

O PROCEDIMENTO FBST PARA DISCRIMINAÇÃO ENTRE AS DISTRIBUIÇÕES LOGNORMAL E WEIBULL

Cachimo Combo ASSANE¹
Basilio de Brangança PEREIRA¹
Carlos Alberto de Bragança PEREIRA²

- RESUMO: Testes para famílias separadas de hipóteses foram inicialmente considerados por Cox (1961, 1962). Neste artigo, examinamos o teste de significância totalmente Bayesiano, o FBST, para discriminar entre os modelos lognormal e Weibull cujas famílias de distribuições são separadas. Aqui, o problema é abordado num contexto de mistura linear dos modelos candidatos. O procedimento do FBST é utilizado para testar as hipóteses sobre os pesos da mistura com o objetivo de calcular as evidências a favor de cada modelo. Neste trabalho, as funções densidades dos componentes da mistura são reparametrizadas em termos da média μ e da variância σ^2 da população, para que todos os modelos compartilhem dos mesmos parâmetros, uma vez que a comparação entre eles é baseada no mesmo conjunto de observações, isto é, na mesma população. Para avaliar o desempenho do método proposto, alguns resultados numéricos baseados em simulações de pontos amostrais são apresentados. Nessas simulações, os resultados do FBST são comparados com os do teste de Cox. Dois exemplos de aplicação ilustrando os procedimentos para conjunto de dados não censurados também são apresentados.
- PALAVRAS-CHAVE: Hipóteses separadas; modelo de mistura linear; distribuição lognormal; distribuição Weibull; testes de significância; teste FBST; teste de Cox.

¹Universidade Federal do Rio de Janeiro - UFRJ, Faculdade de Medicina, Caixa Postal 68507, CEP: 21941-972, Rio de Janeiro, RJ, Brasil. E-mail: *cachimo.assane@uem.mz*; *basilio@hucff.ufrj.br*

²Universidade de São Paulo - USP, Instituto de Matemática e Estatística, Caixa Postal 66281, CEP: 05508-090, São Paulo, SP, Brasil. E-mail: *cpereira@ime.usp.br*

1 Introdução

A escolha entre modelos estatísticos alternativos é um problema importante em análise estatística, especialmente na modelagem da distribuição do tempo até a ocorrência de uma falha (PEREIRA, 1978; LAWLESS, 2002; KUNDU e MANGLICK, 2004; ARAUJO e PEREIRA, 2007). A teoria de teste de hipóteses de Neyman-Pearson pode ser aplicada para comparar modelos pertencentes a mesma família de distribuições. Alternativamente, procedimentos especiais são necessários quando os modelos candidatos pertencem a famílias separadas, no sentido de que um membro arbitrário de uma família não pode ser obtido como limite dos membros da outra. Problemas desse tipo são denominados problemas de *modelos separados* ou de *hipóteses separadas*. As distribuições lognormal e Weibull são exemplos de modelos pertencentes a famílias separadas (PEREIRA, 1978).

Testes para discriminar modelos separados foram inicialmente considerados por Cox (1961, 1962), que desenvolveu resultados gerais e alguns exemplos. Nesses trabalhos fundamentais de Cox, foram sugeridas abordagens alternativas que serviram como base para o desenvolvimento de diversos procedimentos de testes para modelos separados. Para revisão geral e referências, vide Pereira (1977), Araujo e Pereira (2007) e Pereira e Pereira (2017).

Teste de significância é considerado como procedimento estatístico para medir a consistência dos dados com a hipótese sendo testada, por vezes chamada de hipótese nula (COX, 1977). Na estatística frequentista, o p-valor é usualmente considerado como uma medida de evidência a favor da hipótese nula (BERGER e DELAMPADY, 1987). Porém, seu conceito é estabelecido sobre o espaço amostral (os dados) e não sobre o espaço paramétrico onde a hipótese é formulada (PEREIRA et al., 2008; STERN e PEREIRA, 2014). No contexto Bayesiano, Berger e Delampady (1987) consideram o fator de Bayes e a probabilidade a posteriori da hipótese nula como medidas de evidência alternativas do p-valor clássico. Estas medidas, no entanto, apresentam diversas limitações quando são utilizadas para resolver o problema de seleção de modelos (KAMARY et al., 2014). Particularmente, sabe-se que o fator de Bayes, quando utilizado para testar hipóteses precisas ou discriminar modelos separados, é muito sensível a suposições a priori. Por exemplo, quando a informação a priori sobre os parâmetros dos modelos candidatos é fraca e prioris impróprias são usadas, o fator de Bayes torna-se indefinido (GHOSH, 2006). Para contornar as dificuldades com relação ao uso de prioris impróprias, diversas modificações do fator de Bayes têm sido propostas na literatura (ARAUJO e PEREIRA, 2007; PEREIRA e PEREIRA, 2017).

O Teste de Significância Totalmente Bayesiano (FBST, do inglês *Fully Bayesian Significance Test*) foi apresentado por Pereira e Stern (1999) como alternativa do fator de Bayes, da probabilidade a posteriori e dos p-valores clássicos, para calcular a medida de evidência a favor da hipótese nula precisa. O teste é considerado totalmente Bayesiano por fazer cálculos probabilísticos no espaço paramétrico original, em sua plena dimensão. O e-valor, a medida de evidência utilizada para definir o FBST, representa a probabilidade a posteriori relacionada

aos pontos menos prováveis do espaço paramétrico. Essa medida favorece a hipótese nula quando seu valor é elevado. Uma vantagem do FBST é a de não necessitar premissas adicionais, como probabilidade positiva para hipóteses precisas. Para uma revisão geral do FBST e referências, vide Pereira et al. (2008) e Stern e Pereira (2014).

As distribuições lognormal e Weibull estão entre as mais utilizadas na análise de sobrevivência e de confiabilidade para modelagem do tempo até a ocorrência de uma falha (LAWLESS, 2002). A escolha entre essas duas distribuições é um problema que vem sendo bastante estudado na literatura. Vide, por exemplo, Pereira (1978) Kundu e Manglick (2004) e Araujo e Pereira (2007). A distribuição Weibull é a única que tem riscos proporcionais e uma representação do tempo de falha acelerado. Sua função de risco (taxa de falha) é monótona, isto é, ela é crescente, decrescente ou constante, dependendo do parâmetro de forma. A distribuição lognormal também pode ser usada na modelagem do tempo de falha acelerado. Sua função do risco cresce, atinge um valor máximo e depois decresce para zero à medida que o tempo tende ao infinito (COLLETT, 2003).

