



UNIVERSIDADE FEDERAL DO CEARÁ
DEPARTAMENTO DE ESTATÍSTICA E MATEMÁTICA APLICADA
CURSO DE ESTATÍSTICA

RODNEY VASCONCELOS FONSECA

**DISTRIBUIÇÃO BIRNBAUM-SAUNDERS: ASPECTOS CLÁSSICOS E
BAYESIANOS**

FORTALEZA

2014

RODNEY VASCONCELOS FONSECA

**DISTRIBUIÇÃO BIRNBAUM-SAUNDERS: ASPECTOS CLÁSSICOS E
BAYESIANOS**

Monografia submetida à Coordenação do Curso de Estatística do Departamento de Estatística e Matemática Aplicada da Universidade Federal do Ceará, como requisito parcial para obtenção do título de Bacharel em Estatística.

Área de concentração: Estatística e Probabilidade

Orientador: Prof. Dr. Juvêncio Santos Nobre

Co-Orientador: Prof. Dr. Rafael Bráz Azevedo
Farias

FORTALEZA

2014

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca do Curso de Física

-
- F747d Fonseca, Rodney Vasconcelos
Distribuição Birnbaum-Saunders: aspectos clássicos e bayesianos / Rodney Vasconcelos
Fonseca. – Fortaleza, 2014.
82 f. : il. enc.; 30 cm.
- Monografia (Graduação em Estatística) – Universidade Federal do Ceará, Centro de Ciências,
Departamento de Estatística e Matemática Aplicada, Curso de Estatística, Fortaleza, 2014.
Orientação: Prof. Dr. Juvêncio Santos Nobre.
Coorientação: Prof. Dr. Rafael Bráz Azevedo Farias.
Área de concentração: Probabilidade e Estatística.
Inclui bibliografia e apêndices.
1. Estatística. 2. Birnbaum-Saunders, modelo de. 3. Análise de regressão. 4. Teoria bayesiana de decisão estatística. I. Nobre, Juvêncio Santos. II. Farias, Rafael Bráz Azevedo. III. Título.

RODNEY VASCONCELOS FONSECA

**DISTRIBUIÇÃO BIRNBAUM-SAUNDERS: ASPECTOS CLÁSSICOS E
BAYESIANOS**

Monografia submetida à Coordenação do Curso de Estatística do Departamento de Estatística e Matemática Aplicada da Universidade Federal do Ceará, como requisito parcial para obtenção do título de Bacharel em Estatística. Área de concentração: Estatística e Probabilidade

Aprovada em: __/__/____

BANCA EXAMINADORA

Prof. Dr. Juvêncio Santos Nobre
Universidade Federal do Ceará - UFC
Orientador

Prof. Dr. Rafael Bráz Azevedo Farias
Universidade Federal do Ceará - UFC
Co-orientador

Prof. Dr. João Maurício Araújo Mota
Universidade Federal do Ceará - UFC

Aos meus Pais.

AGRADECIMENTOS

Dedico este trabalho aos meus pais, a todos os meus familiares, aos meus amigos, meus professores e a todos que de alguma forma contribuíram para o meu crescimento acadêmico.

”Alea iacta est”

(Júlio César)

RESUMO

A distribuição Birnbaum-Saunders é um modelo utilizado para analisar o tempo de vida de materiais sujeitos a algum dano cumulativo, sendo um exemplo de distribuição de tempo de vida por fadiga. A distribuição tem recebido bastante atenção ultimamente, sendo tema para diversos trabalhos, inclusive com uma proposta de uma forma reparametrizada como objeto de modelagem, em que a média da distribuição é utilizada como um dos seus parâmetros. Isto permite que o ajuste de modelos de regressão de maneira semelhante à realizada em modelos lineares generalizados (MLG) possam ser feitos para a distribuição Birnbaum-Saunders. O objetivo deste trabalho consiste em discutir algumas propriedades da distribuição Birnbaum-Saunders reparametrizada (BSR), assim como formas de estimação dos parâmetros da BSR, pelos métodos de máxima verossimilhança, momentos e momentos modificado. Modelos de regressão Birnbaum-Saunders (MRBS) são apresentados, tratando do seu desenvolvimento e método de estimação. Além da abordagem clássica, a abordagem Bayesiana da distribuição também será discutida no trabalho, tratando das distribuições *a priori* dos parâmetros e das estimativas obtidas com as distribuições *a posteriori*, tanto para casos univariados como em modelos de regressão Bayesianos. Um exemplo de aplicação será feito em dados sobre o tempo de vida de pacientes com leucemia, analisando o ajuste dos modelos de regressão com a distribuição Birnbaum-Saunders aos dados. Com intuito de verificar a sensibilidade dos modelos no ajuste aos dados, neste trabalho é realizada uma comparação entre os métodos de análise de diagnóstico de influência dos modelos, no MRBS clássico e Bayesiano, verificando as similaridades e diferenças entre ambos, obtendo então mais informações para decisões de um modelo a ser adotado.

Palavras-chave: Distribuição Birnbaum-Saunders. Modelos de regressão. Análise de diagnóstico.

ABSTRACT

The Birnbaum-Saunders distribution is a model used to analyze lifetime of objects that suffer from a cyclic cumulative damage, consisting in an example of fatigue lifetime distribution. The distribution has received great attention lately, serving as theme of many works, including one proposal of reparametrization used in modeling, where the mean is used as one of its parameters. This allows the regression model with the Birnbaum-Saunders distribution to be fitted in a similar manner as a generalized linear model (GLM). The objective of this work is to discuss some properties of the Birnbaum-Saunders distribution, also estimation procedures: the maximum likelihood, moments method and modified moments method. Birnbaum-Saunders regression models (BSRM) are presented, with its development and estimation methods. Besides the classical version, the Bayesian approach is discussed in the work, presenting *priori* distributions of the parameters and estimatives obtained with *posteriori* distributions, for both univariate case and Bayesian regression models. An application example is done in a dataset about the lifetime of patients with leukemia, analyzing the adjustment of Birnbaum-Saunders regression models for this data. With the concern to handle a sensitive analyses for the models, influence diagnostic methods in the classic BSRM and the Bayesian model are compared, verifying the similarities and differences of both methodologies in evaluating influential observations, obtaining then more information to decide which model to use.

Keywords: Birnbaum-Saunders distribution. Regression models. Diagnostic analysis.

LISTA DE FIGURAS

Figura 1	Gráfico da função de distribuição acumulada da distribuição $BS(\alpha, \beta)$	18
Figura 2	Gráfico da função densidade da distribuição $BS(\alpha, \beta)$	18
Figura 3	Gráfico da função de distribuição acumulada da BSR.	21
Figura 4	Gráfico da função densidade da distribuição BSR	23
Figura 5	Boxplot das variáveis <i>semanas</i> e <i>IWBC</i> para cada nível de <i>AG</i>	55
Figura 6	Gráfico de dispersão das variáveis <i>semanas</i> e <i>IWBC</i> para cada nível de <i>AG</i>	55
Figura 7	Gráfico dos resíduos vs. os valores ajustados do MRBS para dados de leucemia.	56
Figura 8	Gráfico de envelope simulado para o MRBS dos dados de leucemia.	57
Figura 9	Gráfico de influência local para o MRBS dos dados de leucemia.	58
Figura 10	Gráfico da divergência K-L para o MRBS Bayesiano dos dados de leucemia.	59
Figura 11	Comparação dos gráficos de influência local e da divergência K-L.	60
Figura 12	Comparação dos valores de CPO nos modelos $M_1(a)$ e $M_2(b)$	62

LISTA DE TABELAS

Tabela 1	Estimativas de diferentes valores de μ com $\delta = 5$	30
Tabela 2	Estimativas de diferentes valores de δ com $\mu = 2$	30
Tabela 3	Estimativas dos parâmetros do MRBS na simulação.	43
Tabela 4	Resultados usando priores Uniforme	49
Tabela 5	Resultados usando priores Gama	49
Tabela 6	Resultados usando priores Log-normal	49
Tabela 7	Estimativas dos parâmetros do MRBS Bayesiano simulado.	51
Tabela 8	Análise descritiva.	54
Tabela 9	Estimativas do MRBS para dados de leucemia.	56
Tabela 10	Predições dos parâmetros do MRBS Bayesiano para os dados de leucemia.	58
Tabela 11	Calibração da divergência K-L dados de leucemia.	59
Tabela 12	Observações influentes da Figura 11.	61
Tabela 13	MR(%) para as observações da Tabela 12 no MRBS clássico.	61
Tabela 14	MR(%) para as observações da Tabela 12 no MRBS Bayesiano.	62
Tabela 15	Estimativas dos casos C_1 e C_2 sob enfoque clássico	62
Tabela 16	Estimativas dos casos C_1 e C_2 sob enfoque Bayesiano	63

LISTA DE ABREVIATURAS

BS: Distribuição Birnbaum-Saunders
BSR: Distribuição Birnbaum-Saunders Reparametrizada
fdp: Função densidade de probabilidade
fda: Função de distribuição acumulada
va: Variável aleatória
EMV: Estimador de Máxima Verossimilhança
EMM: Estimador pelo Método dos Momentos
EMMod: Estimador pelo Método dos Momentos Modificado
DP: Desvio Padrão
EQM: Erro Quadrático Médio
MRBS: Modelo de Regressão Birnbaum-Saunders
MCMC: Monte Carlo via Cadeias de Markov
MLG: Modelos Lineares Generalizados
MC error: Erro de Monte Carlo
CPO: *Conditional Predictive Ordinate*
PFB: Pseudo Fator de Bayes
K-L: Kullback-Leibler
AG: Fator indicativo da presença do Bastão de Auer em células da medula
lWBC: Logaritmo do número de células brancas
AIC: Critério de Informação de Akaike
BIC: Critério de Informação Bayesiano
LD(.): Medida de desvio na verossimilhança

LISTA DE SÍMBOLOS

- α : Parâmetro de escala da BS
 β : Parâmetro de forma da BS
 Φ : Função de distribuição acumulada da Normal Padrão
 ϕ : Função densidade de probabilidade da Normal Padrão
 μ : Parâmetro representando a média na BSR
 δ : Parâmetro de dispersão da BSR
 μ_i : Média da i -ésima observação
 β_i : i -ésimo coeficiente de regressão
 $\boldsymbol{\beta}$: Vetor dos coeficientes de regressão
 η : Preditor linear
 \mathbf{x}_i : Vetor de variáveis explicativas da i -ésima observação
 \mathbf{X} : Matriz de especificação do modelo
 $\boldsymbol{\theta}$: Vetor completo dos parâmetros do modelo de regressão
 \mathbf{U} : Vetor escore
 \mathbf{L} : Matriz Observada de Fisher
 $\text{tr}(\cdot)$: Traço de uma matriz
 $\text{diag}(\cdot)$: Matriz diagonal
 \mathbf{K} : Matriz Esperada de Fisher
 \mathbf{W} : Matriz de pesos no MRBS
 \mathcal{Z} : Variável dependente no MRBS
 Δ : Matriz de perturbação
 $K(\cdot, \cdot)$: Divergência de Kullback-Leibler

SUMÁRIO

1	INTRODUÇÃO	14
2	DISTRIBUIÇÃO BIRNBAUM-SAUNDERS	16
3	DISTRIBUIÇÃO BIRNBAUM-SAUNDERS REPARAMETRIZADA	21
3.1	Definições	21
3.2	Propriedades	22
3.3	Função de verossimilhança	25
3.4	Estimadores pelo método dos momentos	27
3.5	Estimador pelo método dos momentos modificado	28
3.6	Simulações	29
4	MODELOS DE REGRESSÃO BIRNBAUM-SAUNDERS	32
4.1	Definições do modelo	32
4.2	Derivadas da função do logaritmo da função verossimilhança	33
4.3	Matriz de informação esperada de Fisher	36
4.4	Simulações	42
4.5	Influência Local	43
5	ABORDAGEM BAYESIANA	47
5.1	Inferência Bayesiana	47
5.1.1	Distribuições a priori gama	47
5.1.2	Distribuições a priori uniforme	47
5.1.3	Distribuições a priori Log-Normal	48
5.1.4	Simulações	48
5.2	Modelo de regressão	49
5.2.1	Simulações	50
5.3	Comparação de modelos	51
5.4	Diagnóstico de influência Bayesiano	52
6	APLICAÇÃO	54
6.1	Pacientes com leucemia	54
6.1.1	Análise descritiva	54

6.1.2	Modelos de regressão clássicos	54
6.1.3	Modelos de regressão Bayesianos	57
6.1.4	Análise confirmatória	60
7	CONSIDERAÇÕES FINAIS	64
	REFERÊNCIAS BIBLIOGRÁFICAS	65
	APÊNDICE A – FUNÇÃO DENSIDADE DE PROBABILIDADE	68
	APÊNDICE B – RELAÇÃO ENTRE AS DISTRIBUIÇÕES <i>BS</i> E NORMAL	69
	APÊNDICE C – MOMENTO DE ORDEM <i>R</i> DA DISTRIBUIÇÃO <i>BS</i>.....	71
	APÊNDICE D – CÓDIGOS DO PROGRAMA <i>R</i> PARA A DISTRIBUIÇÃO <i>BSR</i> ...	73
D.1	Modelos usando somente o <i>R</i>	73
D.2	Modelo usando conjuntamente o OpenBUGS pelo R2OpenBUGS	79
	APÊNDICE E – CÓDIGOS DO PROGRAMA OPENBUGS PARA A DISTRIBUIÇÃO <i>BSR</i>	82

1 INTRODUÇÃO

A distribuição Birnbaum-Saunders é uma distribuição de suporte positivo utilizada em análise de tempo de vida. Ela foi proposta inicialmente por Birnbaum e Saunders (1969a), tendo sua forma gerada baseada em razões físicas do que ocorre com um objeto submetido a um processo de fadiga, sendo posteriormente mostrados por Birnbaum e Saunders (1969b) métodos de estimação dos parâmetros por máxima verossimilhança bem como algumas propriedades da distribuição.

