



## MODELOS DE REGRESSÃO LOG-BIRNBAUM-SAUNDERS

Rafael Baptista de Assis<sup>1</sup>, Michelli Barros<sup>2</sup>

### RESUMO

Os modelos de regressão log-Birnbaum-Saunders tem sido largamente aplicado na análise de dados de sobrevivência e confiabilidade e relaciona o tempo até a ocorrência da falha com algum dano acumulativo que é assumido Gaussiano. Este trabalho teve como objetivo o estudo de algumas medidas de diagnóstico de influência para os modelos de regressão log-Birnbaum-Saunders, bem como o estudo de um teste de hipóteses para outliers baseado no modelo de outlier por deslocamento da média.

**Palavras-chave:** sobrevivência, modelos Birnbaum-Saunders, teste escore.

### LOG-BIRNBAUM-SAUNDERS REGRESSION MODELS

#### ABSTRACT

The Log-Birnbaum-Saunders Regression Models have been widely applied to reliability and survival analysis relating the time until the occurrence of failure to cumulative damage, assumed Gaussian. The main purpose of this work was the study of some influence diagnostics measures for the Birnbaum-Saunders-Regression Models as well as an outlier test of hypothesis based on Mean Shift Outlier Model.

**Keywords:** survival, Birnbaum-Saunders models, score test.

#### INTRODUÇÃO

A análise de sobrevivência é um ramo da estatística que trata da modelagem do tempo até a ocorrência de certo evento (denominado de falha), podendo este ser a morte de um paciente de hospital, ou a ocorrência de rachadura em um determinado equipamento mecânico. O interesse, então, é descrever o tempo de vida (isto é, o tempo ou o número de ciclos até a ocorrência do evento) em função de variáveis explicativas, variáveis estas que vão depender do estudo que será realizado.

A partir de considerações provenientes da Física e da engenharia, Birnbaum e Saunders(1969a) derivaram uma distribuição que hoje é conhecida como distribuição Birnbaum-Saunders e esta vem recebendo muita atenção nas últimas décadas. Esta atenção toda é resultado de seus argumentos teóricos, de suas propriedades atrativas e de sua relação com a distribuição normal. Estes aspectos do modelo Birnbaum-Saunders reforçam a proposta de ser usado como um candidato alternativo ao modelo normal quando analisamos dados não-negativos que seguem uma assimetria positiva. Pela sua origem teórica da engenharia, foi utilizada principalmente na análise de dados de confiabilidade. Contudo, trabalhos recentes como os de Leiva et al.(2007) e Barros et al. (2008), sob o argumento de que muitos problemas médicos têm origem num dano cumulativo (por exemplo, doenças cardíacas ou diferentes tipos de câncer), aplicaram a distribuição à dados de sobrevivência.

<sup>1</sup> Aluno do Curso de Matemática, Unidade Acadêmica de Matemática e Estatística, UFCG, Campina Grande, PB, E-mail: [rafaelbassis@gmail.com](mailto:rafaelbassis@gmail.com)

<sup>2</sup> Profa. Doutora, Unidade Acadêmica de Matemática e Estatística, UFCG, Campina Grande, PB, E-mail: [michelli@dme.ufcg.edu.br](mailto:michelli@dme.ufcg.edu.br)

Em trabalho recente, Díaz-García e Leiva (2005) generalizaram a distribuição Birnbaum-Saunders a partir de distribuições de contornos elípticos, alegando que distribuições com caudas mais leves ou pesadas surgiriam a partir de tal generalização.

Modelagem estatística sob as distribuições Birnbaum-Saunders e Birnbaum-Saunders generalizada tem recebido uma crescente atenção na literatura nos últimos anos. No caso clássico, alguns trabalhos têm sido desenvolvidos, os quais podem ser vistos em Rieck e Nedelman (1991), Owen e Padgett (1999, 2000) e Tsionas (2001). Galea et al.(2004) abordaram o diagnóstico de influência no modelo de regressão log-Birnbaum-Saunders (LBSRM), utilizando aí um argumento físico para a homogeneidade do parâmetro de forma. Xie e Wei (2007) obtiveram algumas medidas de diagnóstico de influência e propuseram um teste para a detecção de outlier, bem como para testar a homogeneidade para o parâmetro de forma em questão. Nosso objetivo nesse trabalho é desenvolver, em linhas gerais, um pouco da teoria envolvendo o modelo de regressão log-Birnbaum-Saunders apresentando alguns resultados obtidos por Xie e Wei (2007) envolvendo diagnóstico de influência.