O objetivo principal deste artigo é formular o FBST para discriminar entre as distribuições lognormal e Weibull. Conforme sugerido por Cox (1961), aqui, o problema é abordado num contexto de mistura linear dos modelos candidatos, ou seja, os modelos a serem comparados são considerados como componentes de um modelo de mistura finita. O procedimento do FBST é utilizado para testar as hipóteses sobre os pesos da mistura com o objetivo de calcular as evidências a favor de cada modelo.

Lauretto *et al.* (2007) utilizaram o FBST para seleção de modelos separados através da mistura, comparando entre as distribuições Weibull e Gompertz. A novidade na nossa abordagem é que as funções densidades dos componentes da mistura são reparametrizadas em termos da média μ e da variância σ^2 da população, para que os modelos a serem comparados compartilhem dos mesmos parâmetros. O mais comum na análise Bayesiana de um modelo de mistura finita é considerar pares de parâmetros diferentes para cada um desses modelos e assumir prioris independentes para cada par de parâmetros e uma priori $Beta(1, 1)$ (uniforme(0, 1)) para o peso da mistura. No entanto, uma vez que a comparação entre os modelos é baseada no mesmo conjunto de observações, i.e, na mesma população, entendemos que seria inconveniente considerar, nesses modelos, médias e variâncias diferentes e ainda por cima com prioris diferentes.

Com base no estudo de simulação, os resultados do FBST são comparados com os do teste de Cox, tomando como medidas de desempenho as taxas de decisões corretas (aceitação do modelo verdadeiro e rejeição do modelo falso). Dois exemplos de aplicação ilustrando os procedimentos para conjunto de dados não censurados também são apresentados.

Este trabalho está organizado como se segue: A seção 2 revê conceitos básicos do teste de Cox e sintetiza as estatísticas do teste para comparação entre as distribuições lognormal e Weibull. A seção 3 apresenta conceitos básicos do FBST e desenvolve a metodologia sobre procedimento do teste para discriminação entre as

distribuições a partir de um modelo de mistura linear. A seção 4 apresenta alguns resultados numéricos baseados em simulações de pontos amostrais. Na seção 5, dois conjuntos de dados reais não censurados são usados para ilustrar os procedimentos. Considerações finais e propostas para estudos futuros são apresentados na seção 6.

2 O teste de Cox

Seja $y = (y_1, \dots, y_n)$ uma amostra aleatória de n observações condicionalmente (aos parâmetros) independentes e identicamente distribuídas (c.i.i.d.). Suponha que a hipótese nula, H_f , especifica que y provém de uma distribuição F , que pertence a uma família de distribuições de probabilidade $\mathcal{F} = \{f(y|\alpha) : \alpha \in \Omega_\alpha\}$. Suponha ainda que a hipótese alternativa, H_g , especifica que y provém de uma distribuição $G \in \mathcal{G} = \{g(y|\beta) : \beta \in \Omega_\beta\}$. Aqui, α e β são vetores de parâmetros desconhecidos e supõe-se que as famílias das distribuições são separadas, i.e, os espaços de parâmetros Ω_α e Ω_β são tais que $\Omega_\alpha \cap \Omega_\beta \neq \Omega_\alpha$ e $\Omega_\beta \cap \Omega_\alpha \neq \Omega_\beta$. O problema consiste em decidir qual dentre as distribuições, F e G , melhor se ajusta aos dados disponíveis.

O teste assintótico desenvolvido por Cox (1961, 1962) é baseado numa modificação da razão de máxima verossimilhança de Neyman-Pearson. A estatística considerada por Cox para testar H_f contra H_g é

$$T_{fg} = \ell_f(\hat{\alpha}) - \ell_g(\hat{\beta}) - n \left[\text{plim}_{n \rightarrow \infty} \frac{\ell_f(\hat{\alpha}) - \ell_g(\hat{\beta})}{n} \right]_{\alpha=\hat{\alpha}}, \quad (1)$$

onde $\hat{\alpha}$ e $\hat{\beta}$ são respectivamente os estimadores de máxima verossimilhança de α e β , $\ell_f(\hat{\alpha}) = \log f(y|\hat{\alpha})$, $\ell_g(\hat{\beta}) = \log g(y|\hat{\beta})$. O operador plim simboliza convergência em probabilidade, e o subscrito α indica que o limite em probabilidade é calculado sob H_f .

Cox mostrou que, assintoticamente, sob a hipótese alternativa, T_{fg} tem média negativa e que, sob a hipótese nula, T_{fg} tem distribuição normal com média zero e variância

$$V_\alpha(T_{fg}) = V_\alpha \{ \ell_f(\alpha) - \ell_g(\beta_\alpha) \} - C'_\alpha I_\alpha^{-1} C_\alpha, \quad (2)$$

onde β_α é o limite em probabilidade, quando $n \rightarrow \infty$, de $\hat{\beta}$ supondo H_f verdadeira,

$$C_\alpha \equiv n \frac{\partial}{\partial \alpha} \left[\text{plim}_{n \rightarrow \infty} \frac{\ell_f(\hat{\alpha}) - \ell_g(\hat{\beta})}{n} \right] \text{ e } I_\alpha \text{ é a matriz de informação de } \alpha.$$

Quando H_g é a hipótese nula e H_f é a hipótese alternativa, resultados análogos são obtidos para uma estatística T_{gf} . Portanto $T_{fg}^* = T_{fg} \{V_\alpha(T_{fg})\}^{-1/2}$ e $T_{gf}^* = T_{gf} \{V_\beta(T_{gf})\}^{-1/2}$, sob H_f e H_g respectivamente, podem ser considerados aproximadamente como variáveis aleatórias com distribuição normal padrão e testes bilaterais podem ser efetuados de acordo com as recomendações de Cox (1962). Interpretações sobre possíveis resultados quando ambos os testes, T_{fg}^* e T_{gf}^* , são aplicados podem ser encontradas em Pereira e Pereira (2017).

