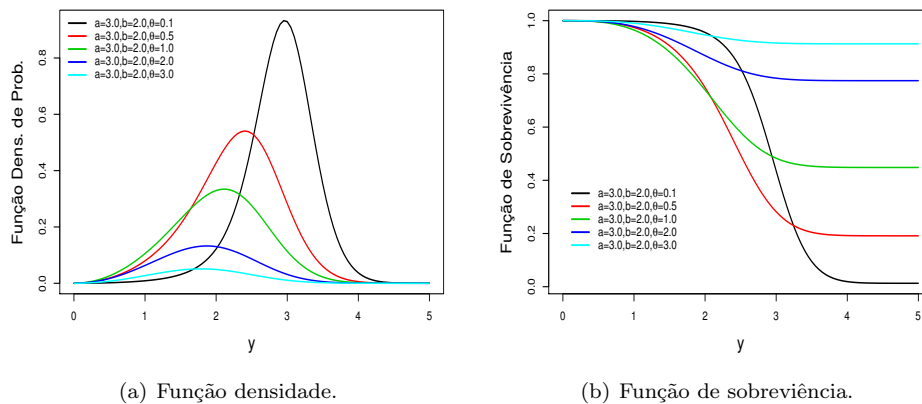
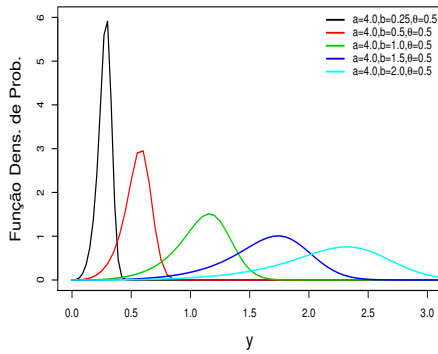
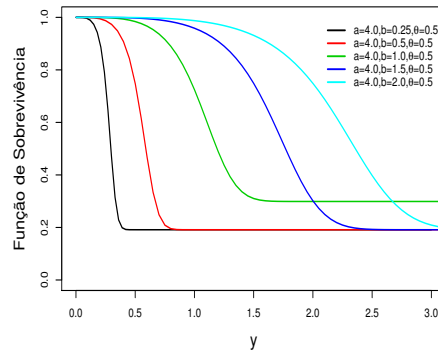


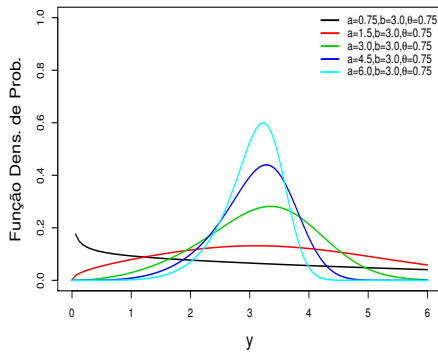
Figura 9 - Função densidade de probabilidade e função de sobrevivência do LWcr sob o esquema de última ativação.



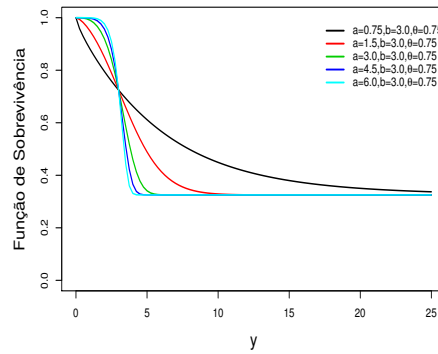
(a) Função densidade.



(b) Função de sobrevivência.



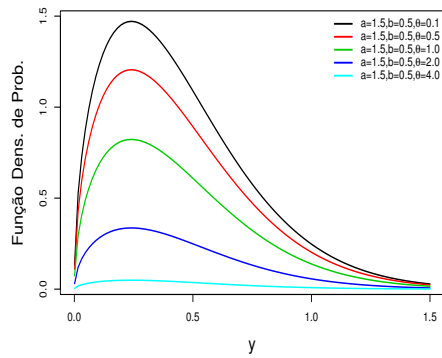
(c) Função densidade.