## DISTRIBUIÇÃO BIRNBAUM-SAUDERS E SENH-NORMAL

Uma variável aleatória  $T$  tem distribuição Birnbaum-Saunders de parâmetro de forma  $\alpha$  e parâmetro de escala  $\beta$  se sua f.d.p. é dada por (nesse caso, escrevemos  $T \sim BS(\alpha, \beta)$ ):

$$f_T(t) = \frac{1}{2\alpha\beta\sqrt{2\pi}} \left[ \left(\frac{\beta}{t}\right)^{1/2} + \left(\frac{\beta}{t}\right)^{3/2} \right] \exp\left\{-\frac{1}{2\alpha^2} \left[\frac{t}{\beta} + \frac{\beta}{t} - 2\right]\right\}$$

Uma distribuição não menos importante para os modelos de regressão log-Birnbaum-Saunders é a distribuição senh-normal, explorada por Rieck (1989). Uma variável aleatória  $Y$  tem distribuição senh-normal com parâmetro de forma  $\alpha$ , parâmetro de localização  $\gamma$  e parâmetro de escala  $\sigma$ , se:

$$Y = \operatorname{arcsenh}(\alpha Z/2)\sigma + \gamma$$

com  $Z \sim N(0, 1)$ . Denotamos este fato por  $Y \sim SN(\alpha, \gamma, \sigma)$ , e sua f.d.p. é dada por:

$$f_Y(y) = \left(\frac{2}{\alpha\sigma\sqrt{2\pi}}\right) \cosh\left(\frac{y-\gamma}{\sigma}\right) \exp\left[-2\alpha^{-2}\operatorname{senh}^2\left(\frac{y-\gamma}{\sigma}\right)\right], \quad y \in \mathbb{R}.$$

Rieck e Nedelman (1991) provaram que, se  $T \sim BS(\alpha, \beta)$ , então  $\log T = Y \sim SN(\alpha, \log \beta, 2)$ . Por causa disto, a distribuição senh-normal é também conhecida como distribuição log-Birnbaum-Saunders.

## MODELO DE REGRESSÃO LOG-BIRNBAUM-SAUDERS

Mostraremos aqui o modelo log-Birnbaum-Saunders como apresentado em Rieck e Nedelman (1991). Considere  $n$  variáveis aleatórias  $Y_1, \dots, Y_n$  independentes com distribuição Birnbaum-Saunders de parâmetro de forma  $\alpha_i$  e parâmetro de escala  $\beta_i$ , de forma que a distribuição de  $Y_i$  dependa de um conjunto de variáveis explicativas  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$  do seguinte modo:

$$\beta_i = \exp(\mathbf{x}_i' \boldsymbol{\theta}), \quad \text{para } i=1, \dots, n,$$

em que  $\boldsymbol{\theta}' = (\theta_1, \dots, \theta_p)$  é um vetor de parâmetros a serem estimados, e  $\alpha_i = \alpha$ , para todo  $i$ , indicando que o parâmetro de forma é independente do vetor explicativo  $\mathbf{x}_i$ . Birnbaum e Saunders (1969a) mostraram que, se  $T \sim BS(\alpha, \beta)$ ,  $cT \sim BS(\alpha, c\beta)$ . Deste fato, segue que  $Y_i = \exp(\mathbf{x}_i' \boldsymbol{\theta}) \delta_i$ , onde  $\delta_i \sim BS(\alpha, 1)$ . Fazendo  $L_i = \log Y_i$ , temos:

$$L_i = \mathbf{x}_i' \boldsymbol{\theta} + \log(\delta_i), \quad \text{para } i=1, \dots, n.$$

Agora fazendo  $\boldsymbol{\varepsilon}_i = \log(\delta_i)$ , temos um modelo log-linear de regressão com erro  $\boldsymbol{\varepsilon}_i \sim SN(\alpha, 0, 2)$ . É conveniente agora mudarmos nossa notação. Faremos então  $\boldsymbol{\beta}' = (\beta_1, \dots, \beta_p)$  o vetor de parâmetros a serem estimados,  $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\alpha})'$ , e denotaremos  $L_i$  simplesmente por  $Y_i$ . Considerando este modelo, Rieck e Nedelman (1991) obtiveram a função de log-verossimilhança abaixo (para um conjunto de observações  $(y_1, \dots, y_n)^T$ ):