Em relação a modelos de regressão, um modelo log-linear Birnbaum-Saunders foi proposto por Rieck e Nedelman (1991), obtendo a distribuição log-Birnbaum-Saunders como caso particular da distribuição Seno-Hiperbólico Normal. Modelos para tempo de vida acelerado foram propostos por Owen e Padget (2000). Posteriormente, diversos trabalhos foram realizados sobre modelos de regressão para esta distribuição, entre eles pode-se destacar Díaz-García e Leiva (2002) que propuseram uma extensão dos modelos anteriores, a distribuição Birnbaum-Saunders generalizada, em que distribuições de contornos elípticos podem ser utilizadas como núcleo da distribuição ao invés do modelo original que tem a Normal como núcleo gerador. Barros, Paula e Leiva (2009) desenvolveram um pacote no programa R para análise para a distribuição Birnbaum-Saunders generalizada, considerando casos de dados com ou sem censura, possibilitando uma maior aplicação da BS em dados de análise de sobrevivência.

No referente à abordagem Bayesiana, Achcar (1991) realizou inferências para a distribuição Birnbaum-Saunders, comparando resultados obtidos pela metodologia Bayesiana com os da abordagem clássica da distribuição. Análises de influência e diagnóstico Bayesianos para modelos de regressão Birnbaum-Saunders com núcleo t-Student foram propostos por Cancho, Ortega e Paula (2010). Uma análise também foi feita por Farias e Lemonte (2011), para um modelo Birnbaum-Saunders não-linear Bayesiano, com avaliação de influência e diagnóstico.

Novas parametrizações foram estudadas por Ahmed *et al* (2008), que avaliaram métodos de estimação e performance de uma reparametrização da distribuição. Na mesma linha, também pode-se citar o trabalho de Santos-Neto *et al* (2012), que avaliaram diversas reparametrizações e suas propriedades. Em relação a modelos de regressão para uma das formas reparametrizadas, Leiva *et al* (2014) propuseram uma nova forma de modelagem para a distribuição Birnbaum-Saunders.

O objetivo deste trabalho é apresentar o modelo proposto por Leiva *et al* (2014), e apresentar uma nova abordagem para o mesmo utilizando a metodologia Bayesiana, comparando ambas versões. No Capítulo 2, será apresentada a distribuição Birnbaum-Saunders na sua versão original, sendo mostradas suas principais propriedades e a demonstração de alguns resultados. No Capítulo 3, a versão reparametrizada pela média será introduzida, sendo mostradas propriedades da distribuição e formas de estimação dos parâmetros. O modelo de regressão Birnbaum-Saunders proposto por Leiva *et al* (2014) será abordado no Capítulo 4, sendo apresentado sua forma, métodos de estimação e análise de influência. No Capítulo 5, será proposta uma abordagem Bayesiana para a versão reparametrizada da distribuição e do modelo de regressão. No modelo de regressão Bayesiano também será discutido métodos de comparação

de modelos e de diagnóstico de influência. No Capítulo 6, será feito um exemplo de aplicação para os modelos de regressão Birnbaum-Saunders nas versões clássica e Bayesiana, sendo comparados ambos modelos, verificando também as semelhanças entre métodos de diagnóstico de influência sob ambas abordagens. Por último, no Capítulo 7, serão feitas as considerações finais do trabalho e propostas para trabalhos futuros.

Nos capítulos 2 a 5 serão feitas simulações dos modelos e distribuições avaliados, sendo utilizado os programas R (R Core Team, 2014) e OpenBUGS (THOMAS *et al*, 2006), ambos de código aberto e de licença livre.

2 DISTRIBUIÇÃO BIRNBAUM-SAUNDERS

A distribuição Birnbaum-Saunders (*BS*) é uma distribuição desenvolvida para modelar tempo de vida por fadiga, a qual relaciona o tempo até a ocorrência de falha devido a algum dano cumulativo, que é assumido seguir uma distribuição normal. Birnbaum e Saunders (1969a) fizeram as seguintes suposições sobre o processo de fadiga:

1. Um material é submetido a um padrão cíclico de força;
2. A falha do material ocorre devido ao desenvolvimento e ao crescimento de uma fissura dominante no material, ou seja, quando o tamanho da fissura excede certo nível de resistência, denotado por w , acontece a falha;
3. A sequência de tensão imposta ao material é a mesma ciclo a ciclo;
4. A extensão incremental da fissura X_i resultante da aplicação da i -ésima oscilação de carga é uma variável aleatória com uma distribuição que só depende da fissura atual causada pela tensão neste ciclo;
5. A extensão da fissura durante o $(j + 1)$ -ésimo ciclo é

$$Y_{j+1} = X_{jm+1} + \dots + X_{jm+i} + \dots + X_{jm+m}, \quad j = 0, 1, \dots$$

em que X_{jm+i} é a extensão incremental da fissura após a i -ésima oscilação de carga do $(j + 1)$ -ésimo ciclo;

6. As extensões das fissuras em diferentes ciclos são independentes;
7. A extensão total da fissura, Y_j , devido ao j -ésimo ciclo é uma variável aleatória que segue uma distribuição com média μ e variância σ^2 , $\forall j \in \mathbb{N}$.

Assim, a extensão total da fissura após n ciclos é dada pela variável aleatória

$$W_n = \sum_{j=1}^n Y_j.$$

Sob as suposições 6 e 7, para n grande, é possível estabelecer pelo Teorema do Limite Central, que

$$\frac{W_n - n\mu}{\sqrt{n\sigma^2}} \stackrel{a}{\sim} \mathcal{N}(0, 1),$$

em que $\stackrel{a}{\sim}$ significa assintoticamente distribuído. Outras versões para a distribuição de W_n também foram discutidas por Patriota (2012), em que foram apresentadas outras distribuições para a extensão total da fissura com base na classe de misturas de escala da distribuição Normal.

Seja N o número de ciclos necessários até que seja observado a falha, a função de distribuição acumulada (fda) é dada por

$$\mathbb{P}(N \leq n) = \mathbb{P}(W \geq w), \quad (2.1)$$

que significa que para termos até N ciclos até a falha, é preciso que a ruptura tenha atingido pelo menos o nível crítico w que causa a falha.

Padronizando os termos da equação (2.1), tem-se que

$$\begin{aligned} \mathbb{P}(N \leq n) &= \mathbb{P}\left(\frac{W_n - n\mu}{\sigma\sqrt{n}} \geq \frac{w - n\mu}{\sigma\sqrt{n}}\right) = 1 - \mathbb{P}\left(\frac{W_n - n\mu}{\sigma\sqrt{n}} < \frac{w - n\mu}{\sigma\sqrt{n}}\right) \\ &= 1 - \Phi\left(\frac{w - n\mu}{\sigma\sqrt{n}}\right) = \Phi\left[-\left(\frac{w - n\mu}{\sigma\sqrt{n}}\right)\right] = \Phi\left(\frac{n\mu - w}{\sigma\sqrt{n}}\right) \\ &= \Phi\left[\frac{\sqrt{w\mu}}{\sigma} \left(\sqrt{\frac{n}{w/\mu}} - \sqrt{\frac{w/\mu}{n}}\right)\right], \end{aligned}$$

em que $\Phi(\cdot)$ representa a função de distribuição acumulada (fda) da Normal Padrão.

De acordo com Birnbaum e Saunders(1969a), se n for substituído por uma variável real não-negativa t , então a va T é a extensão contínua de N , representando o tempo total até que ocorra uma falha. Assim, tomando

$$\alpha = \frac{\sigma}{\sqrt{\mu w}}, \quad \text{e} \quad \beta = \frac{w}{\mu},$$

a fda de T é pode ser reescrita como

$$F(t; \alpha, \beta) = P(T \leq t) = \Phi\left[\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right], \quad t > 0,$$

em que $\alpha > 0$ e $\beta > 0$. Diz-se que T segue uma distribuição Birnbaum-Saunders, com parâmetros α e β , denotando por $T \sim BS(\alpha, \beta)$.

Os parâmetros α e β são, respectivamente, parâmetros de forma e escala. O parâmetro β também se trata da mediana da distribuição Birnbaum-Saunders, propriedade essa que pode ser verificada da seguinte forma,

$$F(\beta) = \Phi\left[\frac{1}{\alpha} \left(\sqrt{\frac{\beta}{\beta}} - \sqrt{\frac{\beta}{\beta}}\right)\right] = \Phi(0) = \frac{1}{2},$$

implicando que β é a mediana da distribuição, independentemente do valor de α , como pode ser observado no gráfico da Figura 1, em que as linhas cheias demarcam as medianas das funções de distribuição. A função densidade de probabilidade da distribuição Birnbaum-Saunders é dada por,

$$f(t; \alpha, \beta) = \frac{t^{-3/2}(t + \beta)}{2\alpha\sqrt{2\pi\beta}} \exp\left\{-\frac{1}{2\alpha^2} \left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right\}, \quad \mathbb{1}(t)_{(0, \infty)},$$

cuja demonstração se encontra no Apêndice A.

Através da análise do gráfico da Figura 2, pode-se observar que à medida que o valor de α cresce, a distribuição fica mais assimétrica, e quando α tende a zero, mais simétrica em torno de β a função densidade fica.

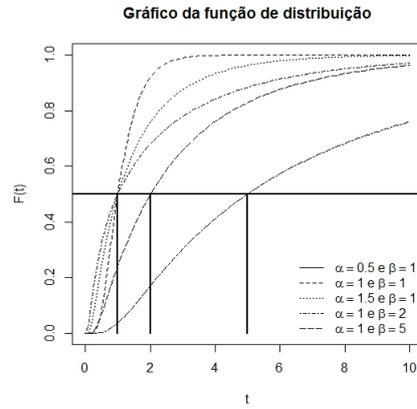


Figura 1: Gráfico da função de distribuição acumulada da distribuição $BS(\alpha, \beta)$.

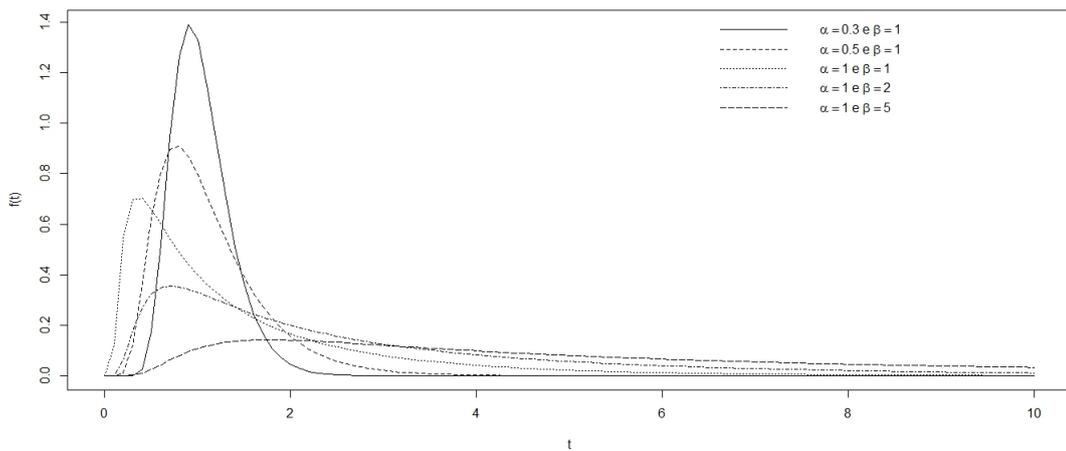


Figura 2: Gráfico da função densidade da distribuição $BS(\alpha, \beta)$.

Como a distribuição Birnbaum-Saunders foi desenvolvida para a análise de tempo de vida por fadiga, se tem interesse na probabilidade do objeto de estudo funcionar por um tempo determinado, ou seja, que a falha não ocorra nesse período. A função de sobrevivência de T é

$$S(t) = 1 - F(t) = 1 - \Phi \left[\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right], \quad t > 0. \quad (2.2)$$

A distribuição BS possui a seguinte relação estocástica com a distribuição Normal

$$T = \beta \left[\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2} \right)^2 + 1} \right]^2, \quad (2.3)$$

em que $Z \sim \mathcal{N}(0, 1)$, e cuja demonstração se encontra no Apêndice B. A partir dessa relação, tem-se também que

$$Z = \frac{1}{\alpha} \left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right).$$

Utilizando a relação entre a distribuição BS e a Normal, pode-se obter uma expressão para o q -ésimo quantil de T , dada por

$$t(q) = F^{-1}(q) = \beta \left[\frac{\alpha z(q)}{2} + \sqrt{\left(\frac{\alpha z(q)}{2} \right)^2 + 1} \right]^2, \quad 0 < q < 1,$$

em que $z(q)$ é o q -ésimo quantil da distribuição normal e $F_T^{-1}(q)$ é o inverso da função de distribuição de T aplicada em q .

A expressão do r -ésimo momento de $T \sim BS(\alpha, \beta)$ é dada por

$$E[T^r] = \beta^r \sum_{m=0}^r \binom{2r}{2m} \sum_{n=0}^m \binom{m}{n} \left(\frac{\alpha}{2} \right)^{2(r-m+n)} \frac{[2(r-m+n)]!}{2^{(r-m+n)}(r-m+n)!}, \quad (2.4)$$

cuja demonstração se encontra no Apêndice C.

Utilizando a Equação (2.4), os quatro primeiros momentos de T são

$$\begin{aligned} \mathbb{E}(T) &= \beta \left(1 + \frac{\alpha^2}{2} \right), \\ \mathbb{E}(T^2) &= \beta^2 \left[1 + 2\alpha^2 + \frac{3\alpha^4}{2} \right], \\ \mathbb{E}(T^3) &= \beta^3 \left[1 + \frac{9}{2}\alpha^2 + 9\alpha^4 + \frac{15}{2}\alpha^6 \right], \\ \mathbb{E}(T^4) &= \beta^4 \left[1 + 8\alpha^2 + 30\alpha^4 + 60\alpha^6 + \frac{105}{2}\alpha^8 \right], \end{aligned}$$

Com isso, as expressões para a esperança, variância, coeficiente de assimetria e coeficiente de curtose de T , são dados respectivamente por:

$$\begin{aligned}\mathbb{E}(T) &= \beta \left(1 + \frac{\alpha^2}{2}\right), \\ \text{Var}(T) &= \mathbb{E}[(T - \mu)^2] = (\alpha\beta)^2 \left(1 + \frac{5\alpha^2}{4}\right), \\ \alpha_3(T) &= \frac{\mathbb{E}[(T - \mu)^3]}{\sigma^3} = \frac{48\alpha + 104\alpha^3}{(4 + 5\alpha^2)^{3/2}}, \\ \alpha_4(T) &= \frac{\mathbb{E}[(T - \mu)^4]}{\sigma^4} = \frac{48 + 360\alpha^2 + 633\alpha^4}{(4 + 5\alpha^2)^2}.\end{aligned}$$

Pode-se observar que os coeficientes de assimetria e curtose da distribuição Birnbaum-Saunders não dependem dos parâmetros de escala β e são sempre positivas, sendo função somente do parâmetro de forma α .