2.1 Testes para as distribuições lognormal e Weibull

Suponhamos que H_L seja a hipótese nula de que a distribuição é lognormal; e que H_W seja a hipótese alternativa de que a distribuição é Weibull, isto é,

$$H_L : f_L(y, \alpha_1, \alpha_2) = \frac{1}{y\sqrt{2\pi\alpha_2}} \exp\left\{-\frac{(\log y - \alpha_1)^2}{2\alpha_2}\right\}, \quad (3)$$

$$H_W : f_W(y, \beta_1, \beta_2) = \left(\frac{\beta_2}{y}\right) \left(\frac{y}{\beta_1}\right)^{\beta_2} \exp\left\{-\left(\frac{y}{\beta_1}\right)^{\beta_2}\right\}.$$

Os limites em probabilidade $\beta_{1\hat{\alpha}}$ e $\beta_{2\hat{\alpha}}$ dos estimadores $\hat{\beta}_1$ e $\hat{\beta}_2$, respectivamente, são soluções de

$$\beta_{1\hat{\alpha}} = \exp\{\hat{\alpha}_1 + \sqrt{\hat{\alpha}_2}/2\}, \quad \beta_{2\hat{\alpha}} = \hat{\alpha}_2^{-1/2}.$$

Então temos

$$T_{LW} = n \left\{ \hat{\beta}_2 \ln \hat{\beta}_1 - \beta_{2\hat{\alpha}} \ln \beta_{1\hat{\alpha}} - \ln \hat{\beta}_2 + \ln \beta_{2\hat{\alpha}} - \hat{\alpha}_1 (\hat{\beta}_2 - \beta_{2\hat{\alpha}}) \right\}, \quad (4)$$

$$V_L(T_{LW}) = 0.2183n.$$

Quando H_L e H_W invertem os papéis, tal que a hipótese nula é H_W e a hipótese alternativa é H_L , teremos

$$\alpha_{1\hat{\beta}} = -0.5772/\hat{\beta}_2 + \ln \hat{\beta}_1; \quad \alpha_{2\hat{\beta}} = 1.6449/\hat{\beta}_2^2, \\ T_{WL} = n \left\{ \hat{\beta}_2 (\hat{\alpha}_1 - \alpha_{1\hat{\beta}}) + \frac{1}{2} \ln \frac{\hat{\alpha}_2}{\alpha_{2\hat{\beta}}} \right\} \quad (5)$$

$$V_W(T_{WL}) = 0.2834n.$$

Portanto, assintoticamente, as estatísticas $T_{LW}^* = T_{LW} \{V_L(T_{LW})\}^{-1/2}$ e $T_{WL}^* = T_{WL} \{V_W(T_{WL})\}^{-1/2}$, sob H_L e H_W , respectivamente, seguem aproximadamente a distribuição normal padrão. Considerando testes bilaterais, fixado um nível de significância γ , a região de rejeição do teste de Cox será definida como $R = \{Y : |T^*| > z_{\gamma/2}\}$, onde $T^* \sim N(0, 1)$.

Interpretações sobre possíveis resultados quando os testes T_{LW}^* e T_{WL}^* são aplicados podem ser encontradas em Pereira e Pereira (2017).

3 O FBST-Teste de Significância Totalmente Bayesiano

O modelo padrão da estatística Bayesiana paramétrica concerne a uma variável aleatória (vetorial), $y \in \mathcal{Y}$, que tem uma distribuição amostral com uma forma funcional especificada, $p(y|\theta)$, indexada por $\theta \in \Theta$. A mesma forma funcional, considerada como função da variável livre θ com argumento fixo y , passa a ser denominada função de verossimilhança do modelo, $l(\theta; y) = p(y|\theta)$. O conjunto de todos os valores possíveis de θ , $\Theta \subset \mathcal{R}^q$, $q \geq 1$, é denominado espaço paramétrico, e o conjunto de todos os valores possíveis de y , $\mathcal{Y} \subset \mathcal{R}^p$, $p \geq 1$, é denominado espaço amostral.

No contexto Bayesiano, o parâmetro θ é considerado como uma variável aleatória latente (não observável). A distribuição conjunta de y e θ é definida como

$$p(\theta, y) = l(\theta; y)p(\theta) = p(\theta|y)p(y). \quad (6)$$

A distribuição de probabilidade a priori, $p_0(\theta)$, representa a informação inicial disponível sobre o parâmetro. Se em seguida observarmos um único evento, y , a distribuição a posteriori para θ , representando a informação disponível sobre o parâmetro após a observação, é dada por

$$p_1(\theta) \propto l(\theta; y)p_0(\theta).$$

Assumindo que as observações são c.i.i.d., a distribuição a posteriori após n observações, $y = (y_1, \dots, y_n)$, será

$$p_n(\theta) \propto l(\theta; y_n)p_{n-1}(\theta) \propto \left[\prod_{i=1}^n l(\theta; y_i) \right] p_0(\theta). \quad (7)$$

3.1 Definições

O FBST, *Fully Bayesian Significance Teste*, foi apresentado por Pereira e Stern (1999) como um teste intuitivo e teoricamente coerente, com caracterização geométrica, e que pode ser implementado usando técnicas modernas de otimização e integração numérica.

Uma hipótese estatística, H , (por vezes chamada de hipótese nula) afirma que o parâmetro θ do modelo estatístico encontra-se dentro do conjunto da hipótese, Θ_H , ou seja,

$$H : \theta \in \Theta_H,$$

onde

$$\Theta_H = \{\theta \in \Theta | g(\theta) \leq 0 \wedge h(\theta) = 0\}.$$

A bem da simplicidade, de agora em diante, usaremos uma notação relaxada, escrevendo H no lugar de Θ_H . Estamos particularmente interessados em hipóteses *precisas*, i.e., aquelas nas quais a dimensão de H é menor do que a dimensão de Θ .

Denotando por \sup_H o supremo de H , definimos o conjunto tangente, T , e a medida de evidência Bayesiana da seguinte forma:

$$p^* = \sup_H p_n(\theta) \text{ e } T = \{\theta : p_n(\theta) > p^*\}. \quad (8)$$

A medida de evidência contra H é a probabilidade a posteriori de T ,

$$\bar{ev} = \Pr(\theta \in T|y) = \int_T p_n(\theta) d\theta. \quad (9)$$

Por conseguinte, $ev = 1 - \bar{ev}$ é a evidência a favor de H . Intuitivamente, um grande volume de T é indício de que a região sobre a hipótese (incluindo θ^*) possui baixa densidade de probabilidade, i.e., θ está entre os pontos menos prováveis a posteriori, indicando “forte” evidência contra H .

É importante notar que $ev(H)$ não é uma evidência contra a hipótese alternativa $\bar{H} = \Theta \setminus H$ (que não é precisa), assim como $\bar{ev}(H)$ não é evidência a favor de \bar{H} embora seja contra H .