$$l(\theta) = \sum_{i=1}^n \log \left\{ \frac{1}{\alpha\sqrt{2\pi}} \cosh\left(\frac{y_i - \mu_i}{2}\right) \exp\left[-\frac{2}{\alpha^2} \sinh^2\left(\frac{y_i - \mu_i}{2}\right)\right] \right\}$$

$$= -\frac{n}{2} \log 8\pi + \sum_{i=1}^n \log \xi_{i1} - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2,$$

em que

$$\mu_i = X_i^T \beta, \text{ para } i = 1, \dots, n$$

$$\xi_{i1} = \frac{2}{\alpha} \cosh\left(\frac{y_i - \mu_i}{2}\right), \quad \xi_{i2} = \frac{2}{\alpha} \sinh\left(\frac{y_i - \mu_i}{2}\right)$$

O estimador de máxima verossimilhança do vetor de parâmetros  $\beta$  desse modelo deve ser obtido por algum procedimento numérico. Já para o parâmetro  $\alpha$ , Rieck e Nedelman (1991) obtiveram

$$\hat{\alpha}^2 = \frac{4}{n} \sum_{i=1}^n \sinh^2\left(\frac{y_i - X_i^T \hat{\beta}}{2}\right)$$

### MODELO DE DELEÇÃO DE CASOS (CDM)

Para se realizar um estudo da influência da  $i$ -ésima observação nas estimativas e inferências do modelo, é comum se utilizar do modelo de deleção de casos (Case Deletion Model ou CDM), que consiste em excluir a  $i$ -ésima observação do modelo e observar o modelo de regressão resultante. Assim, o CDM para o modelo de regressão log-Birnbaum-Saunders é:

$$y_j = X_j^T \beta + \varepsilon_j, \quad j = 1, 2, \dots, n, \quad j \neq i$$

Denotaremos as quantidades relativas ao modelo com exclusão do  $i$ -ésimo caso subscrevendo "(i)", no que segue. Por exemplo,  $l_{(i)}(\theta)$  denota a função de log-verossimilhança para o modelo com o  $i$ -ésimo caso deletado. A idéia aqui é comparar a diferença entre  $\hat{\theta}_{(i)}$  e  $\hat{\theta}$ . Se a deleção do  $i$ -ésimo caso influenciar seriamente as estimativas, mais atenção deverá ser dada a esta observação.

A partir do Modelo de Deleção de Casos (CDM) podemos obter estatísticas que nos mostram o quão influentes são determinadas observações. Para o estimador de máxima verossimilhança  $\hat{\theta} = (\hat{\beta}^T, \hat{\alpha})^T$ , devemos basicamente comparar  $\hat{\theta}_{(i)}$  e  $\hat{\theta}$ . Se a distância entre eles for grande, a  $i$ -ésima observação é influente no modelo. Para reduzir um pouco os cálculos, usamos a aproximação abaixo (proposta por Cook e Weisberg, 1982):

$$\hat{\theta}_{(i)}^1 = \hat{\theta} + \{-\ddot{l}(\hat{\theta})\}^{-1} \dot{l}_{(i)}(\hat{\theta}),$$

em que  $\dot{l}_{(i)}(\hat{\theta}) = \partial l_{(i)}(\theta) / \partial \theta |_{\theta=\hat{\theta}}$ ,  $\ddot{l}(\hat{\theta}) = \partial^2 l(\theta) / \partial \theta \partial \theta^T |_{\theta=\hat{\theta}}$ .

Para obter a matriz de informação observada, passamos às derivadas da função de log-verossimilhança:

$$\frac{\partial^2 l(\theta)}{\partial \beta \partial \beta^T} = \sum_{i=1}^n \left[ \frac{1}{4} \operatorname{sech}^2 \left( \frac{y_i - \mu_i}{2} \right) - \frac{1}{\alpha^2} \cosh(y_i - \mu_i) \right] X_i X_i^T = X^T V X,$$

$$\frac{\partial^2 l(\theta)}{\partial \beta \partial \alpha} = -\frac{2}{\alpha^3} \sum_{i=1}^n \sinh(y_i - \mu_i) X_i = X^T h,$$

$$\frac{\partial^2 l(\theta)}{\partial \alpha^2} = \frac{n}{\alpha^2} - \frac{12}{\alpha^4} \sum_{i=1}^n \sinh^2 \left( \frac{y_i - \mu_i}{2} \right) = \delta.$$

$$V = \operatorname{diag}(v_1, v_2, \dots, v_n), \quad v_i = \frac{1}{4} \operatorname{sech}^2 \left( \frac{y_i - \mu_i}{2} \right) - \frac{1}{\alpha^2} \cosh(y_i - \mu_i),$$

$$h = (h_1, h_2, \dots, h_n)^T, \quad h_i = -\frac{2}{\alpha^3} \sinh(y_i - \mu_i)$$