Outras propriedades interessantes da distribuição BS , é que esta é fechada sob reciprocidade, ou seja, sua recíproca também possui distribuição Birnbaum-Saunders, alterando somente o valor dos parâmetros β ,

$$\frac{1}{T} \sim BS\left(\alpha, \frac{1}{\beta}\right),$$

e também é fechada sob proporcionalidade, em que para um número $c > 0$, novamente somente o valor de β é alterado,

$$cT \sim BS(\alpha, c\beta).$$

Essas propriedades também são mantidas em outras parametrizações da distribuição BS , sendo estes resultados provados na próxima seção, que terá como tema principal uma versão reparametrizada da BS e suas propriedades, características e conceitos.

3 DISTRIBUIÇÃO BIRNBAUM-SAUNDERS REPARAMETRIZADA

3.1 Definições

Santos-Neto *et al* (2012) propuseram uma nova parametrização da distribuição *BS*, que foi reparametrizada pela média. Tomando os parâmetros de $T \sim BS(\alpha, \beta)$ como

$$\beta = \frac{\delta\mu}{1+\delta} \quad \text{e} \quad \alpha = \sqrt{\frac{2}{\delta}},$$

os novos parâmetros da forma reparametrizada da *BS* serão

$$\mu = \beta \left(1 + \frac{\alpha^2}{2}\right) \quad \text{e} \quad \delta = \frac{2}{\alpha^2}.$$

Pode-se observar que o parâmetro μ se trata diretamente da média de T , sendo esta uma característica importante da reparametrização. A fda da distribuição *BS* dessa vez reparametrizada por μ e δ é

$$F(t; \mu, \delta) = \Phi \left[\sqrt{\frac{\delta}{2}} \left(\sqrt{\frac{(1+\delta)t}{\delta\mu}} - \sqrt{\frac{\delta\mu}{(1+\delta)t}} \right) \right], \quad t > 0, \mu > 0, \delta > 0. \quad (3.1)$$

No gráfico da Figura 3, tem-se a função de distribuição da Birnbaum-Saunders reparametrizada (BSR) para δ fixo igual a 5 na Figura 3a, e para μ fixo igual a 2 na Figura 3b.

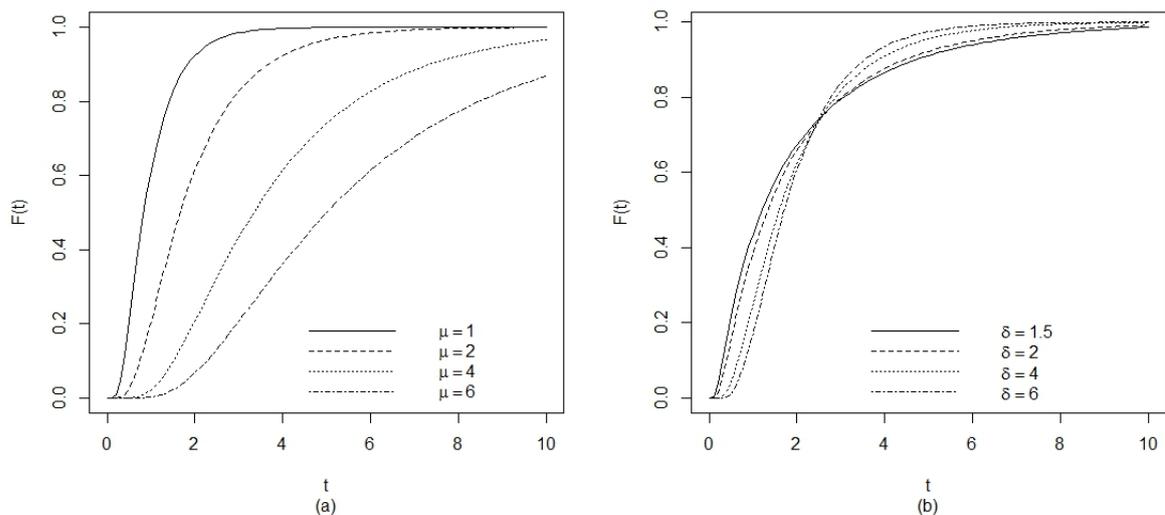


Figura 3: Gráfico da função de distribuição acumulada da BSR.

Derivando a expressão (3.1) obtém-se a fdp de $T \sim BSR(\mu, \delta)$, em que

$$\begin{aligned} \frac{dF(t; \mu, \delta)}{dt} &= \frac{d}{dt} \left\{ \Phi \left[\sqrt{\frac{\delta}{2}} \left(\sqrt{\frac{(1+\delta)t}{\delta\mu}} - \sqrt{\frac{\delta\mu}{(1+\delta)t}} \right) \right] \right\} \\ &= \left(\frac{1}{2} \sqrt{\frac{(1+\delta)}{2\mu}} \frac{1}{t^{1/2}} + \frac{1}{2} \sqrt{\frac{\delta^2\mu}{2(1+\delta)}} \frac{1}{t^{3/2}} \right) \phi \left[\sqrt{\frac{\delta}{2}} \left(\sqrt{\frac{(1+\delta)t}{\delta\mu}} - \sqrt{\frac{\delta\mu}{(1+\delta)t}} \right) \right] \\ &= \frac{1}{2} \sqrt{\frac{(1+\delta)}{2\mu}} \left(t^{-1/2} + \sqrt{\frac{2\mu}{(1+\delta)}} \sqrt{\frac{\delta^2\mu}{2(1+\delta)}} t^{-3/2} \right) \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{A}{2} \right\}, \end{aligned}$$

sendo $\phi(\cdot)$ a fdp da Normal Padrão e o termo A sendo dado por

$$\begin{aligned} A &= \left[\sqrt{\frac{\delta}{2}} \left(\sqrt{\frac{(1+\delta)t}{\delta\mu}} - \sqrt{\frac{\delta\mu}{(1+\delta)t}} \right) \right]^2 \\ &= \frac{\delta}{2} \left(\frac{(1+\delta)t}{\delta\mu} - 2 \sqrt{\frac{(1+\delta)t}{\delta\mu}} \sqrt{\frac{\delta\mu}{(1+\delta)t}} + \frac{\delta\mu}{(1+\delta)t} \right) \\ &= \frac{\delta}{2} \left(\frac{(1+\delta)t}{\delta\mu} + \frac{\delta\mu}{(1+\delta)t} - 2 \right), \end{aligned}$$

logo, a fdp de T é

$$f(t; \mu, \delta) = \frac{\exp\{\delta/2\} \sqrt{1+\delta}}{4\sqrt{\pi\mu}} t^{-3/2} \left(t + \frac{\delta\mu}{(1+\delta)} \right) \exp \left\{ -\frac{\delta}{4} \left(\frac{(1+\delta)t}{\delta\mu} + \frac{\delta\mu}{(1+\delta)t} \right) \right\}, \quad t > 0. \quad (3.2)$$

O gráfico da Figura 3.2 da fdp da BSR, foi feito com δ fixo igual a 5 na Figura 3.2a, e para μ fixo igual a 2 em 3.2b. Pode-se observar na Figura 3.2a que o parâmetro μ mesmo sendo a média da distribuição também se comporta como um parâmetro de forma, o que pode ser devido ao fato desse parâmetro ser obtido utilizando os parâmetros de forma e escala da distribuição Birnbaum-Saunders original. Na Figura 3.2b se observa que para valores pequenos de δ a distribuição BSR tem sua densidade mais afilada, destacando a influência também desse parâmetro na forma da distribuição.

3.2 Propriedades

Verificando as propriedades da BSR, tem-se que o valor esperado de T é

$$\mathbb{E}(T) = \beta \left(1 + \frac{\alpha^2}{2} \right) = \mu, \quad (3.3)$$

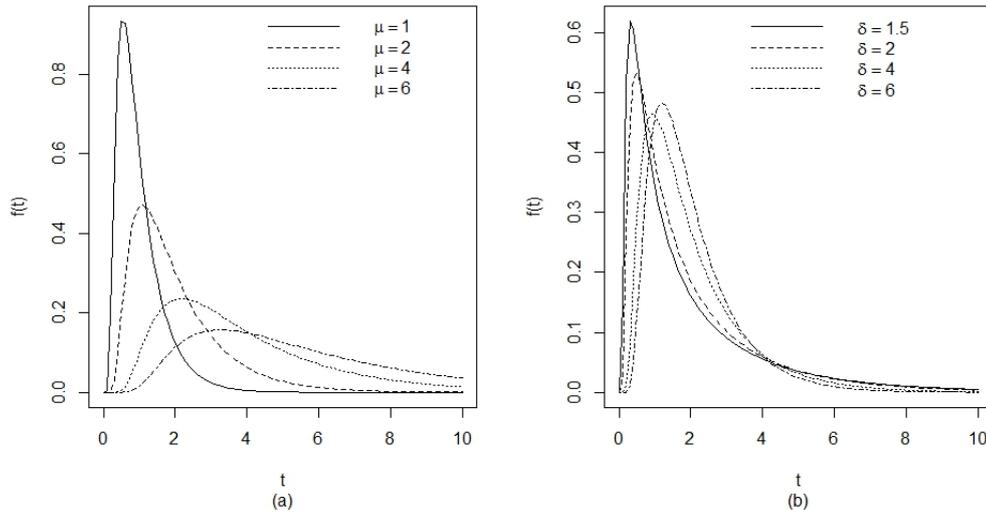


Figura 4: Gráfico da função densidade da distribuição BSR

e a variância é

$$\begin{aligned}
 \text{Var}(T) &= (\alpha\beta)^2 \left(1 + \frac{5\alpha^2}{4}\right) \\
 &= \alpha^2\beta^2 \left(\frac{2+\alpha^2}{2}\right)^2 \left(\frac{2}{2+\alpha^2}\right)^2 \left(\frac{4+5\alpha^2}{4}\right) \\
 &= \left[\beta \left(\frac{2+\alpha^2}{2}\right)\right]^2 \left(\frac{4\alpha^2+5\alpha^4}{4+4\alpha^2+\alpha^4}\right) \\
 &= \mu^2 \left(\frac{4\frac{2}{\delta} + 5\frac{4}{\delta^2}}{4 + 4\frac{2}{\delta} + \frac{4}{\delta^2}}\right) \\
 &= \mu^2 \left(\frac{2\delta+5}{\delta^2+2\delta+1}\right) \\
 &= \mu^2 \frac{(2\delta+5)}{(\delta+1)^2},
 \end{aligned}$$

logo

$$\text{Var}(T) = \mu^2 \frac{(2\delta+5)}{(\delta+1)^2}. \quad (3.4)$$

A variância também pode ser expressa de uma outra forma, formada por uma função da média e uma função do parâmetro δ , $\text{Var}(T) = V(\mu)h(\delta)$, tal que

$$V(\mu) = \mu^2 \quad \text{e} \quad h(\delta) = \frac{(2\delta+5)}{(\delta+1)^2}. \quad (3.5)$$

A distribuição (BSR) é fechada sob proporcionalidade, em que, para $a > 0$ e $Y = aT$,

tem-se que $T = \frac{Y}{a}$ e

$$F(y/a; \mu, \delta) = \mathbb{P}(T \leq y/a) = \Phi \left[\sqrt{\frac{\delta}{2}} \left(\sqrt{\frac{(1+\delta)y}{a\delta\mu}} - \sqrt{\frac{a\delta\mu}{(1+\delta)y}} \right) \right],$$

ou seja, $Y \sim BSR(a\mu, \delta)$.

Os demais momentos podem ser obtidos analogamente através dos momentos da distribuição $BS(\alpha, \beta)$, sendo por exemplo

$$\mathbb{E}[T^3] = \frac{\mu^3}{(1+\delta)^3} (\delta^3 + 9\delta^2 + 36\delta + 60),$$

e

$$\mathbb{E}[T^4] = \frac{\mu^4}{(1+\delta)^4} (\delta^4 + 16\delta^3 + 120\delta^2 + 360\delta + 840).$$

A distribuição BSR também é fechada sob reciprocidade, em que, para $X = \frac{1}{T}$, tem-se que $T = \frac{1}{X}$ e

$$\begin{aligned} F(1/x; \mu, \delta) &= \mathbb{P}(T \leq 1/x) \\ &= \Phi \left[\sqrt{\frac{\delta}{2}} \left(\sqrt{\frac{(1+\delta)x}{\delta\mu}} - \sqrt{\frac{\delta\mu}{(1+\delta)}} \right) \right] \\ &= \Phi \left[\sqrt{\frac{\delta}{2}} \left(\sqrt{\frac{\delta}{(1+\delta)x} \frac{(1+\delta)^2}{\delta^2\mu}} - \sqrt{\frac{(1+\delta)x}{\delta} \frac{\delta^2\mu}{(1+\delta)}} \right) \right], \end{aligned}$$

ou seja, $X \sim BSR\left(\frac{(\delta+1)^2}{\delta^2\mu}, \delta\right)$.