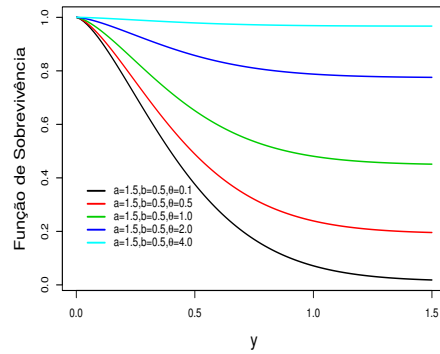
(d) Função de sobrevivência.

Figura 10 - Função densidade de probabilidade e função de sobrevivência do LWCer sob o esquema de última ativação.

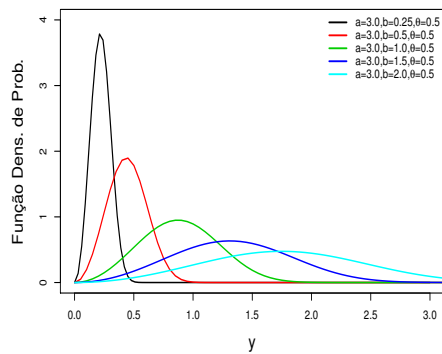
De forma similar, a Figura 11 mostra os gráficos da função densidade de probabilidade e da função de sobrevivência para o esquema de ativação aleatória. As Figuras 11(a, b) mostram os gráficos para alguns valores de θ com $a = 1,50$ e $b = 0,50$; as Figuras 11(c, d) mostram os gráficos para alguns valores de b com $a = 3,00$ e $\theta = 0,50$; e as Figuras 11(e, f) mostram os gráficos para alguns valores de a com $b = 1,50$ e $\theta = 0,75$.



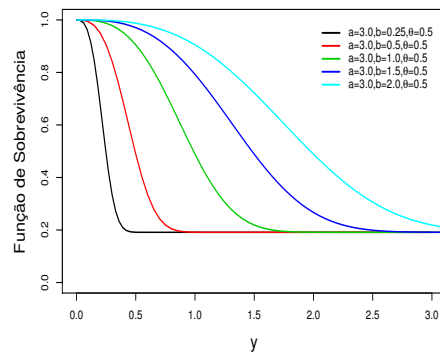
(a) Função densidade.



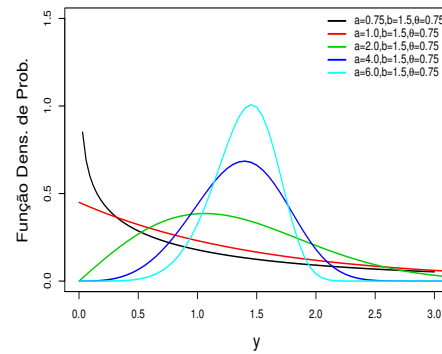
(b) Função de sobrevivência.



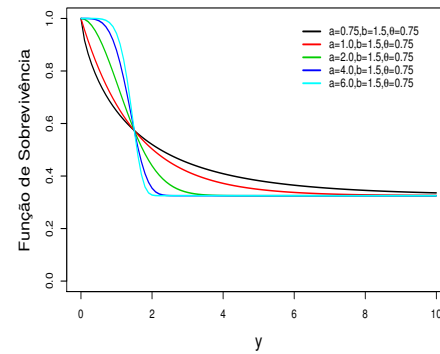
(c) Função densidade.



(d) Função de sobrevivência.



(e) Função densidade.



(f) Função de sobrevivência.

Figura 11 - Função densidade de probabilidade e função de sobrevivência do LWcr sob o esquema de ativação aleatória.