Obtendo assim a matriz de informação observada:

$$\ddot{l}(\hat{\theta}) = \begin{bmatrix} \frac{\partial^2 l(\theta)}{\partial \beta \partial \beta^T} & \frac{\partial^2 l(\theta)}{\partial \beta \partial \alpha} \\ \frac{\partial^2 l(\theta)}{\partial \alpha \partial \beta^T} & \frac{\partial^2 l(\theta)}{\partial \alpha^2} \end{bmatrix}_{\hat{\theta}} = \begin{bmatrix} X^T V X & X^T h \\ h^T X & \delta \end{bmatrix}_{\hat{\theta}}.$$

Uma matriz em blocos, que podemos inverter com técnicas que podem ser vistas em qualquer livro de álgebra de matrizes, obtendo então:

$$-\ddot{l}^{-1}(\hat{\theta}) = \begin{pmatrix} -X^T A X^{-1} & \delta^{-1} (X^{-1} A^{-1} h) \\ \delta^{-1} h^T A^{-1} (X^T)^{-1} & -(\delta - h^T V^{-1} h)^{-1} \end{pmatrix}$$

$$A = V - h \delta^{-1} h^T$$

Onde:

Agora calculamos as derivadas da função de verossimilhança com a  $i$ -ésima observação excluída. Como

$$\frac{\partial l(\theta)}{\partial \beta} |_{(\hat{\beta}, \hat{\alpha})} = 0, \quad \frac{\partial l(\theta)}{\partial \alpha} |_{(\hat{\beta}, \hat{\alpha})} = 0,$$

segue diretamente que:

$$\frac{\partial l_{(i)}(\theta)}{\partial \beta} |_{(\hat{\beta}, \hat{\alpha})} = X_i (\xi_{i2} / (2\xi_{i1}) - \xi_{i1} \xi_{i2} / 2) |_{(\hat{\beta}, \hat{\alpha})} = (-X_i r_i) |_{(\hat{\beta}, \hat{\alpha})},$$

$$\frac{\partial l_{(i)}(\theta)}{\partial \alpha} |_{(\hat{\beta}, \hat{\alpha})} = (1/\alpha - \xi_{i2}^2 / \alpha) |_{(\hat{\beta}, \hat{\alpha})} = (-\eta_i) |_{(\hat{\beta}, \hat{\alpha})}.$$

Sendo

$$r_i = \xi_{i1} \xi_{i2} / 2 - \xi_{i2} / (2\xi_{i1}), \quad \eta_i = -1/\alpha + \xi_{i2}^2 / \alpha.$$

Realizando a multiplicação de matrizes, vem:

$$\hat{\beta}_{(i)}^1 = \hat{\beta} + \{(X^TAX)^{-1}(X_i r_i - X^T h \delta^{-1} \eta_i)\}_{\hat{\theta}},$$

$$\hat{\alpha}_{(i)}^1 = \hat{\alpha} + \{[\delta^{-1} + \delta^{-1} h^T X (X^TAX)^{-1} X^T h \delta^{-1}] \eta_i - \delta^{-1} h^T X (X^TAX)^{-1} X_i r_i\}_{\hat{\theta}}$$

Este resultado é importante porque nos permite comparar o estimador de máxima verossimilhança do modelo de regressão original e o do CDM com a  $i$ -ésima observação excluída, a partir do que veremos ser a distância de Cook Generalizada ( $GD_i$ ). A definição da Distância Generalizada de Cook é:

$$GD_i = (\hat{\theta}_{(i)} - \hat{\theta})^T M (\hat{\theta}_{(i)} - \hat{\theta})$$

Ou seja, é a norma padronizada da diferença entre os estimadores, porque  $M$  é uma matriz não-negativa definida. A escolha da matriz  $M$  foi considerada por Cook e Weisberg(1982), e foram citadas algumas possibilidades. Neste caso foi escolhida a matriz de informação observada. Com algum cálculo (a partir das aproximações feitas acima) chegamos à fórmula abaixo (que é uma aproximação):

$$GD_i^1 = l_{(i)}(\hat{\theta})^T \{-\ddot{l}(\hat{\theta})\}^{-1} l_{(i)}(\hat{\theta})$$