Através das relações da distribuição Birnbaum-Saunders original, a distribuição BSR possui a seguinte relação estocástica com a distribuição Normal

$$T = \frac{\delta\mu}{\delta+1} \left[\frac{Z}{\sqrt{2\delta}} + \sqrt{\left(\frac{Z}{\sqrt{2\delta}}\right)^2 + 1} \right]^2, \quad (3.6)$$

em que $Z \sim \mathcal{N}(0, 1)$, e também que

$$Z = \sqrt{\frac{\delta}{2}} \left[\sqrt{\frac{(\delta+1)T}{\mu\delta}} - \sqrt{\frac{\mu\delta}{(\delta+1)T}} \right].$$

A função quantil para a distribuição BSR é dada por

$$t(q) = \frac{\delta\mu}{\delta+1} \left[\frac{z(q)}{\sqrt{2\delta}} + \sqrt{\left(\frac{z(q)}{\sqrt{2\delta}}\right)^2 + 1} \right]^2,$$

em que $z(q)$ é o q-ésimo quantil da distribuição Normal.

3.3 Função de verossimilhança

Através de uma amostra proveniente de uma população que possa admitir uma distribuição Birnbaum-Saunders reparametrizada, é possível construir sua função de verossimilhança, a qual será útil por exemplo para obter estimativas para os parâmetros e verificar quais valores de μ e δ são mais prováveis para tal amostra.

A função de verossimilhança $L(\mu, \delta)$ da distribuição BSR é dada por

$$\begin{aligned}
 L(\mu, \delta) &= \prod_{i=1}^n f(t_i; \mu, \delta) \\
 &= \prod_{i=1}^n \frac{\exp\{\delta/2\} \sqrt{1+\delta}}{4\sqrt{\pi\mu}} t_i^{-3/2} \left(t_i + \frac{\delta\mu}{(1+\delta)} \right) \exp \left\{ -\frac{\delta}{4} \left(\frac{(1+\delta)t_i}{\delta\mu} + \frac{\delta\mu}{(1+\delta)t_i} \right) \right\} \\
 &= \frac{\exp\{n\delta/2\} (1+\delta)^{n/2}}{4^n (\pi\mu)^{n/2}} \left(\prod_{i=1}^n t_i \right)^{-3/2} \prod_{i=1}^n \left(t_i + \frac{\delta\mu}{(1+\delta)} \right) \\
 &\times \exp \left\{ -\frac{(1+\delta)}{4\mu} \sum_{i=1}^n t_i - \frac{\delta^2\mu}{4(1+\delta)} \sum_{i=1}^n \frac{1}{t_i} \right\},
 \end{aligned}$$

e com isso, o logaritmo da função de verossimilhança para uma observação é dado por

$$\begin{aligned}
 l(t_i; \mu, \delta) &= \frac{\delta}{2} + \frac{1}{2} \log(\delta + 1) - \frac{1}{2} \log(16\pi\mu) \\
 &\quad - \frac{3}{2} \log(t_i) + \log \left(t_i + \frac{\delta\mu}{\delta + 1} \right) - \frac{t_i(\delta + 1)}{4\mu} - \frac{\delta^2\mu}{4(1+\delta)t_i} \\
 &= \frac{\delta}{2} - \frac{1}{2} \log(\delta + 1) - \frac{1}{2} \log(16\pi\mu) \\
 &\quad - \frac{3}{2} \log(t_i) + \log(\delta t_i + t_i + \delta\mu) - \frac{t_i(\delta + 1)}{4\mu} - \frac{\delta^2\mu}{4(1+\delta)t_i},
 \end{aligned}$$

sendo a função de verossimilhança da amostra

$$\begin{aligned}
 l(\mu, \delta) &= \sum_{i=1}^n l(t_i; \mu, \delta) = \frac{n\delta}{2} - \frac{n}{2} \log(\delta + 1) - \frac{n}{2} \log(16\pi\mu) \\
 &\quad - \frac{3}{2} \sum_{i=1}^n \log(t_i) + \sum_{i=1}^n \log(\delta t_i + t_i + \delta\mu) - \frac{(\delta + 1)}{4\mu} \sum_{i=1}^n t_i - \frac{\delta^2\mu}{4(1+\delta)} \sum_{i=1}^n \frac{1}{t_i},
 \end{aligned}$$

Derivando $l(\mu, \delta)$ em relação a μ , tem-se que

$$\frac{\partial l(\mu, \delta)}{\partial \mu} = -\frac{n}{2\mu} + \sum_{i=1}^n \frac{\delta}{(t_i\delta + t_i + \delta\mu)} + \frac{(\delta + 1)}{4\mu^2} \sum_{i=1}^n t_i - \frac{\delta^2}{4(\delta + 1)} \sum_{i=1}^n \frac{1}{t_i}.$$

Derivando $l(\mu, \delta)$ agora em relação à δ , tem-se que

$$\begin{aligned} \frac{\partial l(\mu, \delta)}{\partial \delta} &= \frac{n}{2} - \frac{n}{2(\delta + 1)} + \sum_{i=1}^n \frac{(t_i + \mu)}{(\delta t_i + t_i + \delta \mu)} \\ &\quad - \frac{1}{4\mu} \sum_{i=1}^n t_i - \frac{(2\delta(\delta + 1) - \delta^2)\mu}{4(1 + \delta)^2} \sum_{i=1}^n \frac{1}{t_i} \\ &= \frac{n\delta}{2(\delta + 1)} + \sum_{i=1}^n \frac{(t_i + \mu)}{(\delta t_i + t_i + \delta \mu)} - \frac{1}{4\mu} \sum_{i=1}^n t_i - \frac{2\delta(\delta + 2)\mu}{4(1 + \delta)^2} \sum_{i=1}^n \frac{1}{t_i}. \end{aligned}$$

Não é possível obter os estimadores de máxima verossimilhança (EMV) de forma explícita na distribuição BSR, assim como ocorre com a forma original da distribuição em função dos parâmetros β e α .

A derivada segunda de $l(\mu, \delta)$ em relação a μ e a δ é

$$\begin{aligned} \frac{\partial^2 l(\mu, \delta)}{\partial \mu \partial \delta} &= \sum_{i=1}^n \frac{(\delta t_i + t_i + \delta \mu) - (t_i + \delta \mu)\delta}{(\delta t_i + t_i + \delta \mu)^2} + \frac{1}{4\mu^2} \sum_{i=1}^n t_i - \frac{\delta(\delta + 2)}{4(\delta + 1)^2} \sum_{i=1}^n \frac{1}{t_i} \\ &= \sum_{i=1}^n \frac{t_i}{(\delta t_i + t_i + \delta \mu)^2} + \frac{1}{4\mu^2} \sum_{i=1}^n t_i - \frac{\delta(\delta + 2)}{4(\delta + 1)^2} \sum_{i=1}^n \frac{1}{t_i}. \end{aligned}$$

As esperanças das derivadas segundas de $l(\mu, \delta)$ são

$$\begin{aligned} \mathbb{E} \left[\frac{\partial^2 l(\mu, \delta)}{\partial \mu^2} \right] &= \sum_{i=1}^n \left[-\frac{\delta}{2\mu^2} - \frac{\delta^2}{(\delta + 1)^2} \mathcal{J}(\delta) \right] \\ \mathbb{E} \left[\frac{\partial^2 l(\mu, \delta)}{\partial \delta^2} \right] &= \sum_{i=1}^n \left[-\frac{(\delta^2 + 3\delta + 1)}{2(\delta + 1)^2 \delta^2} - \frac{\mu^2}{(\delta + 1)^4} \mathcal{J}(\delta) \right] \\ \mathbb{E} \left[\frac{\partial^2 l(\mu, \delta)}{\partial \delta \partial \mu} \right] &= \sum_{i=1}^n \left[-\frac{1}{2\mu(\delta + 1)} - \frac{\delta \mu}{(\delta + 1)^3} \mathcal{J}(\delta) \right] \end{aligned}$$

em que

$$\mathcal{J}(\delta) = \int_0^\infty \frac{\exp\{\delta/2\} \sqrt{1+\delta}}{4\sqrt{\pi\mu}} t^{-3/2} \left(t + \frac{\delta\mu}{(1+\delta)} \right)^{-1} \exp \left\{ -\frac{\delta}{4} \left(\frac{(1+\delta)t}{\delta\mu} + \frac{\delta\mu}{(1+\delta)t} \right) \right\} dt.$$

As derivadas da função de verossimilhança são utilizadas para obter o vetor escore

$\mathbf{U} = [\mathbf{U}_\mu, \mathbf{U}_\delta]^\top$, e a matriz de informação de Fisher $\mathcal{J}_{\mathcal{F}}$, em que

$$\mathbf{U}_\mu = \frac{\partial l(\mu, \delta)}{\partial \mu}, \quad \mathbf{U}_\delta = \frac{\partial l(\mu, \delta)}{\partial \delta}$$

e

$$\mathcal{J}_{\mathcal{F}} = \begin{bmatrix} -\mathbb{E} \left[\frac{\partial^2 l(\mu, \delta)}{\partial \mu^2} \right] & -\mathbb{E} \left[\frac{\partial^2 l(\mu, \delta)}{\partial \mu \partial \delta} \right] \\ -\mathbb{E} \left[\frac{\partial^2 l(\mu, \delta)}{\partial \delta \partial \mu} \right] & -\mathbb{E} \left[\frac{\partial^2 l(\mu, \delta)}{\partial \delta^2} \right] \end{bmatrix}.$$

Dessa forma, é possível estimar o valor dos parâmetros μ e δ pelo método Scoring de Fisher, em que sendo $\hat{\boldsymbol{\theta}}^{[k]} = [\hat{\mu}, \hat{\delta}]^\top$ o vetor de estimadores dos parâmetros no passo k do algoritmo, novas estimativas são obtidas por

$$\hat{\boldsymbol{\theta}}^{[k+1]} = \hat{\boldsymbol{\theta}}^{[k]} + \mathcal{J}_{\mathcal{F}}^{-1} \mathbf{U}.$$

Sobre devidas condições de regularidade (ver Leite e Singer (1990)), tem-se que

$$\hat{\boldsymbol{\theta}} \stackrel{a}{\sim} \mathcal{N}(\boldsymbol{\theta}, \mathcal{J}_{\mathcal{F}}^{-1}),$$

e com esse resultado é possível construir intervalos de confiança para $\boldsymbol{\theta}$.

3.4 Estimadores pelo método dos momentos

Uma outra opção de estimação dos parâmetros da distribuição BSR é o método dos momentos (EMM), sendo os EMM's dados pela solução do sistema entre as expressões da média e do segundo momento populacional e os dois primeiros momentos amostrais.

Através das expressões (3.3) e (3.4) tem-se que

$$\mathbb{E}(T^2) = \text{Var}(T) + \mathbb{E}^2(T) = \frac{\mu^2(2\delta + 5)}{(\delta + 1)^2} + \mu^2 = \mu^2 \left[\frac{2\delta + 5}{(\delta + 1)^2} \right] = \frac{\mu^2(\delta^2 + 4\delta + 6)}{(\delta + 1)^2}.$$

Com isso, os EMM's são obtidos através da solução do seguinte sistema

$$\begin{cases} \frac{1}{n} \sum_{i=1}^n t_i = \bar{t} = \mu, \\ \frac{1}{n} \sum_{i=1}^n t_i^2 = \frac{\mu^2(2\delta+5)}{(\delta+1)^2} + \mu^2. \end{cases}$$

Logo, o EMM de μ é dado por $\tilde{\mu} = \bar{t}$, sendo então a seguinte igualdade da forma

$$\begin{aligned} \frac{\bar{t}^2(2\delta + 5)}{(\delta + 1)^2} + \bar{t}^2 &= \frac{1}{n} \sum_{i=1}^n t_i^2 \\ \frac{\bar{t}^2(2\delta + 5)}{(\delta + 1)^2} &= \frac{1}{n} \sum_{i=1}^n t_i^2 - \bar{t}^2 = s^2 \\ \frac{(2\delta + 5)}{(\delta + 1)^2} &= \frac{s^2}{\bar{t}^2} = k \\ (2\delta + 5) &= k(\delta + 1)^2 = k(\delta^2 + 2\delta + 1) = k\delta^2 + 2k\delta + k \\ 0 &= k\delta^2 - (2 - 2k)\delta - (5 - k) \\ 0 &= A\delta^2 - B\delta - C, \end{aligned}$$

em que

$$A = k = \frac{s^2}{\bar{t}^2}, \quad B = (2 - 2k) = 2 - 2\frac{s^2}{\bar{t}^2} \quad \text{e} \quad C = 5 - k = 5 - \frac{s^2}{\bar{t}^2},$$

e com isso, o EMM de δ é obtido através da solução da equação do segundo grau obtida no sistema, sendo a sua raiz

$$\tilde{\delta} = \frac{B + \sqrt{B^2 + 4AC}}{2A},$$

pois deve-se ter $\delta > 0$. Caso fosse utilizado a outra raiz da solução, teria-se um valor negativo para a estimativa de δ , o que não faz sentido para este parâmetro.

Dessa forma,

$$\begin{aligned} \tilde{\delta} &= \frac{1}{2} \frac{\bar{t}^2}{s^2} \left[2 - 2\frac{s^2}{\bar{t}^2} + \sqrt{\left(2 - 2\frac{s^2}{\bar{t}^2}\right)^2 + 4\frac{s^2}{\bar{t}^2} \left(5 - \frac{s^2}{\bar{t}^2}\right)} \right] \\ &= \frac{1}{2s^2} \left[2\bar{t}^2 - 2s^2 + \sqrt{4\bar{t}^4 + 12s^2\bar{t}^2} \right] \\ &= \frac{1}{2s^2} \left[2\bar{t}^2 - 2s^2 + 2\sqrt{\bar{t}^4 + 3s^2\bar{t}^2} \right] \\ &= \frac{\bar{t}^2 - s^2 + \sqrt{\bar{t}^4 + 3s^2\bar{t}^2}}{s^2}. \end{aligned}$$

Os estimadores pelo método dos momentos por serem mais simples de serem calculados podem ser utilizados como estimativas iniciais para a estimação no EMV quando for usado o método de Scoring de Fisher.

3.5 Estimador pelo método dos momentos modificado

Levando em consideração os problemas com os estimadores pelo método dos momentos convencional para a distribuição $BS(\alpha, \beta)$, Ng, Kundu e Balakrishna (2003) propuseram outras formas de obtenção de estimadores mais simples por um método dos momentos

modificado, que utiliza o fato da distribuição Birnbaum-Saunders ser fechada sob reciprocidade.