Considere a densidade acumulada da evidência contra uma hipótese, $\bar{V}(c) = \Pr(\bar{ev} \leq c)$. Dado θ^0 , o valor verdadeiro do parâmetro, sob condições apropriadas de regularidade, para tamanhos de amostras crescentes, $n \rightarrow \infty$, pode-se afirmar que:

- Se H é falsa, $\theta^0 \notin H$, então $\bar{ev}(H)$ converge (em probabilidade) para 1, i.e., $\bar{V}(0 < c < 1) \rightarrow 0$.
- Se H é verdadeira, $\theta^0 \in H$, então $\bar{V}(c)$, o nível de confiança, converge (em distribuição) para função

$$Q(t, h, c) = F_{t-h}[F_t^{-1}(c)], \quad (10)$$

onde $t = \dim(\Theta)$, $h = \dim(H)$, $F_g(x)$ é a densidade acumulada da distribuição qui-quadrada com g graus de liberdade e F^{-1} é o quantil da qui-quadrado.

Esta propriedade sugere um critério assintótico para rejeição de H com um nível de confiança $1 - \gamma$, adotando-se

$$c = F_t[F_{t-h}^{-1}(1 - \gamma)], \quad (11)$$

ou seja, escolhendo-se c tal que $Q(t, h, c) = 1 - \gamma$. Portanto, o procedimento do FBST rejeita H se $\bar{ev}(H) > c$.

Diniz *et al.* (2012) mostram que, assintoticamente, existe uma relação entre $\bar{ev}(H)$ e o p-valor baseado no teste da razão de verossimilhança. Assim, assumindo a propriedade de normalidade assintótica, com base na equação (11), tem-se $\bar{ev}(H) \approx F_t[F_{t-h}^{-1}(1 - \text{p-valor})]$. Invertendo a equação, teremos

$$\text{p-valor} = 1 - F_{t-h}[F_t^{-1}(1 - ev(H))]. \quad (12)$$

3.2 FBST para mistura das distribuições lognormal e Weibull

O procedimento de teste para famílias separadas de hipóteses baseado na mistura de modelos foi sugerido por Cox(1961). Neste trabalho adotamos a mistura linear (QUANDT, 1974) para formular o FBST para discriminar entre as distribuições lognormal e Weibull.

Consideremos os modelos especificados pelas hipóteses H_L e H_W (3). A função densidade de probabilidade de y , que provém da mistura das duas distribuições é representada como

$$f(y|\theta) = pf_L(y, \alpha_1, \alpha_2) + (1 - p)f_W(y, \beta_1, \beta_2), \quad (13)$$

onde o vetor $\theta = (p, \alpha_1, \alpha_2, \beta_1, \beta_2)$ contém todos os parâmetros do modelo, sendo p a proporção ou peso da mistura.

Neste trabalho, as densidades de probabilidade das distribuições lognormal e Weibull são reparametrizadas em termos da média, μ , e da variância, σ^2 , da população. A principal razão para esta reparametrização é que, uma vez que a comparação entre as distribuições é baseada no mesmo conjunto de observações, isto é, na mesma população, entendemos que não faria muito sentido considerar, nesses modelos, médias e variâncias diferentes e ainda por cima com prioris diferentes.

Portanto, para distribuição lognormal, temos

$$\begin{cases} \mu = E(y, \alpha_1, \alpha_2) = e^{\alpha_1 + \alpha_2/2} \\ \sigma^2 = Var(y, \alpha_1, \alpha_2) = (e^{\alpha_2} - 1)e^{2\alpha_1 + \alpha_2} \end{cases} \Rightarrow \begin{cases} \alpha_1 = \log \frac{\mu^2}{\sqrt{\mu^2 + \sigma^2}} \\ \alpha_2 = \sqrt{\log \frac{\mu^2 + \sigma^2}{\mu^2}}, \end{cases} \quad (14)$$

e para distribuição Weibull,

$$\begin{cases} \mu = E(y, \beta_1, \beta_2) = \beta_1 \Gamma(1 + 1/\beta_2) \\ \sigma^2 = Var(y, \beta_1, \beta_2) = \beta_1^2 \Gamma(1 + 2/\beta_2) - \beta_1^2 \Gamma^2(1 + 1/\beta_2) \end{cases} \Rightarrow \begin{cases} \beta_1 = \frac{\mu}{\Gamma(1 + 1/\beta_2)} \\ 2 \log \Gamma(1 + 1/\beta_2) - \log \Gamma(1 + 2/\beta_2) + \log \frac{\mu^2 + \sigma^2}{\mu^2} = 0. \end{cases} \quad (15)$$

O valor de β_2 é calculado por um método numérico. Neste trabalho, recorreremos ao método de Newton-Raphson, via função `nleqslv` do pacote com mesmo nome, do *software R* (R CORE TEAM, 2016).

Observe que, inicialmente (13), θ continha cinco parâmetros a serem estimados. Após a reparametrização, a quantidade dos parâmetros diminui para três, ou seja, $\theta = (\mu, \sigma^2, p)$.

A função de verossimilhança condicionada aos dados observados, considerando a reparametrização, é dada por:

$$l(\theta; y_n) = \prod_{i=1}^n [pf_L(y_i, \mu, \sigma) + (1-p)f_W(y_i, \mu, \sigma)]. \quad (16)$$

Assumindo independência a priori entre os parâmetros, a função densidade conjunta a priori do vetor $\theta = (\mu, \sigma^2, p)$ é definida como

$$\pi(\theta) = \pi(p)\pi(\mu)\pi(\sigma^2). \quad (17)$$

Por conseguinte, temos a distribuição a posteriori conjunta dada por

$$f(\theta|y) = l(\theta; y_n)\pi(\theta). \quad (18)$$

Neste trabalho, a incerteza a prior para os parâmetros μ e σ^2 é modelada pela distribuição gama de média 1 e variância 100, ou seja, $\mu, \sigma^2 \sim \text{gama}(0.01, 100)$ (PEREIRA e PEREIRA, 2017). Para o peso da mistura, consideramos a distribuição a priori uniforme(0, 1) (*Beta*(1, 1)).

Para calcular as evidências a favor das hipóteses H_L e H_W , formuladas em (3), testes sobre o peso da mistura são efetuados. Assim, a hipótese de que y tem a função densidade f_L é equivalente à hipótese

$$H_L : p = 1,$$

e a hipótese de que y tem função densidade f_W equivale à

$$H_W : p = 0.$$

O procedimento do FBST é utilizado para testar H_L e H_W de acordo com as expressões (8) and (9). No passo de otimização, utilizamos o método de gradiente conjugado (Fletcher e Reeves, 1964), via função `optim` do pacote `stats` do software R, versão 3.2.2.