Utilizando os resultados obtidos anteriormente, é fácil ver que:

$$GD_i^1 = \{-r_i X_i^T (X^TAX)^{-1} X_i r_i + 2r_i X_i^T (X^TAX)^{-1} X^T h \delta^{-1} \eta_i - \eta_i^T [\delta^{-1} + \delta^{-1} h^T X (X^TAX)^{-1} X^T h \delta^{-1}] \eta_i\}_{\hat{\theta}},$$

O que nos dá uma forma de calcular a influência da  $i$ -ésima observação nas estimativas. Agora, se quisermos a influência nas estimativas de um dos parâmetros apenas, é conveniente ter as seguintes medidas:

$$GD_i^1(\beta) = \{l_{(i)\beta}(\hat{\theta})^T \{(\mathbf{I}_p, \mathbf{0})[-\ddot{l}(\hat{\theta})]^{-1} (\mathbf{I}_p, \mathbf{0})^T\} l_{(i)\beta}(\hat{\theta})\}_{\hat{\theta}},$$

$$GD_i^1(\alpha) = \{l_{(i)\alpha}(\hat{\theta})^T \{(\mathbf{0}_p, 1)[- \ddot{l}(\hat{\theta})]^{-1} (\mathbf{0}_p, 1)^T\} l_{(i)\alpha}(\hat{\theta})\}_{\hat{\theta}},$$

Onde  $l_{(i)\beta}(\hat{\theta}) = \partial l_{(i)}(\theta) / \partial \beta|_{\theta=\hat{\theta}}$ ,  $l_{(i)\alpha}(\hat{\theta}) = \partial l_{(i)}(\theta) / \partial \alpha|_{\theta=\hat{\theta}}$ ,  $\mathbf{I}_p$  é a matriz identidade de ordem  $p$ ,  $\mathbf{0}_p$  é a matriz  $p \times p$  com entradas nulas,  $\mathbf{0} = (0, \dots, 0)^T$  e  $\mathbf{1} = (1, \dots, 1)^T$ . Tais medidas revelam o impacto da  $i$ -ésima observação nos estimadores de máxima verossimilhança separadamente. Computando o resultado, temos:

$$GD_i^1(\beta) = \{-r_i X_i^T (X^TAX)^{-1} X_i r_i\}_{\hat{\theta}},$$

$$GD_i^1(\alpha) = \{-\eta_i^T [\delta^{-1} + \delta^{-1} h^T X (X^TAX)^{-1} X^T h \delta^{-1}] \eta_i\}_{\hat{\theta}},$$

Outra medida muito usada é a distância de verossimilhanças. Sua definição vem logo abaixo:

$$LD_i = 2\{l(\hat{\theta}) - l(\hat{\theta}_{(i)})\}$$

Podemos utilizar diretamente os resultados anteriores para calculá-la para o nosso modelo. Temos, então:

$$LD_i^1 = 2\{l(\hat{\theta}) - l(\hat{\theta} + \{-\ddot{l}(\hat{\theta})\}^{-1} l_{(i)}(\hat{\theta}))\}$$

## TESTE PARA OUTLIERS

O Modelo de Outlier por Deslocamento da Média (Mean Shift Outlier Model ou MSOM) é construído adicionando-se um parâmetro a uma das observações no modelo original, a saber:

$$\begin{cases} y_j = X_j^T \beta + \varepsilon_j, & j = 1, 2, \dots, n, \quad j \neq i, \\ y_i = X_i^T \beta + \gamma + \varepsilon_i, \end{cases}$$

Claramente, se o parâmetro  $\gamma$  for não-nulo, a  $i$ -ésima observação pode ser considerada um outlier, pois não corresponde ao modelo original.