No caso da distribuição Birnbaum-Saunders reparametrizada, os estimadores de μ e δ pelo método dos momentos modificado são obtidos através da solução do seguinte sistema

$$\begin{cases} \bar{t} = \hat{\mu}, \\ \bar{t}_h = \frac{\hat{\delta}^2 \hat{\mu}}{(\hat{\delta}+1)^2}, \end{cases}$$

sendo \bar{t} a média amostral de uma população da variável aleatória $T \sim BSR(\mu, \delta)$, e $\bar{t}_h = \left[\frac{1}{n} \sum_{i=1}^n \frac{1}{t_i} \right]^{-1}$ a média harmônica de T , levando em consideração $1/T \sim BSR\left(\frac{(\delta+1)^2}{\delta^2 \mu}, \delta\right)$. Dessa forma, os estimadores pelo método dos momentos modificados são

$$\hat{\mu} = \bar{t} \quad \text{e} \quad \hat{\delta} = \left[\sqrt{\frac{\bar{t}}{\bar{t}_h}} - 1 \right]^{-1}.$$

Pela simplicidade da obtenção das estimativas por esse método, ele também pode ser utilizado como palpite inicial nos estimadores de máxima verossimilhança, assim como o método dos momentos convencional.

3.6 Simulações

Um estudo de simulação foi realizado para comparar as estimativas obtidas pelos diferentes métodos apresentados, método dos momentos (EMM), método dos momentos modificado (EMMod) e máxima verossimilhança (EMV). Valores aleatórios da distribuição BSR foram gerados no programa R a partir de valores gerados da distribuição Normal padrão, utilizando a relação da Equação (3.6). Foram geradas 1000 réplicas de valores de amostras com tamanhos $n \in \{10, 20, 50, 100\}$, para valores de μ iguais a 2, 4 e 6, com $\delta = 5$ fixo, e também com $\mu = 2$ fixo para os diferentes valores de δ iguais a 2, 4 e 6. Foram calculados a média, o desvio-padrão (DP) e erro quadrático médio (EQM) das estimativas.

Na Tabela 1 estão as estimativas obtidas para o parâmetro μ . Pode-se observar que para tamanhos de amostra de 10 e 20, os métodos geram estimativas com valores próximos, sendo o método com menor EQM e desvio padrão o EMV, com uma diferença maior entre os resultados para $\mu = 6$, em que o EMV tem menor EQM e desvio padrão que os outros dois, enquanto o EMM tem os maiores EQM e desvio padrão, podendo então considerar a performance do EMV melhor que os demais em tal situação. Para tamanhos de amostra maiores, as estimativas e estatísticas são próximas para os três valores de μ , sem grandes diferenças entre os três métodos.

Na Tabela 2 estão resultados obtidos para a simulação da estimativa do parâmetro δ . Para um tamanho de amostra igual a 10, os resultados obtidos com o EMM mostram maiores valores de desvio padrão e EQM obtidos entre os três métodos para os valores de δ analisados, e por outro lado, o EMMod tem os menores valores de desvio padrão e EQM, indicando que

Tabela 1: Estimativas de diferentes valores de μ com $\delta = 5$.

μ	n	EMM			EMMod			EMV		
		Média	DP	EQM	Média	DP	EQM	Média	DP	EQM
$\mu = 2$	$n = 10$	1,98	0,41	0,17	1,99	0,42	0,18	1,98	0,41	0,17
	$n = 20$	2,00	0,28	0,08	1,99	0,30	0,09	2,01	0,28	0,08
	$n = 50$	1,99	0,19	0,03	2,00	0,18	0,03	1,99	0,18	0,03
	$n = 100$	2,01	0,13	0,02	2,00	0,13	0,02	2,00	0,13	0,02
$\mu = 4$	$n = 10$	4,01	0,83	0,68	4,02	0,81	0,66	4,05	0,82	0,68
	$n = 20$	4,03	0,56	0,31	3,98	0,58	0,34	4,01	0,59	0,35
	$n = 50$	3,98	0,36	0,13	4,00	0,36	0,13	4,01	0,37	0,14
	$n = 100$	4,00	0,26	0,07	4,00	0,26	0,07	4,00	0,26	0,07
$\mu = 6$	$n = 10$	6,02	1,26	1,60	6,05	1,21	1,48	6,07	1,29	1,68
	$n = 20$	6,01	0,88	0,78	5,97	0,87	0,76	5,98	0,85	0,72
	$n = 50$	6,00	0,55	0,30	6,00	0,54	0,30	5,98	0,52	0,27
	$n = 100$	6,00	0,37	0,14	6,01	0,38	0,15	6,00	0,39	0,15

Tabela 2: Estimativas de diferentes valores de δ com $\mu = 2$.

δ	n	EMM			EMMod			EMV		
		Média	DP	EQM	Média	DP	EQM	Média	DP	EQM
$\delta = 2$	$n = 10$	3,72	2,30	8,24	2,85	1,71	3,63	2,92	1,83	4,21
	$n = 20$	2,85	1,29	2,38	2,41	0,85	0,90	2,38	0,84	0,86
	$n = 50$	2,29	0,74	0,63	2,14	0,46	0,23	2,15	0,45	0,23
	$n = 100$	2,17	0,53	0,31	2,05	0,29	0,08	2,06	0,31	0,10
$\delta = 4$	$n = 10$	6,77	4,53	28,17	5,90	3,59	16,45	5,65	3,18	12,81
	$n = 20$	5,24	2,19	6,33	4,78	1,75	3,67	4,83	1,80	3,92
	$n = 50$	4,50	1,23	1,77	4,29	0,94	0,97	4,27	0,90	0,88
	$n = 100$	4,23	0,87	0,80	4,15	0,63	0,42	4,12	0,60	0,37
$\delta = 6$	$n = 10$	9,96	7,09	66,03	8,52	5,17	33,08	8,86	5,69	40,59
	$n = 20$	7,68	3,18	12,92	7,05	2,56	7,67	7,00	2,50	7,22
	$n = 50$	6,66	1,66	3,18	6,34	1,35	1,93	6,43	1,39	2,11
	$n = 100$	6,33	1,18	1,50	6,16	0,86	0,77	6,20	0,93	0,90

nesta situação este método pode ser mais eficaz que os demais. Para os demais tamanhos de amostra, em geral tem-se que os resultados do EMM são piores que dos outros métodos, e o EMMod e EMV apresentam comportamento semelhante, sendo seus valores de desvio padrão e EQM próximos para n igual a 20 e 100, e com o EMV obtendo melhores resultados para n igual a 50.

Nesta capítulo foi apresentada a distribuição RBS, com suas características e métodos para análise de dados. Porém, em alguns casos o interesse pode ser de verificar a influência de outros fatores em uma variável estuda através da distribuição RBS. Isso é possível através de modelos de regressão para a distribuição RBS, que serão abordados no próximo capítulo.

4 MODELOS DE REGRESSÃO BIRNBAUM-SAUNDERS

4.1 Definições do modelo

Em uma situação na qual se possui uma amostra com uma variável de interesse e variáveis explicativas (covariáveis) associadas, em que a distribuição BSR possa ser utilizada para análise da variável resposta, o efeito das covariáveis na resposta pode ser estudado através de modelos de regressão Birnbaum-Saunders.

Considere uma amostra aleatória T_1, \dots, T_n de uma população em que cada indivíduo tem distribuição $BSR(\mu_i, \delta)$, $i = 1, \dots, n$, a fdp das variáveis é dada por

$$f(t_i; \mu_i, \delta) = \frac{\exp\{\delta/2\}\sqrt{1+\delta}}{4\sqrt{\pi\mu_i}} t_i^{-3/2} \left(t_i + \frac{\delta\mu_i}{(1+\delta)} \right) \exp \left\{ -\frac{\delta}{4} \left(\frac{(1+\delta)t_i}{\delta\mu_i} + \frac{\delta\mu_i}{(1+\delta)t_i} \right) \right\}.$$

Como mostrado anteriormente, essa nova reparametrização possui um parâmetro que é a própria média da variável, em que $\mathbb{E}(T_i) = \mu_i$, e que $\text{Var}(T_i) = V(\mu_i)/h(\delta)$, com $V(\mu_i)$ sendo uma função da média e $h(\delta)$ uma função do parâmetro δ , considerado parâmetro de dispersão e constante para todas as observações. No modelo de regressão Birnbaum-Saunders (MRBS), proposto por Leiva *et al* (2014), a forma reparametrizada da distribuição Birnbaum-Saunders é utilizada para construir um modelo de regressão semelhante à abordagem feita em modelos lineares generalizados (MLG), no qual trabalha-se com uma função da média e com o seguinte componente sistemático

$$g(\mu_i) = \mathbf{x}_i^\top \boldsymbol{\beta} = \eta_i,$$

em que $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^\top$ é um vetor de variáveis explicativas, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ é um vetor de parâmetros a serem estimados e η_i é o preditor linear.

Como nesse caso uma função da média é que possui relação linear com as variáveis explicativas e o vetor de parâmetros, tem-se que o valor esperado da variável resposta é dado por $\mu = g^{-1}(\mathbf{x}_i^\top \boldsymbol{\beta})$, sendo $g: \mathbb{R} \rightarrow \mathbb{R}^+$ uma função de ligação estritamente monótona e pelo menos duas vezes diferenciável. O logaritmo da função verossimilhança do modelo para o vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \delta)^\top$ tem a seguinte forma

$$l(\boldsymbol{\theta}) = \log \left[\prod_{i=1}^n f(\mu_i, \delta) \right] = \sum_{i=1}^n l_i(t_i; \mu_i, \delta).$$

Com o intuito de simplificar os cálculos, pode-se trabalhar com o logaritmo da

função verossimilhança para uma observação, dada por

$$\begin{aligned}
l_i(\mu_i, \delta) &= \frac{\delta}{2} + \frac{1}{2} \log(\delta + 1) - \frac{1}{2} \log(16\pi\mu_i) \\
&\quad - \frac{3}{2} \log(t_i) + \log\left(t_i + \frac{\delta\mu_i}{\delta + 1}\right) - \frac{t_i(\delta + 1)}{4\mu_i} - \frac{\delta^2\mu_i}{4(1 + \delta)t_i} \\
&= \frac{\delta}{2} - \frac{1}{2} \log(\delta + 1) - \frac{1}{2} \log(16\pi\mu_i) \\
&\quad - \frac{3}{2} \log(t_i) + \log(\delta t_i + t_i + \delta\mu_i) - \frac{t_i(\delta + 1)}{4\mu_i} - \frac{\delta^2\mu_i}{4(1 + \delta)t_i}.
\end{aligned}$$

4.2 Derivadas da função do logaritmo da função verossimilhança

Assim como apresentado em Leiva *et al* (2014), os estimadores de máxima verossimilhança do MRBS, podem ser obtidos derivando $l(\boldsymbol{\theta})$ em relação aos parâmetros componentes de $\boldsymbol{\theta}$.

Sendo a derivada em relação a μ_i do logaritmo da função verossimilhança para uma observação dada por

$$\frac{\partial l_i(\mu_i, \delta)}{\partial \mu_i} = -\frac{1}{2\mu_i} + \frac{\delta}{(\delta t_i + t_i + \delta\mu_i)} + \frac{t_i(\delta + 1)}{4\mu_i^2} - \frac{\delta^2}{4t_i(\delta + 1)}, \quad (4.1)$$

a derivada da função de $l_i(\mu_i, \delta)$ para uma observação em relação a um parâmetro β_j , é dada por

$$\begin{aligned}
\frac{\partial l_i(\mu_i, \delta)}{\partial \beta_j} &= \frac{\partial l_i(t_i; \mu_i, \delta)}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j} \\
&= \left[-\frac{1}{2\mu_i} + \frac{\delta}{(\delta t_i + t_i + \delta\mu_i)} + \frac{t_i(\delta + 1)}{4\mu_i^2} - \frac{\delta^2}{4t_i(\delta + 1)} \right] \frac{1}{\partial \eta_i / \partial \mu_i} x_{ij} \\
&= \left[-\frac{1}{2\mu_i} + \frac{\delta}{(\delta t_i + t_i + \delta\mu_i)} + \frac{t_i(\delta + 1)}{4\mu_i^2} - \frac{\delta^2}{4t_i(\delta + 1)} \right] \frac{1}{g'(\mu_i)} x_{ij}.
\end{aligned}$$

Dessa forma, os componentes do vetor escore para o vetor de parâmetros $\boldsymbol{\beta}$ serão

$$\mathbf{U}_{\beta_j}(\boldsymbol{\theta}) = \sum_{i=1}^n \left[-\frac{1}{2\mu_i} + \frac{\delta}{(\delta t_i + t_i + \delta\mu_i)} + \frac{t_i(\delta + 1)}{4\mu_i^2} - \frac{\delta^2}{4t_i(\delta + 1)} \right] \frac{1}{g'(\mu_i)} x_{ij}, \quad j = 1, \dots, p.$$

Seja $\mathbf{D}(\mathbf{a})_{n \times n} = \text{diag}\{a_1, \dots, a_n\}$, com

$$a_i = \frac{1}{g'(\mu_i)}, \quad (4.2)$$

e $\mathbf{z} = (z_1, \dots, z_n)^\top$, com

$$z_i = \frac{-1}{2\mu_i} + \frac{\delta}{\delta t_i + t_i + \delta\mu_i} + \frac{t_i(\delta + 1)}{4\mu_i^2} - \frac{\delta^2}{4t_i(\delta + 1)}, \quad (4.3)$$

e \mathbf{X} uma matriz de especificação do modelo de dimensão $(n \times p)$ onde a i -ésima linha é o vetor da covariável \mathbf{x}_i^\top , a forma matricial da função escore de $\boldsymbol{\beta}$ é dada por

$$\mathbf{U}_{\boldsymbol{\beta}}(\boldsymbol{\theta})_{p \times 1} = \mathbf{X}^\top \mathbf{D}(\mathbf{a})\mathbf{z}.$$