Na etapa de integração, adotamos o algoritmo Metropolis adaptativo (AM, do inglês *Adaptive Metropolis*), proposto por Haario *et al.* (2001), para simular as amostras da distribuição a posteriori dos parâmetros de interesse. Foi simulada, para cada parâmetro, uma amostra com 80 mil valores da distribuição a posteriori. Após descartarmos cinco mil valores iniciais (*burn-in*) e tomando um salto de tamanho cinquenta, obtemos uma amostra final de 1500 observações que são utilizadas para calcular as estimativas.

Os valores da posteriori conjunta são obtidos a partir dos valores dos parâmetros das distribuições lognormal e Weibull que são recalculados, com base nas expressões (14) e (15), usando observações de μ e σ^2 mantidas no método Markov Chain Monte Carlo (MCMC).

Aqui, a implementação do modelo Bayesiano foi desenvolvida usando o pacote `LaplacesDemon`, do software R (R CORE TEAM, 2016). O `LaplacesDemon` é um pacote estatístico gratuito, de código-fonte aberto e livremente distribuído que proporciona um ambiente completo para simulação em inferência Bayesiana (STATISTICAT, LCC, 2016).

4 Simulações

Nesta seção apresentamos alguns resultados numéricos, baseados na simulação de pontos amostrais, para avaliar o desempenho do FBST na escolha entre as distribuições lognormal e Weibull. Os resultados dos testes foram obtidos usando o software R (R CORE TEAM, 2016). Todos os programas computacionais foram rodados em um Intel(R) Core(TM) i7-5500U CPU@ 2.40GHz.

4.1 Cenários para simulação de pontos amostrais

Para cada uma das distribuições, simulamos 1000 amostras de tamanho $n = 20, 40, 60, 80, 100, 150, 200$ e testamos as hipóteses H_L e H_W através do FBST e do teste de Cox. O objetivo nesses testes é verificar as taxas de decisões corretas (aceitação do modelo verdadeiro e rejeição do modelo falso).

Aqui, as amostras foram geradas usando diferentes valores dos parâmetros para as distribuições lognormal e Weibull, obtidos com base nos cenários apresentados na Tabela 1.

Tabela 1 - Cenários para geração das amostras

Cenário	Média (μ)	Variância (σ^2)
1	10	9^2
2	10	12^2
3	150	15^2
4	200	12^2

Para aceitação/rejeição das hipóteses, calculamos o nível crítico c de acordo com o critério assintótico, apresentado na seção 3.1, com um nível de significância de $\gamma = 5\%$. Uma vez que o modelo de mistura completo (reparametrizado) e o modelo restrito (uma componente) possuem, respectivamente, 3 e 2 graus de liberdade, temos $c = F_3[F_1^{-1}(0.95)] = 0.72$. Portanto, rejeitamos H se $\bar{e}v > 0.72$, ou equivalentemente, se $ev < 0.28$.

No teste de Cox, adotando $\gamma = 5\%$, define-se a região de rejeição como $R = \{y : |T^*| > 1.96\}$, onde $T^* \sim N(0, 1)$. Os estimadores de máxima verossimilhança são calculados via função `fitdistr`, do pacote `MASS`, enquanto que as estimativas dos limites em probabilidade, quando necessárias, são obtidos com recurso ao método de Newton-Raphson, via função `nleqslv` do pacote com mesmo nome, do *software* R.

É importante observar que no teste de Cox, quando a hipótese nula é H_L , consequentemente, H_W representa a hipótese alternativa e vice-versa. Já no FBST, a hipótese alternativa é a mistura entre as duas densidades ($0 < p < 1$). Adicionalmente, a rejeição da hipótese nula tem um significado distinto em cada teste. No teste de Cox, por exemplo, se T_{LW}^* é significativamente negativo, implica a rejeição de H_L na direção (a favor) da hipótese nula H_W . Se T_{LW}^* é significativamente positivo, implica rejeitar H_L na direção oposta de H_W . Enquanto que no FBST, a evidência que leva a rejeição da hipótese nula, digamos H_L , de

maneira alguma favorece a hipótese alternativa (que não é precisa) e nem a hipótese H_W (que não é a alternativa).

4.2 Resultados da simulação

As Tabelas 2, 3, 4 e 5 mostram os resultados dos testes para dados gerados a partir das distribuições lognormal e Weibull, considerando valores dos parâmetros definidos com base nos diferentes cenários.

Observe que, em todos os cenários considerados, os resultados dos testes apresentam um comportamento similar. Isso era esperado, uma vez que a medida de evidência do FBST é invariante com respeito a parametrizações da hipótese nula e com respeito a transformações regulares do sistema de coordenadas (MADRUGA et al., 2003). Adicionalmente, devido à invariância da razão de verossimilhança, a distribuição da estatísticas do teste de Cox é independente dos valores dos parâmetros α_1 e α_2 da lognormal e, também, dos parâmetros β_1 e β_2 da Weibull (PEREIRA, 1978).

Os dois procedimentos apresentam altas taxas de aceitação do modelo verdadeiro para todos os tamanhos das amostra. Observa-se que, embora as taxas de aceitação calculadas através do teste de Cox sejam, em geral, maiores do que as do FBST, ambos os testes alcançam níveis de significância muito próximos de 5%, que corresponde ao nível de significância estabelecido pelo critério assintótico. Isso é o que se espera numa aplicação específica (PEREIRA, 1978).

Quanto à rejeição do modelo falso, nota-se que, à medida que o tamanho da amostra cresce, a taxa converge para 1. As taxas de rejeição fornecidas pelo FBST são, em geral, maiores do que as do teste de Cox, principalmente quando n tende a ser pequeno. Isso significa que o FBST apresenta maior poder discriminatório em comparação com o teste de Cox. Note que o teste de Cox é um procedimento formulado justamente para alcançar um alto poder contra a hipótese nula falsa (COX, 1961).