Um resultado interessante o seguinte: se os estimadores de máxima verossimilhança dos modelos CDM e MSOM são únicos, então

$$\hat{\beta}_{mi} = \hat{\beta}_{(i)}, \quad \hat{\alpha}_{mi}^2 \approx \hat{\alpha}_{(i)}^2.$$

Indicamos com subscrito “mi” os estimadores de máxima verossimilhança do modelo MSOM (utilizaremos essa notação para demonstrar o resultado acima). Para mostrar que o resultado é verdadeiro, consideramos as equações abaixo no CDM (válidas para  $\hat{\beta}_{(i)}$  e  $\hat{\alpha}_{(i)}$ ):

$$\frac{\partial l_{(i)}(\theta)}{\partial \beta} = \sum_{j \neq i} \left( \frac{\xi_{j1} \xi_{j2}}{2} - \frac{\xi_{j2}}{2 \xi_{j1}} \right) X_j = 0,$$

$$\frac{\partial l_{(i)}(\theta)}{\partial \alpha} = -\frac{n-1}{\alpha} + \frac{1}{\alpha} \sum_{j \neq i} \xi_{j2}^2 = 0.$$

Para o modelo MSOM, temos:

$$\begin{aligned} l_{mi}(\theta) &= -\frac{n}{2} \log 8\pi + \left( \sum_{j \neq i} \log \xi_{j1} - \frac{1}{2} \sum_{j \neq i} \xi_{j2}^2 \right) + \left( \log \bar{\xi}_{i1} - \frac{1}{2} \bar{\xi}_{i2}^2 \right) \\ &= l_{(i)}(\theta) - \frac{1}{2} \log 8\pi + \left( \log \bar{\xi}_{i1} - \frac{1}{2} \bar{\xi}_{i2}^2 \right), \end{aligned}$$

com

$$\bar{\xi}_{i1} = \frac{2}{\alpha} \cosh \left( \frac{y_i - X_i^T \beta - \gamma}{2} \right), \quad \bar{\xi}_{i2} = \frac{2}{\alpha} \sinh \left( \frac{y_i - X_i^T \beta - \gamma}{2} \right).$$

Agora, note que os estimadores de máxima verossimilhança pelo MSOM satisfazem:

$$\frac{\partial l_{mi}(\theta)}{\partial \beta} = \frac{\partial l_{(i)}(\theta)}{\partial \beta} - \left\{ \frac{1}{\alpha \bar{\xi}_{i1}} \sinh \left( \frac{y_i - X_i^T \beta - \gamma}{2} \right) - \frac{\bar{\xi}_{i2}}{\alpha} \cosh \left( \frac{y_i - X_i^T \beta - \gamma}{2} \right) \right\} X_i = 0,$$

$$\frac{\partial l_{mi}(\theta)}{\partial \gamma} = - \left\{ \frac{1}{\alpha \bar{\xi}_{i1}} \sinh \left( \frac{y_i - X_i^T \beta - \gamma}{2} \right) - \frac{\bar{\xi}_{i2}}{\alpha} \cosh \left( \frac{y_i - X_i^T \beta - \gamma}{2} \right) \right\} = 0.$$

Que combinadas nos dão:

$$\frac{\partial l_{mi}(\theta)}{\partial \beta} = \frac{\partial l_{(i)}(\theta)}{\partial \beta} = 0.$$

De onde vem a primeira igualdade, sob a hipótese de unicidade de estimadores de máxima verossimilhança. Agora, para a segunda parte, usamos um resultado de Rieck e Nedelman (1991):

$$\hat{\alpha}_{(i)}^2 = \frac{4}{n-1} \sum_{j \neq i} \sinh^2 \left( \frac{y_j - X_j^T \hat{\beta}_{(i)}}{2} \right),$$

$$\begin{aligned} \hat{\alpha}_{mi}^2 &= \frac{4}{n} \sum_{j \neq i} \sinh^2 \left( \frac{y_j - X_j^T \hat{\beta}_{mi}}{2} \right) + \frac{4}{n} \sinh^2 \left( \frac{y_i - X_i^T \hat{\beta}_{mi} - \hat{\gamma}_{mi}}{2} \right) \\ &= \left(1 - \frac{1}{n}\right) \hat{\alpha}_{(i)}^2 + \frac{4}{n} \sinh^2 \left( \frac{y_i - X_i^T \hat{\beta}_{mi} - \hat{\gamma}_{mi}}{2} \right). \end{aligned}$$

Que, com alguma manipulação algébrica nos dá

$$\hat{\alpha}_{mi}^2 = (1 - 1/n) \hat{\alpha}_{(i)}^2 \approx \hat{\alpha}_{(i)}^2, \text{ or } \hat{\alpha}_{mi}^2 = \hat{\alpha}_{(i)}^2 - 4/(n-1) \approx \hat{\alpha}_{(i)}^2$$

quando  $n$  tende ao infinito.