A derivada de $l_i(\mu_i, \delta)$ em relação à δ é

$$\frac{\partial l_i(\mu_i, \delta)}{\partial \delta} = \frac{\delta}{2(\delta + 1)} + \frac{t_i + \mu_i}{\delta t_i + t_i + \delta\mu_i} - \frac{t_i}{4\mu_i} - \frac{\delta(\delta + 2)\mu_i}{4t_i(\delta + 1)^2},$$

logo a função escore de δ é dada por

$$\mathbf{U}_{\delta}(\boldsymbol{\theta}) = \sum_{i=1}^n \frac{\partial l_i(\mu_i, \delta)}{\partial \delta}.$$

Seja $\mathbf{D}(\mathbf{b}) = \text{diag}\{b_1, \dots, b_n\}$, com

$$b_i = \frac{\delta}{2(\delta + 1)} + \frac{t_i + \mu_i}{\delta t_i + t_i + \delta\mu_i} - \frac{t_i}{4\mu_i} - \frac{\delta(\delta + 2)\mu_i}{4t_i(\delta + 1)^2}, \quad (4.4)$$

então podemos reescrever a função escore como

$$\mathbf{U}_{\delta}(\boldsymbol{\theta})_{1 \times 1} = \text{tr}[\mathbf{D}(\mathbf{b})].$$

Tomando a segunda derivada de $l(\boldsymbol{\theta})$ em relação à β_j e β_l tem-se que

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_l} &= \sum_{i=1}^n \left[\frac{\partial^2 l_i(\mu_i, \delta)}{\partial \mu_i^2} \frac{\partial \mu_i}{\partial \eta_i} + \frac{\partial l_i(\mu_i, \delta)}{\partial \mu_i} \frac{\partial}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \right] \frac{\partial \mu_i}{\partial \eta_i} x_{ij} x_{il} \\ &= \sum_{i=1}^n \left[\frac{\partial^2 l_i(\mu_i, \delta)}{\partial \mu_i^2} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 + \frac{\partial l_i(\mu_i, \delta)}{\partial \mu_i} \left(\frac{\partial}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \right) \frac{\partial \mu_i}{\partial \eta_i} \right] x_{ij} x_{il}, \end{aligned}$$

sendo

$$\frac{\partial^2 l_i(\mu_i, \delta)}{\partial \mu_i^2} = \frac{1}{2\mu_i^2} - \frac{\delta^2}{(\delta t_i + t_i + \delta\mu_i)^2} - \frac{t_i(\delta + 1)}{2\mu_i^3}. \quad (4.5)$$

Seja $\mathbf{D}(\mathbf{c}) = \text{diag}(c_1, \dots, c_n)$, com

$$c_i = \frac{\partial^2 l_i(\mu_i, \delta)}{\partial \mu_i^2} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 + \frac{\partial l_i(\mu_i, \delta)}{\partial \mu_i} \left(\frac{\partial}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \right) \frac{\partial \mu_i}{\partial \eta_i}, \quad (4.6)$$

então na forma matricial tem-se que a matriz hessiana para o vetor $\boldsymbol{\beta}$ será

$$\ddot{\mathbf{L}}_{\boldsymbol{\beta}\boldsymbol{\beta}} = \mathbf{X}^\top \mathbf{D}(\mathbf{c}) \mathbf{X}.$$

A derivada segunda de $l(\boldsymbol{\theta})$ com respeito a β_j e a δ é dada por

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_j \partial \delta} = \sum_{i=1}^n \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu_i \partial \delta} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j},$$

mas sendo

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mu_i \partial \delta} &= \frac{(\delta t_i + t_i + \delta \mu_i) - (t_i + \delta \mu_i) \delta}{(\delta t_i + t_i + \delta \mu_i)^2} + \frac{t_i}{4\mu_i^2} - \frac{\delta(\delta + 2)}{4(\delta + 1)^2 t_i} \\ &= \frac{t_i}{(\delta t_i + t_i + \delta \mu_i)^2} + \frac{t_i}{4\mu_i^2} - \frac{\delta(\delta + 2)}{4(\delta + 1)^2 t_i}, \end{aligned}$$

com esse resultado tem-se que

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_j \partial \delta} = \sum_{i=1}^n \left[\frac{t_i}{(\delta t_i + t_i + \delta \mu_i)^2} + \frac{t_i}{4\mu_i^2} - \frac{\delta(\delta + 2)}{4(\delta + 1)^2 t_i} \right] \frac{1}{g'(\mu_i)} x_{ij}.$$

Seja $\mathbf{m} = (m_1, \dots, m_n)^\top$, com

$$m_i = \frac{t_i}{(\delta t_i + t_i + \delta \mu_i)^2} + \frac{t_i}{4\mu_i^2} - \frac{\delta(\delta + 2)}{4(\delta + 1)^2 t_i},$$

então a forma matricial pode ser escrita como

$$\ddot{\mathbf{L}}_{\boldsymbol{\beta}\boldsymbol{\delta}} = \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{m}.$$

A derivada de segunda ordem de $l(\boldsymbol{\theta})$ com respeito a δ é dada por

$$\sum_{i=1}^n \frac{\partial^2 l(\mu_i, \delta)}{\partial \delta^2} = \sum_{i=1}^n \left[\frac{1}{2(\delta + 1)^2} - \frac{(t_i + \mu_i)^2}{(\delta t_i + t_i + \delta \mu_i)^2} - \frac{\mu_i}{2t_i(\delta + 1)^3} \right].$$

Seja $\mathbf{D}(\mathbf{d}) = \text{diag}(d_1, \dots, d_n)$, com

$$d_i = \frac{1}{2(\delta + 1)^2} - \frac{(t_i + \mu_i)^2}{(\delta t_i + t_i + \delta \mu_i)^2} - \frac{\mu_i}{2t_i(\delta + 1)^3},$$

a forma matricial será

$$\ddot{\mathbf{L}}_{\boldsymbol{\delta}\boldsymbol{\delta}} = \text{tr}[\mathbf{D}(\mathbf{d})].$$

Com isso a matriz hessiana do modelo, que também é a matriz observada de Fisher,

é dada por

$$\dot{\mathbf{L}}_{\theta\theta} = \begin{bmatrix} \dot{\mathbf{L}}_{\beta\beta} & \dot{\mathbf{L}}_{\beta\delta} \\ \dot{\mathbf{L}}_{\beta\delta} & \dot{\mathbf{L}}_{\delta\delta} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^\top \mathbf{D}(\mathbf{c}) \mathbf{X} & \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{m} \\ \mathbf{m}^\top \mathbf{D}(\mathbf{a}) \mathbf{X} & \text{tr}[\mathbf{D}(\mathbf{d})] \end{bmatrix}.$$

A inversa de $\dot{\mathbf{L}}_{\theta\theta}$ é dada por (ver Magnus e Neudecker (2007), por exemplo)

$$\dot{\mathbf{L}}_{\theta\theta}^{-1} = \begin{bmatrix} \mathbf{M}^{-1} + \mathbf{A} \mathbf{E}^{-1} \mathbf{A}^\top & -\mathbf{A} \mathbf{E}^{-1} \\ \mathbf{E}^{-1} \mathbf{A}^\top & \mathbf{E}^{-1} \end{bmatrix},$$

em que

$$\begin{aligned} \mathbf{E} &= \text{tr}[\mathbf{D}(\mathbf{d})] - \mathbf{m}^\top \mathbf{D}(\mathbf{a}) \mathbf{X} [\mathbf{X}^\top \mathbf{D}(\mathbf{c}) \mathbf{X}]^{-1} \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{m} \\ &= \text{tr}[\mathbf{D}(\mathbf{d})] - \mathbf{m}^\top \mathbf{D}(\mathbf{a}) \mathbf{X} \mathbf{M}^{-1} \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{m} \\ &= \text{tr}[\mathbf{D}(\mathbf{d})] - \mathbf{m}^\top \mathbf{D}(\mathbf{a}) \mathbf{A}, \end{aligned}$$

e $\mathbf{M} = \mathbf{X}^\top \mathbf{D}(\mathbf{c}) \mathbf{X}$ e $\mathbf{A} = \mathbf{M}^{-1} \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{m}$.

4.3 Matriz de informação esperada de Fisher

A matriz de informação esperada de Fisher é obtida a partir da esperança da matriz observada $\dot{\mathbf{L}}_{\theta\theta}$, sendo os seguintes resultados úteis nos cálculos (SANTOS-NETO *et al*, 2012):

- $\mathbb{E}[T] = \mu,$
- $\mathbb{E}\left[\frac{1}{T}\right] = \frac{(\delta+1)^2}{\mu\delta^2},$
- $\mathbb{E}\left[\left(T + \frac{\mu\delta}{(\delta+1)}\right)^{-1}\right] = \frac{(\delta+1)}{2\mu\delta},$
- $\mathbb{E}\left[\left(T + \frac{\mu\delta}{(\delta+1)}\right)^{-2}\right] = \mathcal{J}(\delta),$

sendo

$$\mathcal{J}(\delta) = \int_0^\infty \frac{\exp\{\delta/2\} \sqrt{1+\delta}}{4\sqrt{\pi\mu}} t^{-3/2} \left(t + \frac{\delta\mu}{(1+\delta)}\right)^{-1} \exp\left\{-\frac{\delta}{4} \left(\frac{(1+\delta)t}{\delta\mu} + \frac{\delta\mu}{(1+\delta)t}\right)\right\} dt.$$

Os dois primeiros resultados citados tratam da esperança de uma variável aleatória $T \sim BS(\mu, \delta)$ e da esperança da sua recíproca.

O terceiro resultado pode ser obtido observando que, sob certas condições de regularidade, a esperança da função escore é igual a zero, como utilizado em Lemonte (2006), ou

seja,

$$\begin{aligned}
\mathbb{E} \left[\frac{\partial l_i(\mu_i, \delta)}{\partial \mu_i} \right] &= 0 \\
\mathbb{E} \left[-\frac{1}{2\mu_i} + \frac{\delta}{(\delta t_i + t_i + \delta \mu_i)} + \frac{t_i(\delta + 1)}{4\mu_i^2} - \frac{\delta^2}{4t_i(\delta + 1)} \right] &= 0 \\
-\frac{1}{2\mu_i} + \delta \mathbb{E} \left[\frac{1}{(\delta t_i + t_i + \delta \mu_i)} \right] + \frac{(\delta + 1)}{4\mu_i^2} \mathbb{E}[t_i] - \frac{\delta^2}{4(\delta + 1)} \mathbb{E} \left[\frac{1}{t_i} \right] &= 0 \\
-\frac{1}{2\mu_i} + \frac{\delta}{(\delta + 1)} \mathbb{E} \left[\left(t_i + \frac{\delta \mu_i}{(\delta + 1)} \right)^{-1} \right] + \frac{(\delta + 1)}{4\mu_i^2} \mu_i - \frac{\delta^2}{4(\delta + 1)} \frac{(\delta + 1)^2}{\mu_i \delta^2} &= 0 \\
\frac{\delta}{(\delta + 1)} \mathbb{E} \left[\left(t_i + \frac{\delta \mu_i}{(\delta + 1)} \right)^{-1} \right] &= \frac{1}{2\mu_i} \\
\mathbb{E} \left[\left(t_i + \frac{\delta \mu_i}{(\delta + 1)} \right)^{-1} \right] &= \frac{(\delta + 1)}{2\delta \mu_i}.
\end{aligned}$$

Tomando a esperança de $\dot{\mathbf{L}}_{\beta\beta}$, tem-se

$$\mathbb{E} [\dot{\mathbf{L}}_{\beta\beta}] = \mathbb{E} [\mathbf{X}^\top \mathbf{D}(\mathbf{c}) \mathbf{X}] = \mathbf{X}^\top \mathbb{E} [\mathbf{D}(\mathbf{c})] \mathbf{X},$$

em que $\mathbb{E} [\mathbf{D}(\mathbf{c})] = \text{diag}\{\mathbb{E}[c_1], \dots, \mathbb{E}[c_n]\}$. Assim, a esperança de um termo genérico $c_i, i = 1, \dots, n$ da matriz é dado por

$$\begin{aligned}
\mathbb{E}[c_i] &= \mathbb{E} \left[\frac{\partial^2 l_i(\mu_i, \delta)}{\partial \mu_i^2} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 + \frac{\partial l_i(\mu_i, \delta)}{\partial \mu_i} \left(\frac{\partial}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \right) \frac{\partial \mu_i}{\partial \eta_i} \right] \\
&= \mathbb{E} \left[\frac{\partial^2 l_i(\mu_i, \delta)}{\partial \mu_i^2} \right] \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2 + \mathbb{E} \left[\frac{\partial l_i(\mu_i, \delta)}{\partial \mu_i} \right] \left(\frac{\partial}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \right) \frac{\partial \mu_i}{\partial \eta_i}
\end{aligned}$$

Porém, como sob certas condições de regularidade (ver Leite e Singer (1990)) a derivada da função score é igual a zero, então

$$\mathbb{E} \left[\frac{\partial l_i(\mu_i, \delta)}{\partial \mu_i} \right] = 0,$$

e por (4.5), tem-se que

$$\begin{aligned}
\mathbb{E}[C_i] &= \mathbb{E} \left[\frac{1}{2\mu_i^2} - \frac{\delta^2}{(\delta t_i + t_i + \delta\mu_i)^2} - \frac{t_i(\delta+1)}{2\mu_i^3} \right] \left(\frac{1}{g'(\mu_i)} \right)^2 \\
&= \left[\frac{1}{2} \frac{1}{\mu_i^2} - \frac{\delta^2}{(\delta+1)^2} \mathbb{E} \left[\frac{1}{\left(t_i + \frac{\delta\mu_i}{\delta+1}\right)^2} \right] - \frac{(\delta+1)}{2\mu_i^3} \mathbb{E}[t_i] \right] \left(\frac{1}{g'(\mu_i)} \right)^2 \\
&= \left[\frac{1}{2} \frac{1}{\mu_i^2} - \frac{\delta^2}{(\delta+1)^2} \mathbb{E} \left[\left(t_i + \frac{\delta\mu_i}{\delta+1}\right)^{-2} \right] - \frac{(\delta+1)\mu_i}{2\mu_i^3} \right] \left(\frac{1}{g'(\mu_i)} \right)^2 \\
&= -\frac{\delta}{2\mu_i^2} \frac{1}{[g'(\mu_i)]^2} - \frac{\delta^2}{(\delta+1)^2} \frac{1}{[g'(\mu_i)]^2} \mathcal{J}(\delta).
\end{aligned}$$