Tabela 2 - Taxa de aceitação da hipótese nula verdadeira, H_L . Dados gerados da lognormal, considerando diferentes cenários

n	Cenário 1		Cenário 2		Cenário 3		Cenário 4	
	FBST	Cox	FBST	Cox	FBST	Cox	FBST	Cox
20	0.970	0.990	0.940	0.984	0.992	0.994	0.984	0.992
40	0.964	0.990	0.918	0.978	0.978	0.988	0.948	0.998
60	0.948	0.966	0.932	0.968	1.000	1.000	0.996	0.996
80	0.976	0.974	0.950	0.980	0.994	0.990	0.982	0.984
100	0.980	0.980	0.940	0.974	1.000	1.000	0.940	1.000
150	0.966	0.928	0.934	0.962	0.952	0.992	0.976	0.984
200	0.990	0.964	0.956	0.972	0.930	0.978	0.990	0.966

Tabela 3 - Taxa de rejeição da hipótese nula falsa, H_W . Dados gerados da lognormal, considerando diferentes cenários

n	Cenário 1		Cenário 2		Cenário 3		Cenário 4	
	FBST	Cox	FBST	Cox	FBST	Cox	FBST	Cox
20	0.196	0.116	0.214	0.114	0.472	0.162	0.316	0.106
40	0.478	0.386	0.474	0.390	0.662	0.398	0.608	0.384
60	0.730	0.680	0.698	0.672	0.800	0.676	0.704	0.670
80	0.864	0.864	0.848	0.868	0.840	0.794	0.930	0.822
100	0.954	0.976	0.906	0.946	0.878	0.878	0.998	0.938
150	0.972	1.000	0.976	0.996	1.000	1.000	1.000	1.000
200	1.000	1.000	0.992	1.000	1.000	1.000	1.000	1.000

Tabela 4 - Taxa de aceitação da hipótese nula verdadeira, H_W . Dados gerados da Weibull, considerando diferentes cenários

n	Cenário 1		Cenário 2		Cenário 3		Cenário 4	
	FBST	Cox	FBST	Cox	FBST	Cox	FBST	Cox
20	0.986	0.992	0.982	0.984	0.970	0.988	0.982	1.000
40	0.968	0.974	0.978	0.986	1.000	1.000	0.964	0.998
60	0.992	0.988	0.970	0.970	0.976	0.958	1.000	0.966
80	0.980	0.964	0.980	0.970	0.966	0.966	0.944	0.966
100	0.992	0.974	0.974	0.974	0.944	0.944	1.000	1.000
150	0.988	0.968	0.980	0.946	0.950	0.918	0.998	0.942
200	1.000	0.998	0.980	0.952	0.998	0.976	0.938	0.904

Tabela 5 - Taxa de rejeição da hipótese nula falsa, H_L . Dados gerados da Weibull, considerando diferentes cenários

n	Cenário 1		Cenário 2		Cenário 3		Cenário 4	
	FBST	Cox	FBST	Cox	FBST	Cox	FBST	Cox
20	0.434	0.280	0.458	0.240	0.360	0.272	0.392	0.300
40	0.736	0.554	0.826	0.544	0.712	0.498	0.686	0.594
60	0.890	0.798	0.896	0.726	0.856	0.808	0.802	0.798
80	0.924	0.858	0.974	0.888	0.886	0.844	0.888	0.884
100	0.976	0.960	0.988	0.930	0.944	0.942	1.000	1.000
150	0.994	0.998	0.998	0.992	0.982	0.982	1.000	1.000
200	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

5 Exemplos de aplicação

Nesta seção analisamos dois conjuntos de dados não censurados e utilizamos o FBST e o teste de Cox para discriminar entre as distribuições lognormal e Weibull. O objetivo é decidir qual dentre as duas distribuições se ajusta melhor aos dados.

Exemplo 1. Lawless (2002, p.99) apresenta dados obtidos num teste de resistência para um conjunto de 23 rolamentos rígidos de esferas. Cada observação representa o número em milhões de revoluções (rotações) que um rolamento pode suportar antes de ocorrer a falha:

17.88	28.92	33.00	41.52	42.12	45.60	48.80	51.84	51.96	54.12
55.56	67.80	68.44	68.64	68.88	84.12	93.12	98.64	105.12	105.84
127.92	128.04	173.40							

A Tabela 6 apresenta as estimativas dos parâmetros do modelo da mistura ajustado aos dados de rolamentos pontuais. Aqui, D.P., 2.5% e 97.5% denotam o desvio padrão e percentis da distribuição a posteriori dos parâmetros, respectivamente. Na Tabela 7 são apresentados os resultados de teste de hipótese para o peso da mistura: e-valor e p-valor são, respectivamente, as medidas de evidência a favor da hipótese nula calculadas através do FBST e do teste de Cox. O p-valor com asterisco foi calculado segundo Diniz et al. (2012), com base na equação 12.

Tabela 6 - Sumário da distribuição a posteriori dos parâmetros da mistura

Parâmetro	Média	D.P.	2.5%	Mediana	97.5%
p	0.565	0.276	0.047	0.590	0.981
μ	72.777	8.687	58.038	71.953	92.182
σ^2	1725.010	980.355	731.408	1466.264	4244.149

Para a hipótese $H_W : p = 0$, temos e-valor = 0.604 e $T_{WL}^* = -1.0283$, que corresponde ao p-valor de 0.304. Para $H_L : p = 1$, e-valor = 0.914 e $T_{LW}^* = -0.5900$, que corresponde ao p-valor de 0.555. Os resultados indicam que ambas as distribuições fornecem um bom ajuste para os dados, embora a evidência a favor do modelo lognormal seja mais forte em relação ao modelo Weibull. Esta conclusão está de acordo com Kundu e Manglick (2004), que utilizaram o mesmo conjunto de dados e compararam os modelos através da razão de máxima verossimilhança.

Ao testarmos a hipótese $H : p = 0.5$, obtemos um e-valor de 0.932 e p-valor correspondente de 0.507, indicando que talvez o modelo de mistura seja mais adequado para a modelagem desses dados.

A Figura 1 apresenta as curvas de densidades acumuladas empírica e ajustadas pelos modelos lognormal, Weibull e de mistura. Aparentemente, a função acumulada estimada pelo modelo lognormal ajusta-se adequadamente aos dados.

Tabela 7 - Teste de hipótese para o peso da mistura

Hipótese	e-valor	p-valor
$p = 0$	0.604	0.304
$p = 1$	0.914	0.555
$p = 0.5$	0.932	0.507*

*p-valor calculado segundo Diniz et al. (2012)

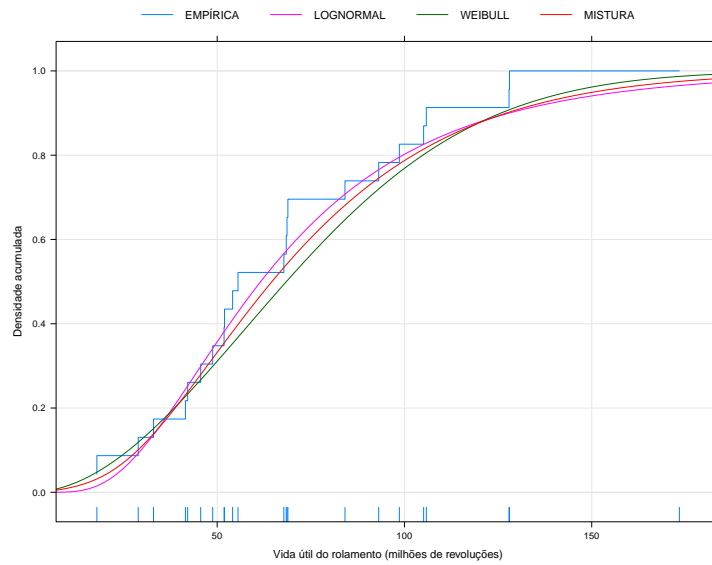


Figura 1 - Funções acumuladas empírica e ajustadas pelos modelos lognormal, Weibull e de mistura.