Como mencionado antes, se o parâmetro extra ( $\gamma$ ) for não-nulo indica um outlier. Então, é interessante testar a hipótese para identificar outlier  $H_0: \gamma=0$  contra  $H_1: \gamma \neq 0$ . Se a hipótese  $H_0$  for rejeitada então o  $i$ -ésimo caso pode ser um possível outlier. Para testar as hipóteses acima consideramos a estatística escore, como em Xie e Wei (2007). Para obtenção da estatística escore, devemos achar as duas primeiras derivadas da função  $l_{mi}(\theta)$  sob  $H_0$ . As derivadas com relação a  $\alpha$  e  $\beta$  já foram encontradas. Agora, basta calcular as derivadas abaixo:

$$\frac{\partial l_{mi}(\theta)}{\partial \gamma} \Big|_{\hat{\theta}} = \left( -\frac{1}{2} \frac{\xi_{i2}}{\xi_{i1}} + \frac{1}{2} \xi_{i1} \xi_{i2} \right)_{\hat{\theta}} = (r_i)_{\hat{\theta}},$$

$$\frac{\partial^2 l_{mi}(\theta)}{\partial \gamma^2} \Big|_{\hat{\theta}} = \left[ -\frac{1}{4} \left( \frac{\xi_{i2}}{\xi_{i1}} \right)^2 + \frac{1}{4} - \frac{1}{4} (\xi_{i1}^2 + \xi_{i2}^2) \right]_{\hat{\theta}} = (u_i)_{\hat{\theta}},$$

$$\frac{\partial^2 l_{mi}(\theta)}{\partial \gamma \partial \beta^T} \Big|_{\hat{\theta}} = \left[ -\frac{1}{4} \left( \frac{\xi_{i2}}{\xi_{i1}} \right)^2 + \frac{1}{4} - \frac{1}{4} (\xi_{i1}^2 + \xi_{i2}^2) \right]_{\hat{\theta}} X_i^T = (u_i)_{\hat{\theta}} X_i^T,$$

$$\frac{\partial^2 l_{mi}(\theta)}{\partial \gamma \partial \alpha} \Big|_{\hat{\theta}} = \left( -\frac{1}{\alpha} \xi_{i1} \xi_{i2} \right)_{\hat{\theta}} = (d_i)_{\hat{\theta}}.$$

Obtendo a matriz de informação observada:

$$\ddot{l}(\hat{\theta}) = \begin{bmatrix} \frac{\partial^2 l(\theta)}{\partial \beta \partial \beta^T} & \frac{\partial^2 l(\theta)}{\partial \beta \partial \alpha} & \frac{\partial^2 l(\theta)}{\partial \beta \partial \gamma} \\ \frac{\partial^2 l(\theta)}{\partial \alpha \partial \beta^T} & \frac{\partial^2 l(\theta)}{\partial \alpha^2} & \frac{\partial^2 l(\theta)}{\partial \alpha \partial \gamma^T} \\ \frac{\partial^2 l(\theta)}{\partial \gamma \partial \beta^T} & \frac{\partial^2 l(\theta)}{\partial \gamma \partial \alpha} & \frac{\partial^2 l(\theta)}{\partial \gamma^2} \end{bmatrix}_{\hat{\theta}} = \begin{bmatrix} X^T V X & X^T h & u_i X_i \\ h^T X & \delta & d_i \\ u_i X_i^T & d_i & u_i \end{bmatrix}.$$

A qual pode ser particionada por blocos de forma que as derivadas com relação a  $\alpha$  e  $\beta$  exclusivamente fiquem num mesmo bloco 2x2. Assim, precisamos apenas obter o termo da segunda linha e da segunda coluna da inversa da matriz observada (colocada por blocos), e utilizando as técnicas de inversão mencionadas antes, temos a estatística escore abaixo:

$$SC_i = \{r_i^2 \rho_i^{-1}\}_{\hat{\theta}},$$

Com

$$\rho_i = u_i^2 X_i^T K X_i + n^{-1} \alpha^2 u_i d_i h^T X K X_i - (2n)^{-1} \alpha^2 d_i^2 + (4n^2)^{-1} \alpha^4 d_i^2 h^T X K X^T h - u_i, \quad d_i = -\xi_{i1} \xi_{i2} / \alpha,$$

$$K = [X^T (V + (2n)^{-1} \hat{\alpha}^2 h h^T) X]^{-1}, \quad u_i = (1 - \xi_{i1}^2 - \xi_{i2}^2 - \xi_{i2}^2 / \xi_{i1}^2) / 4,$$

e  $r_i$  e  $h$  são dados na distância de Cook.