Seja agora

$$v_i = -\mathbb{E}[C_i] = \frac{\delta}{2\mu_i^2} \frac{1}{[g'(\mu_i)]^2} + \frac{\delta^2}{(\delta+1)^2} \frac{1}{[g'(\mu_i)]^2} \mathcal{J}(\delta),$$

em forma matricial tem-se que $\mathbf{D}(\mathbf{v}) = \text{diag}\{v_1, \dots, v_n\}$ e

$$\mathbf{K}_{\beta\beta} = -\mathbb{E}[\ddot{\mathbf{L}}_{\beta\beta}] = -\mathbf{X}^\top \mathbb{E}[\mathbf{D}(\mathbf{c})] \mathbf{X} = \mathbf{X}^\top \mathbf{D}(\mathbf{v}) \mathbf{X}.$$

Tomando a esperança de $\ddot{\mathbf{L}}_{\delta\delta}$, tem-se

$$\mathbb{E}[\ddot{\mathbf{L}}_{\delta\delta}] = \mathbb{E}[\text{tr}[\mathbf{D}(\mathbf{d})]] = \text{tr}[\mathbb{E}[\mathbf{D}(\mathbf{d})]],$$

em que $\mathbb{E}[\mathbf{D}(\mathbf{d})] = \text{diag}\{\mathbb{E}[d_1], \dots, \mathbb{E}[d_n]\}$. Assim, a esperança de um termo genérico $d_i, i = 1, \dots, n$ da matriz é dado por

$$\begin{aligned}
\mathbb{E}[d_i] &= \mathbb{E} \left[\frac{\partial^2 l(\mu_i, \delta)}{\partial \delta^2} \right] \\
&= \mathbb{E} \left[\frac{1}{2(\delta+1)^2} - \frac{(t_i + \mu_i)^2}{(\delta t_i + t_i + \delta\mu_i)^2} - \frac{\mu_i}{2t_i(\delta+1)^3} \right] \\
&= \frac{1}{2(\delta+1)^2} - \mathbb{E} \left[\frac{t_i^2 + 2t_i\mu_i + \mu_i^2}{(\delta t_i + t_i + \delta\mu_i)^2} \right] - \frac{\mu_i}{2(\delta+1)^3} \mathbb{E} \left[\frac{1}{t_i} \right] \\
&= \frac{1}{2(\delta+1)^2} - \mathbb{E} \left[\frac{t_i^2}{(\delta t_i + t_i + \delta\mu_i)^2} \right] - 2\mu_i \mathbb{E} \left[\frac{t_i}{(\delta t_i + t_i + \delta\mu_i)^2} \right] \\
&\quad - \mu_i^2 \mathbb{E} \left[\frac{1}{(\delta t_i + t_i + \delta\mu_i)^2} \right] - \frac{\mu_i}{2(\delta+1)^3} \frac{(\delta+1)^2}{\mu \delta^2} \\
&= \frac{\delta^2 - \delta - 1}{2(\delta+1)^2 \delta^2} - \frac{1}{(\delta+1)^2} \mathbb{E} \left[\frac{t_i^2}{\left(t_i + \frac{\delta\mu_i}{\delta+1}\right)^2} \right] - \frac{2\mu_i}{(\delta+1)^2} \mathbb{E} \left[\frac{t_i}{\left(t_i + \frac{\delta\mu_i}{\delta+1}\right)^2} \right] \\
&\quad - \frac{\mu_i^2}{(\delta+1)^2} \mathbb{E} \left[\frac{1}{\left(t_i + \frac{\delta\mu_i}{\delta+1}\right)^2} \right]
\end{aligned}$$

Agora, utilizando frações parciais, como realizado em Lemonte (2006), tem-se os seguintes resultados

$$\mathbb{E} \left[\frac{T}{(T+\theta)^2} \right] = \mathbb{E} \left[\frac{1}{T+\theta} \left(1 - \frac{\theta}{T+\theta} \right) \right] = \mathbb{E} \left[\frac{1}{T+\theta} \right] - \theta \mathbb{E} \left[\frac{1}{(T+\theta)^2} \right],$$

e também que

$$\begin{aligned} \mathbb{E} \left[\frac{T^2}{(T+\theta)^2} \right] &= \mathbb{E} \left[\frac{T}{T+\theta} \left(1 - \frac{\theta}{T+\theta} \right) \right] \\ &= \mathbb{E} \left[\frac{T}{T+\theta} \right] - \theta \mathbb{E} \left[\frac{T}{(T+\theta)^2} \right] \\ &= \mathbb{E} \left[1 - \frac{\theta}{T+\theta} \right] - \theta \left(\mathbb{E} \left[\frac{1}{T+\theta} \right] - \theta \mathbb{E} \left[\frac{1}{(T+\theta)^2} \right] \right) \\ &= 1 - \theta \mathbb{E} \left[\frac{1}{T+\theta} \right] - \theta \mathbb{E} \left[\frac{1}{T+\theta} \right] + \theta^2 \mathbb{E} \left[\frac{1}{(T+\theta)^2} \right] \\ &= 1 - 2\theta \mathbb{E} \left[\frac{1}{T+\theta} \right] + \theta^2 \mathbb{E} \left[\frac{1}{(T+\theta)^2} \right]. \end{aligned}$$

Com isso,

$$\begin{aligned} \mathbb{E} \left[\frac{t_i^2}{(t_i + \frac{\delta\mu_i}{\delta+1})^2} \right] &= 1 - 2 \frac{\delta\mu_i}{(\delta+1)} \mathbb{E} \left[\frac{1}{(t_i + \frac{\delta\mu_i}{\delta+1})} \right] + \frac{\delta^2\mu_i^2}{(\delta+1)^2} \mathbb{E} \left[\frac{1}{(t_i + \frac{\delta\mu_i}{\delta+1})^2} \right] \\ &= 1 - 2 \frac{\delta\mu_i}{(\delta+1)} \frac{(\delta+1)}{2\mu_i\delta} + \frac{\delta^2\mu_i^2}{(\delta+1)^2} \mathcal{J}(\delta) \\ &= \frac{\delta^2\mu_i^2}{(\delta+1)^2} \mathcal{J}(\delta), \end{aligned}$$

e

$$\begin{aligned} \mathbb{E} \left[\frac{t_i}{(t_i + \frac{\delta\mu_i}{\delta+1})^2} \right] &= \mathbb{E} \left[\frac{1}{(t_i + \frac{\delta\mu_i}{\delta+1})} \right] - \frac{\delta\mu_i}{(\delta+1)} \mathbb{E} \left[\frac{1}{(t_i + \frac{\delta\mu_i}{\delta+1})^2} \right] \\ &= \frac{(\delta+1)}{2\mu_i\delta} - \frac{\delta\mu_i}{(\delta+1)} \mathcal{J}(\delta). \end{aligned}$$

Portanto,

$$\begin{aligned}
\mathbb{E}[d_i] &= \frac{\delta^2 - \delta - 1}{2(\delta + 1)^2 \delta^2} - \frac{\mu_i^2}{(\delta + 1)^2} \mathcal{J}(\delta) - \frac{2\mu_i}{(\delta + 1)^2} \left[\frac{(\delta + 1)}{2\mu_i \delta} - \frac{\delta \mu_i}{(\delta + 1)} \mathcal{J}(\delta) \right] \\
&- \frac{1}{(\delta + 1)^2} \frac{\delta^2 \mu_i^2}{(\delta + 1)^2} \mathcal{J}(\delta) \\
&= \frac{\delta^2 - \delta - 1}{2(\delta + 1)^2 \delta^2} - \frac{1}{(\delta + 1)\delta} + \frac{\mu_i^2}{(\delta + 1)^2} \mathcal{J}(\delta) + \frac{2\delta \mu_i^2}{(\delta + 1)^3} \mathcal{J}(\delta) - \frac{\delta^2 \mu_i^2}{(\delta + 1)^4} \mathcal{J}(\delta) \\
&= \frac{\delta^2 - \delta - 1 - 2(\delta + 1)\delta}{2(\delta + 1)^2 \delta^2} - \left[\frac{\delta^2}{(\delta + 1)^2} - \frac{2\delta}{(\delta + 1)} + 1 \right] \mathcal{J}(\delta) \frac{\mu_i^2}{(\delta + 1)^2} \\
&= \frac{-(\delta^2 + 3\delta + 1)}{2(\delta + 1)^2 \delta^2} - \left[\frac{\delta^2 - 2\delta^2 - 2\delta + \delta^2 + 2\delta + 1}{(\delta + 1)} \right] \mathcal{J}(\delta) \frac{\mu_i^2}{(\delta + 1)^2} \\
&= -\frac{(\delta^2 + 3\delta + 1)}{2(\delta + 1)^2 \delta^2} - \mathcal{J}(\delta) \frac{\mu_i^2}{(\delta + 1)^4}
\end{aligned}$$

Seja

$$u_i = -\mathbb{E}[d_i] = \frac{(\delta^2 + 3\delta + 1)}{2(\delta + 1)^2 \delta^2} + \mathcal{J}(\delta) \frac{\mu_i^2}{(\delta + 1)^4},$$

em forma matricial tem-se que $\mathbf{D}(\mathbf{u}) = \text{diag}\{u_1, \dots, u_n\}$ e

$$\mathbf{K}_{\delta\delta} = -\mathbb{E}[\ddot{\mathbf{L}}_{\delta\delta}] = -\text{tr}[\mathbb{E}[\mathbf{D}(\mathbf{d})]] = \text{tr}[\mathbf{D}(\mathbf{u})].$$

Tomando a esperança de $\ddot{\mathbf{L}}_{\beta\delta}$, tem-se

$$\mathbb{E}[\ddot{\mathbf{L}}_{\beta\delta}] = \mathbb{E}[\mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{m}] = \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbb{E}[\mathbf{m}],$$

em que $\mathbb{E}[\mathbf{m}] = \text{diag}\{\mathbb{E}[m_1], \dots, \mathbb{E}[m_n]\}$. Assim, a esperança de um termo genérico $m_i, i = 1, \dots, n$ da matriz é dado por

$$\begin{aligned}
\mathbb{E}[m_i] &= \mathbb{E} \left[\frac{t_i}{(\delta t_i + t_i + \delta \mu_i)^2} + \frac{t_i}{4\mu_i^2} - \frac{\delta(\delta + 2)}{4(\delta + 1)^2 t_i} \right] \\
&= \frac{1}{(\delta + 1)^2} \mathbb{E} \left[\frac{t_i}{\left(t_i + \frac{\delta \mu_i}{(\delta + 1)}\right)^2} \right] + \frac{1}{4\mu_i^2} \mathbb{E}[t_i] - \frac{\delta(\delta + 2)}{4(\delta + 1)^2} \mathbb{E} \left[\frac{1}{t_i} \right] \\
&= \frac{1}{(\delta + 1)^2} \left[\frac{(\delta + 1)}{2\mu_i \delta} + \frac{\delta \mu_i}{(\delta + 1)} \mathcal{J}(\delta) \right] + \frac{1}{4\mu_i} - \frac{(\delta + 2)}{4\mu_i \delta} \\
&= \frac{1}{2\mu_i \delta (\delta + 1)^2} - \frac{1}{2\mu_i \delta} - \frac{\delta \mu_i}{(\delta + 1)^3} \mathcal{J}(\delta) \\
&= -\frac{1}{2\mu_i (\delta + 1)} - \frac{\delta \mu_i}{(\delta + 1)^3} \mathcal{J}(\delta).
\end{aligned}$$

Seja

$$s_i = -\mathbb{E}[m_i] = \frac{1}{2\mu_i(\delta+1)} + \frac{\delta\mu_i}{(\delta+1)^3} \mathcal{J}(\delta),$$

em forma matricial tem-se que $\mathbf{s} = \{s_1, \dots, s_n\}^\top$ e

$$\mathbf{K}_{\beta\delta} = \mathbf{K}_{\delta\beta}^\top = -\mathbb{E}[\mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{m}] = -\mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbb{E}[\mathbf{m}] = \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{s}.$$

Com isso, a matriz de informação esperada de Fisher é dada por

$$\mathbf{K}_{\theta\theta} = -\mathbb{E}[\ddot{\mathbf{L}}_{\theta\theta}] = \begin{bmatrix} \mathbf{K}_{\beta\beta} & \mathbf{K}_{\beta\delta} \\ \mathbf{K}_{\delta\beta} & \mathbf{K}_{\delta\delta} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^\top \mathbf{D}(\mathbf{v}) \mathbf{X} & \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{s} \\ \mathbf{s}^\top \mathbf{D}(\mathbf{a}) \mathbf{X} & \text{tr}[\mathbf{D}(\mathbf{u})] \end{bmatrix}.$$

Realizando os procedimentos para encontrar a inversa de $\mathbf{K}_{\theta\theta}$, tem-se que

$$\mathbf{K}_{\theta\theta}^{-1} = \begin{bmatrix} \mathbf{D}^{-1} + \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top & -\mathbf{B}\mathbf{R}^{-1} \\ \mathbf{R}^{-1}\mathbf{B}^\top & \mathbf{R}^{-1} \end{bmatrix},$$

em que

$$\begin{aligned} \mathbf{R} &= \text{tr}[\mathbf{D}(\mathbf{u})] - \mathbf{s}^\top \mathbf{D}(\mathbf{a}) \mathbf{X} [\mathbf{X}^\top \mathbf{D}(\mathbf{v}) \mathbf{X}]^{-1} \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{s} \\ &= \text{tr}[\mathbf{D}(\mathbf{u})] - \mathbf{s}^\top \mathbf{D}(\mathbf{a}) \mathbf{X} \mathbf{D}^{-1} \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{s} \\ &= \text{tr}[\mathbf{D}(\mathbf{u})] - \mathbf{s}^\top \mathbf{D}(\mathbf{a}) \mathbf{B}, \end{aligned}$$

e $\mathbf{D} = \mathbf{X}^\top \mathbf{D}(\mathbf{v}) \mathbf{X}$ e $\mathbf{B} = \mathbf{D}^{-1} \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{s}$.