Exemplo 2. Quesenberry e Kent (1982) apresentam dados referentes a tensão de fios de poliéster/viscose para estudar o problema de ruptura durante a tecelagem. O experimento consiste de 100 amostras de fios submetidos a 80 ciclos por minutos ao nível de 2.3% de tensão. O percentual de tensão é com respeito a resistência de falha do fio. Os dados são os ciclos que os fios suportam antes de ocorrer a falha:

86	146	251	653	98	249	400	292	131	169	175	176	76	264
15	364	195	262	88	264	157	220	42	321	180	198	38	20
61	121	282	224	149	180	325	250	196	90	229	166	38	337
65	151	341	40	40	135	597	246	211	180	93	315	353	571
124	279	81	186	497	182	423	185	229	400	338	290	398	71
246	185	188	568	55	55	61	244	20	284	393	396	203	829
239	236	286	194	277	143	198	264	105	203	124	137	135	350
193	188												

Neste exemplo, os testes indicam uma forte evidência contra o modelo lognormal, uma vez que a hipótese $H_W : p = 0$ não é rejeitada. Além disso, ambos os testes rejeitam as hipóteses $H_L : p = 1$ e $H : p = 0.5$ (vide Tabela 9). As estatísticas do teste de Cox para esse conjunto de dados são: $T_{LW}^* = -3.048$ e $T_{WL}^* = -0.548$. Em Quesenberry e Kent (1982), a Weibull também é preferida em relação a lognormal. Eles usam um método de seleção que exige um conjunto de dados não censurados, e que consiste, essencialmente, em comparar estatísticas que são invariantes sob transformação na escala dos dados e selecionam aquela com mais alto valor. Araujo e Pereira (2007) calcularam os fatores de Bayes intrínseco e fracionário para comparar as mesmas distribuições usando este conjunto de dados. Eles também encontraram uma decisiva evidência contra a distribuição lognormal.

Tabela 8 - Sumário da distribuição a posteriori dos parâmetros da mistura

Parâmetro	Média	D.P.	2.5%	Mediana	97.5%
p	0.244	0.171	0.010	0.217	0.633
μ	220.785	14.250	193.451	220.289	249.620
σ^2	20324.692	3895.387	1422.089	19735.266	29489.616

Tabela 9 - Teste de hipótese para o peso da mistura

Hipótese	e-valor	p-valor
$p = 0$	0.874	0.583
$p = 1$	0.000	0.002
$p = 0.5$	0.232	0.038*

*p-valor calculado segundo Diniz et al. (2012)

Com base nas curvas densidades de acumuladas, representadas na Figura 2, parece claro que o modelo Weibull produz uma boa estimativa da função acumulada.

Tabela 10 - Sumário da distribuição a posteriori dos parâmetros do modelo Weibull

Parâmetro	Média	D.P.	2.5%	Mediana	97.5%
μ	220.232	13.667	194.226	219.959	248.247
σ^2	19718.773	3149.243	14629.004	19402.684	27032.125

Tabela 11 - Sumário da distribuição a posteriori dos parâmetros do modelo lognormal

Parâmetro	Média	D.P.	2.5%	Mediana	97.5%
μ	219.537	16.380	189.172	218.725	253.107
σ^2	34655.941	8895.685	20893.080	33567.197	54674.254

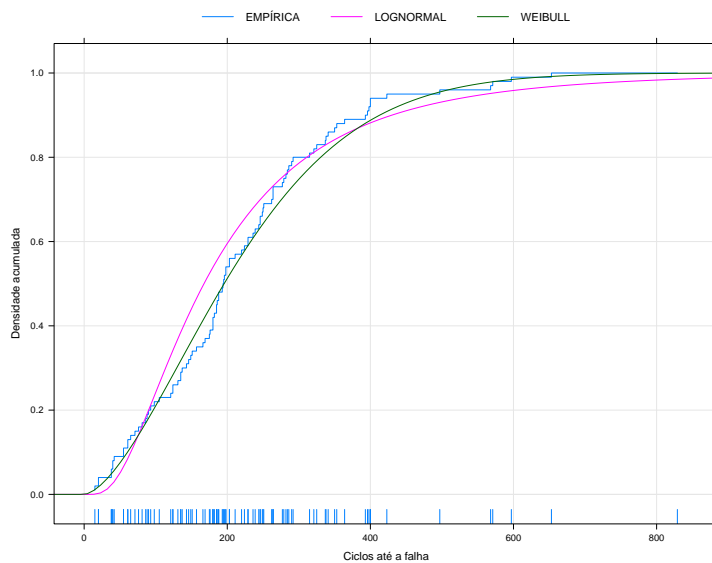


Figura 2 - Funções acumuladas empírica e ajustadas pelos modelos lognormal, Weibull e de mistura.

6 Conclusões

Neste trabalho consideramos o FBST como procedimento de teste para escolha entre as distribuições lognormal e Weibull cujas famílias são separadas. O teste foi formulado no contexto da mistura linear dessas distribuições, conforme sugerido por Cox (1961). O cálculo da evidência a favor de cada modelo é baseado nos testes de hipóteses sobre o peso da mistura. Neste trabalho, as funções densidades das distribuições são reparametrizadas em termos da média μ e da variância σ^2 da população, para que todos os modelos tenham os mesmos parâmetros, já que a comparação entre eles é baseada no mesmo conjunto de dados, i.e, na mesma população.

O padrão geral dos resultados das simulações indica um comportamento do FBST similar ao do teste de Cox. Não obstante, o FBST apresenta maior poder discriminatório em relação ao teste de Cox, principalmente para amostras de tamanho pequeno. Nos dois exemplos de aplicação apresentados neste trabalho, os resultados do FBST levaram às mesmas conclusões anunciadas em outros trabalhos, que abordaram o mesmo problema utilizando outros procedimentos de seleção. Portanto, a abordagem proposta neste trabalho pode ser utilizada para discriminação entre as distribuições lognormal e Weibull mesmo para amostra de tamanho pequeno.