## APLICAÇÃO

Consideramos o conjunto de dados usado por Rieck e Nedelman (1991) e Xie e Wei (2007) como exemplo da aplicabilidade do teste escore para detecção de outlier. Este se refere a dados de fadiga biaxial sob o tempo até a ruptura de um pedaço de metal sujeito a tensões cíclicas. Aqui, a variável resposta é o número de ciclos até a ocorrência de falha, e a variável explicativa, o trabalho realizado até então. Para mais detalhes, veja Rieck e Nedelman (1991). Aplicamos a teoria descrita acima e detectamos os pontos 4, 5, 12, 32 e 46 como possivelmente influentes, confirmando a análise feita em Galea et al. (2004). A Figura 1 apresenta o gráfico da estatística escore contra os índices das observações. Como podemos observar os casos ditos acima são considerados como outlier, pois  $SC_i = 4,01; 3,94; 4,07; 4,23; \text{ e } 3,12$ , para os casos 4, 5, 12, 32 e 46, respectivamente. Como podemos observar se fixamos o nível de significância em 10% temos que o valor crítico é  $\chi_1^2 = 2,706$  e como  $SC_i$  é maior que  $\chi_1^2$  rejeitamos a hipótese  $H_0: \gamma = 0$ . Este resultado foi consistente com o resultado de Xie e Wei (2007).

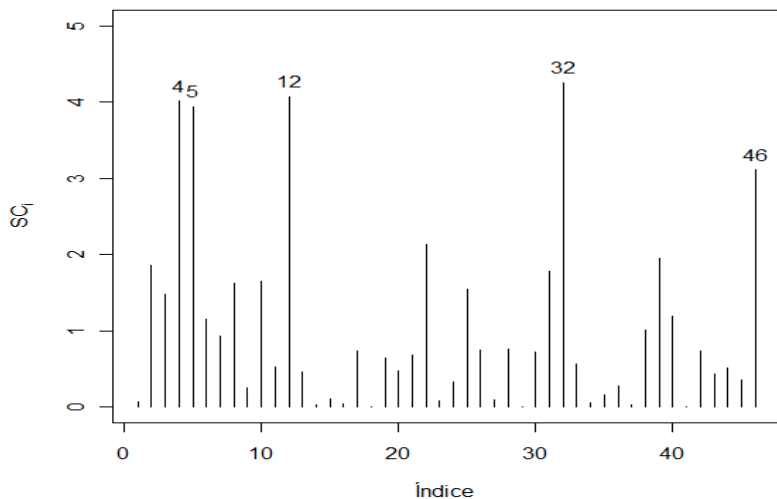


Figura 1. Estatística escore contra os índices das observações.



## AGRADECIMENTOS

Ao CNPq pela bolsa de Iniciação Científica.

## REFERÊNCIAS BIBLIOGRÁFICAS

LEIVA, V., Barros, M., Paula, G. A. e Galea, M. (2007). Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. *Computational Statistics and Data Analysis* 51, 5694-5707.

BARROS, M., Paula, G.A., Leiva, V. (2008) A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. *Lifetime Data Analysis* (in press).

OWEN, W. J. e Padgett, W. J. (2000). A Birnbaum-Saunders accelerated life model. *IEEE Transactions on Reliability* 49, 224-229.

RIECK, J. R. e Nedelman, J. R. (1991). A log-linear model for the Birnbaum-Saunders distribution. *Technometrics* 33, 51-60.

TSIONAS, E. G. (2001). Bayesian inference in Birnbaum-Saunders regression. *Communications in Statistics: Theory and Methods* 30, 179-193.

GALEA, M., Leiva, V. e Paula, G. A. (2004). Influence diagnostics in log-Birnbaum-Saunders regression models. *Journal of Applied Statistics* 31, 1049-1064.

BIRNBAUM, Z. W. e Saunders, S. C. (1969a). A new family of life distributions. *Journal of Applied Probability* 6, 319-327.

DÍAZ-GARCÍA, J.A., Leiva, V. (2005) A new family of life distributions based on elliptically contoured distributions. *J. Stat. Plan. Infer.* 128:445-457; Erratum, *J. Stat. Plan. Infer.* 137:1512-1513.

XIE, Feng-Chang e Wei, Bo-Cheng (2007). Diagnostics analysis for log-Birnbaum-Saunders regression models. *Computational Statistics and Data Analysis.* 51, 4692-4706.