Como trabalho futuro, sugere-se a utilização do procedimento proposto neste artigo para a análise do modelo de mistura lognormal-gama-Weibull aplicado tanto aos dados simulados quanto a uma base dados reais sujeitos à censura.

7 Agradecimentos

Ao CNPq pelo apoio financeiro e aos dois revisores e editores pelos comentários e sugestões que contribuíram para melhoria do nosso trabalho

ASSANE, C. C.; PEREIRA, B. B.; PEREIRA, C. A. B. The FBST procedure for discriminating between the Lognormal and the Weibull distributions. *Rev. Bras. Biom.*, Lavras, v.36, n.1, p.188-206, 2018.

- **ABSTRACT:** *Tests of separate families of hypotheses were initially considered by Cox (1961,1962). In this paper we examine the Fully Bayesian Significance Test, FBST, for discriminating between the lognormal and Weibull models whose families of distributions are separate. Here, we analyze this problem in the context of linear mixture models. The FBST procedure is used for testing the hypotheses on the mixture weights in order to calculate the evidence measure in favor of each model. In this work, the density functions of the mixture components are reparametrized in terms of the common parameters, the mean μ and the variance σ^2 of the population, since the comparison between the models is based on the same dataset, i.e, on the same population. In order to evaluate the performance of the proposed method, some numerical results based on simulations of sample points are given. In these simulations, the results of FBST are compared with those of the Cox test. Two application examples illustrating the procedures for uncensored data set are also presented.*

- **KEYWORDS:** *Separate hypotheses; linear mixture model; lognormal distribution; Weibull distribution; significance tests; FBST test; Cox test.*

Referências

- ARAUJO, M. I.; PEREIRA, B. B. A Comparison of Bayes Factors for Separated Models: Some Simulation Results. *Communications in Statistics–Simulation and Computation*, v.36, n.2, p.297–309, 2007.
- BERGER, J.; DELAMPADY, M. Testing precise hypothesis. *Statistical Science*, v.2, n.3, p. 317–352, 1987.
- COLLETT, D. *Modelling Survival Data in Medical Research*. 2.ed. London: Chapman Hall, 2003.
- COX, D. R. Tests of separate families of hypotheses. *Proceedings 4th Berkeley Symposium in Mathematical Statistics and Probability*, v.1, p.105–123, 1961.
- COX, D. R. Further results on test of separate families of hypotheses. *Journal of the Royal Statistical Society*, v.24, n.2, p.406–424, 1962.
- COX, D. R. The role of significance tests. *Scandinavian Journal of Statistics*, v.4, n.2, p.49–70, 1977.
- DINIZ, M.; PEREIRA, C. A. B.; POLPO, A.; STERN, J. M.; WECHSLER, S. Relationship between Bayesian and Frequentist significance indices. *International Journal for Uncertainty Quantification*, v.2, n.2, p.161–172, 2012.
- FLETCHER, R.; REEVES, C. M. Function minimization by conjugate gradients. *Computer Journal*, v.7, p.148–154, 1964.
- GHOSH, J. K.; DELAMPADY, M.; SAMANTA, T. *An Introduction to bayesian analysis: Theory and methods*. New York: Springer, 2006.
- GILKS, W. R.; BEST, N. G.; TAN, K. K. C. Adaptive rejection Metropolis sampling within Gibbs sampling. *Journal of the Royal Statistical Society*, v.44, n.4, p.455–472, 1995.
- HAARIO, H.; SAKSMAN, E.; TAMMINEN, J. An adaptive Metropolis algorithm. *Bernoulli*, v.7, p.223–242, 2001.
- KAMARY, K.; MENGERSEN, K.; ROBERT, C.P.; ROUSSEAU, J. Testing hypotheses via a mixture estimation model. *arXiv:1412.2044v2*, 2014.
- KUNDU, D.; MANGLICK, A. Discriminating between the Weibull and LogNormal Distributions. *Naval Research Logistics*, v.51, n.6, p.893–905, 2004.
- LAWLESS, J. F. *Statistical Models and Methods for Lifetime Data*. 2.ed. New Jersey: John Wiley & Sons, 2002.
- LAURETTO, M. S.; FARIA JR, S. R.; PEREIRA, B. B.; PEREIRA, C. A. B.; STERN, J. M. The problem of separate hypotheses via mixture models. In: KNUTH, A.; CATICHA, J. L.; CENTER JR, A.; GRIFFIN,

- A. and RODRIGUES, C. C. (Ed.). *Bayesian and maximum entropy methods in science and Engineering*. American Institute of Physics Proceedings, 2007. v.954, p.268–275.
- MADRUGA, M.; PEREIRA, C. A. B.; STERN, J. M. Bayesian evidence test for precise hypotheses. *Journal of Statistical Planning and Inference*, v.117, n.2, p.185–198, 2003.
- PEREIRA, B. B. Discriminating among separate models: A bibliography. *International Statistical Review*, v.45, n.2, p.163–172, 1977.
- PEREIRA, B. B. Empirical comparisons of some tests of separate families of hypotheses. *Metrika*, v.25, n.1, p. 219–234, 1978.
- PEREIRA, B. B.; PEREIRA, C. A. B. *Model choice in nonnested families*, 1.ed. Berlin: Springer, 2017.
- PEREIRA, C. A. B.; STERN, J. Evidence and credibility: Full Bayesian significance test for precise hypotheses. *Entropy*, v.1, n.4, p.69–80, 1999.
- PEREIRA, C. A. B.; STERN, J.; WECHSLER, S. Can a significance test be genuinely Bayesian?. *Bayesian Analysis*, v.3, n.1, p.79–100, 2008.
- QUANDT, R. E. A comparison of methods for testing nonnested hypotheses. *The Review of Economics and Statistics*, v.56, n.1, p. 92–99, 1974.
- QUESENBERRY, C. P.; KENT, J. Selecting among Probability Distributions Used in Reliability. *Technometrics*, v.24, n.1, p. 59–65, 1982.
- R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. 2016.
- STATISTICAT, LCC *LaplacesDemon: A Complete Environment for Bayesian Inference within R*. R Package version 17.07.2016. <https://cran.r-project.org/web/packages/LaplacesDemon/LaplacesDemon.pdf>.
- STERN, J.; PEREIRA, C. A. B. Bayesian epistemic values: Focus on surprise, measure probability. *Logic Journal of The IGPL*, v.22, n.2, p.236–254, 2014.

Recebido em 15.08.2016.

Aprovado após revisão em 24.02.2017.