A matriz $\mathbf{K}_{\theta\theta}$ pode ainda ser particionada de uma outra forma, como a forma quadrática de uma matriz do modelo e uma matriz de pesos,

$$\mathbf{K}_{\theta\theta} = \tilde{\mathbf{X}}^\top \tilde{\mathbf{W}} \tilde{\mathbf{X}},$$

em que

$$\tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} \quad \text{e} \quad \tilde{\mathbf{W}} = \begin{bmatrix} \mathbf{D}(\mathbf{v}) & \mathbf{D}(\mathbf{a}) \mathbf{s} \\ \mathbf{s}^\top \mathbf{D}(\mathbf{a}) & \text{tr}[\mathbf{D}(\mathbf{u})] \end{bmatrix},$$

sendo $\mathbf{0}$ um vetor de zeros de dimensão apropriada.

Dessa forma, a estimação dos parâmetros do MRBS pode ser feita utilizando o método Scoring de Fisher, como proposto em Leiva *et al* (2014), sendo o algoritmo de estimação

$$\begin{aligned}
\boldsymbol{\theta}^{(m+1)} &= \boldsymbol{\theta}^{(m)} + (\mathbf{K}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1})^m \mathbf{U}(\boldsymbol{\theta})^{(m)} \\
&= \boldsymbol{\theta}^{(m)} + \left(\tilde{\mathbf{X}}^\top \tilde{\mathbf{W}} \tilde{\mathbf{X}} \right)^{-1} \begin{bmatrix} \mathbf{X} \mathbf{D}(\mathbf{a})^{(m)} \mathbf{z}^{(m)} \\ \text{tr}[\mathbf{D}(\mathbf{b})]^{(m)} \end{bmatrix} \\
&= \boldsymbol{\theta}^{(m)} + \left(\tilde{\mathbf{X}}^\top \tilde{\mathbf{W}} \tilde{\mathbf{X}} \right)^{-1} \tilde{\mathbf{X}}^\top \begin{bmatrix} \mathbf{D}(\mathbf{a})^{(m)} & \mathbf{0} \\ \mathbf{0}^\top & \text{tr}[\mathbf{D}(\mathbf{b})]^{(m)} \end{bmatrix} \begin{bmatrix} \mathbf{z}^{(m)} \\ 1 \end{bmatrix} \\
&= \left(\tilde{\mathbf{X}}^\top \tilde{\mathbf{W}} \tilde{\mathbf{X}} \right)^{-1} \tilde{\mathbf{X}}^\top \tilde{\mathbf{W}} \boldsymbol{\zeta}^{(m)}
\end{aligned}$$

em que $m = 1, 2, \dots$ é o passo do algoritmo e

$$\boldsymbol{\zeta}^{(m)} = \tilde{\mathbf{X}} \boldsymbol{\theta}^{(m)} + \left(\tilde{\mathbf{W}}^{(m)} \right)^{-1} \begin{bmatrix} \mathbf{D}(\mathbf{a})^{(m)} & \mathbf{0} \\ \mathbf{0}^\top & \text{tr}[\mathbf{D}(\mathbf{b})]^{(m)} \end{bmatrix} \begin{bmatrix} \mathbf{z}^{(m)} \\ 1 \end{bmatrix}.$$

Com isso, a estimação de $\boldsymbol{\theta}$ é semelhante à realizada em um modelo linear generalizado, através de mínimos quadrados ponderados. Sob certas condições de regularidade tem-se que $\hat{\boldsymbol{\theta}} \stackrel{a}{\sim} \mathcal{N}_{p+1}(\boldsymbol{\theta}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$ e $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ é a matriz de variância-covariância de $\hat{\boldsymbol{\theta}}$, que pode ser obtida de $\mathbf{K}_{\boldsymbol{\theta}\boldsymbol{\theta}}$.

4.4 Simulações

Um estudo de simulação foi realizado para o MRBS no programa R, verificando a estimação de parâmetros para um modelo da forma

$$\log(\mu_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}, \quad i = 1, \dots, n$$

em que a função de ligação é a logarítmica, n é o tamanho de amostra e os valores das covariáveis x_{1i} e x_{2i} foram gerados das distribuições Uniforme(0, 1) e $\mathcal{N}(0, 1)$, respectivamente.

Foram tomados os valores de parâmetro $\beta_0 = 2$, $\beta_1 = 1,5$, $\beta_2 = 0,5$ e $\delta = 10$. Os tamanhos de amostra considerados foram $n \in \{15, 30, 50\}$, e para cada tamanho de amostra os valores das covariáveis foram gerados e mantidos fixos. Posteriormente, a variável resposta foi simulada gerando 1000 valores aleatórios da distribuição BSR com média $\exp(\mathbf{x}_i^\top \boldsymbol{\beta})$ e parâmetro de dispersão δ , em que $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^\top$ e $\mathbf{x}_i = (1, x_{1i}, x_{2i})^\top$. Foram calculadas a média, o desvio-padrão (DP) e erro quadrático médio (EQM) das estimativas dos parâmetros do MRBS, contidas na Tabela 3.

Analisando os dados da Tabela 3, pode-se observar que as estimativas dos coeficientes de regressão são mais precisas que do parâmetros δ , sendo os menores valores de DP e EQM obtidos para $\hat{\beta}_2$. No contexto geral, verifica-se que o tamanho de amostra melhora as estimativas obtidas, principalmente para δ , que se tem uma redução do EQM de 74,4 para 7,92.

Tabela 3: Estimativas dos parâmetros do MRBS na simulação.

Parâmetro	n	Média	DP	EQM
$\hat{\beta}_0$	$n = 15$	1,97	0,28	0,08
	$n = 30$	1,98	0,16	0,03
	$n = 50$	2,00	0,11	0,01
$\hat{\beta}_1$	$n = 15$	1,52	0,48	0,24
	$n = 30$	1,51	0,26	0,07
	$n = 50$	1,48	0,20	0,04
$\hat{\beta}_2$	$n = 15$	0,50	0,13	0,02
	$n = 30$	0,50	0,08	0,01
	$n = 50$	0,50	0,07	0,01
$\hat{\delta}$	$n = 15$	14,97	7,05	74,40
	$n = 30$	12,06	3,60	17,23
	$n = 50$	11,17	2,56	7,92

4.5 Influência Local

O método de influência local foi proposto por Cook (1986) para analisar a diferença de um modelo adotado e outro sob perturbações nos dados, ou no modelo.

Sendo $\boldsymbol{\theta}$ um vetor de dimensão $p \times 1$ dos parâmetros do modelo, e $\boldsymbol{\theta}_w$ o vetor de parâmetros do modelo sob perturbações, pode ser obtida uma medida da influência ocasionada pelo vetor de perturbação $\mathbf{w} \subset \mathcal{R}^n$, sendo n o número de observações,

$$LD(\mathbf{w}) = 2(l(\boldsymbol{\theta}) - l_w(\boldsymbol{\theta})),$$

em que $l(\boldsymbol{\theta})$ é o logaritmo da função de verossimilhança para os dados não perturbados, $l_w(\boldsymbol{\theta})$ é o logaritmo da função de verossimilhança para o modelo sob perturbação, sendo $LD(\mathbf{w})$ uma medida dos desvios na verossimilhança causada por \mathbf{w} .

Sendo \mathbf{w}_0 um vetor de não perturbação tal que $l_{\mathbf{w}_0}(\boldsymbol{\theta}) = l(\boldsymbol{\theta})$, segundo Paula (2013), o método consiste em avaliar o comportamento de $LD(\mathbf{w})$ em torno de \mathbf{w}_0 . A análise de influência é feita utilizando uma direção unitária $\|\mathbf{d}\| = 1$, e avaliando o gráfico de $LD(\mathbf{w}_0 + a\mathbf{d})$ versus a , $a \in \mathcal{R}$, sendo observadas as curvaturas normais $C_d(\boldsymbol{\theta})$ em torno de $a = 0$.

A curvatura normal é dada por

$$C_d(\boldsymbol{\theta}) = 2|\mathbf{d}^\top \boldsymbol{\Delta}^\top \ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1} \boldsymbol{\Delta} \mathbf{d}|,$$

sendo $-\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}$ a matriz observada de Fisher e $\boldsymbol{\Delta}$ a matriz de perturbação cujos elementos são

$$\Delta_{ij} = \frac{\partial^2 l_w(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i \partial \mathbf{w}_j},$$

avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, com $i = 1, \dots, p$ e $j = 1, \dots, n$. Com isso, é considerada a direção \mathbf{d}_{max} correspondente a maior curvatura C_{max} , sendo esta última tomada como o maior autovalor de

$\Delta^\top \ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1} \Delta$ e \mathbf{d}_{max} seu autovetor correspondente. A análise gráfica de influência consiste em avaliar o gráfico de \mathbf{d}_{max} versus a ordem das observações, podendo então serem observados os pontos com maior influência no modelo perturbado. Para maiores detalhes ver Cook (1986) e Paula (2013), por exemplo.

Ponderação de casos

A análise de influência local para o MRBS foi realizado por Leiva *et al* (2014), sendo considerado diversos casos de perturbação. Na situação de ponderação de casos são atribuídos pesos aos valores do logaritmo da função de verossimilhança de cada observação, sendo a verossimilhança perturbada dada por

$$\begin{aligned} l_{\mathbf{w}}(\boldsymbol{\theta}) &= \sum_{i=1}^n w_i l(t_i; \mu_i, \delta) \\ &= \sum_{i=1}^n \left[\frac{w_i \delta}{2} - \frac{w_i}{2} \log(\delta + 1) - \frac{w_i}{2} \log(16\pi\mu_i) - \frac{3w_i}{2} \log(t_i) + w_i \log(\delta t_i + t_i + \delta\mu_i) - \right. \\ &\quad \left. - \frac{w_i t_i (\delta + 1)}{4\mu_i} - \frac{w_i \delta^2 \mu_i}{4(1 + \delta)t_i} \right], \end{aligned}$$

sendo $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \delta)^\top$ o vetor de parâmetros do MRBS e $w_i \in \mathcal{R}$, $i = 1, \dots, n$.

Avaliando as derivadas tem-se que

$$\frac{\partial l_{\mathbf{w}}(\boldsymbol{\theta})}{\partial w_i} = \frac{\delta}{2} - \frac{1}{2} \log(\delta + 1) - \frac{1}{2} \log(16\pi\mu_i) - \frac{3}{2} \log(t_i) + \log(\delta t_i + t_i + \delta\mu_i) - \frac{t_i(\delta + 1)}{4\mu_i} - \frac{\delta^2 \mu_i}{4(1 + \delta)t_i},$$

$$\frac{\partial^2 l_{\mathbf{w}}(\boldsymbol{\theta})}{\partial w_i \partial \beta_j} = \frac{\partial^2 l_{\mathbf{w}}(\boldsymbol{\theta})}{\partial w_i \partial \mu_i} \frac{\mu_i}{\eta_i} \frac{\eta_i}{\beta_j} = \left[-\frac{1}{2\mu_i} + \frac{\delta}{(\delta t_i + t_i + \delta\mu_i)} + \frac{t_i(\delta + 1)}{4\mu_i^2} - \frac{\delta^2}{4t_i(\delta + 1)} \right] \frac{1}{g'(\mu_i)^{x_{ij}}},$$

e

$$\frac{\partial^2 l_{\mathbf{w}}(\boldsymbol{\theta})}{\partial w_i \partial \delta} = \frac{\delta}{2(\delta + 1)} + \frac{t_i + \mu_i}{\delta t_i + t_i + \delta\mu_i} - \frac{t_i}{4\mu_i} - \frac{\delta(\delta + 2)\mu_i}{4t_i(\delta + 1)^2}.$$

Com isso, tem-se que a forma matricial da matriz de derivadas dada por

$$\mathbf{\Delta} = \begin{bmatrix} \mathbf{X}^\top \mathbf{D}(\hat{\mathbf{a}}) \mathbf{D}(\hat{\mathbf{z}}) \\ \hat{\mathbf{b}}^\top \end{bmatrix}$$

sendo $\hat{\mathbf{a}}$, $\hat{\mathbf{z}}$ e $\hat{\mathbf{b}}$ os termos definidos em (4.2), (4.3) e (4.4), para o modelo adotado, tal que $\mathbf{\Delta}$ tenha dimensão $p \times n$.

Perturbação na variável resposta

A análise de influência local para o caso de perturbação na variável resposta também foi obtida em Leiva *et al* (2014), sendo adotada uma perturbação aditiva da forma $t_i(w_i) = t_i + w_i s(t_i)$, em que $w_i \in \mathcal{R}$ é o peso utilizado e $s(t_i)$ é um fator de escala que pode ser tomado como o desvio padrão da variável resposta, ou seja, $s(t_i) = \sqrt{r(\mu_i)/h(\delta)}$, como definido em (3.5). Com isso, a verossimilhança perturbada é dada por

$$\begin{aligned} l_{\mathbf{w}}(\boldsymbol{\theta}) &= \sum_{i=1}^n l_{w_i}(t_i(w_i); \mu_i, \delta) \\ &= \sum_{i=1}^n \left[\frac{\delta}{2} - \frac{1}{2} \log(\delta + 1) - \frac{1}{2} \log(16\pi\mu_i) - \frac{3}{2} \log[t_i(w_i)] \right. \\ &\quad \left. + \log[t_i(w_i)(\delta + 1) + \delta\mu_i] - \frac{(\delta + 1)t_i(w_i)}{4\mu_i} - \frac{\delta^2\mu_i}{4(1 + \delta)t_i(w_i)} \right], \end{aligned}$$

e sendo a derivada $\frac{dt_i(w_i)}{dw_i} = s(w_i)$, então as derivadas do logaritmo da verossimilhança são

$$\begin{aligned} \frac{\partial l_{\mathbf{w}}(\boldsymbol{\theta})}{\partial w_i} &= -\frac{3}{2} \frac{s(t_i)}{t_i(w_i)} + \frac{(\delta + 1)s(t_i)}{t_i(w_i)(\delta + 1) + \delta\mu_i} - \frac{(\delta + 1)s(t_i)}{4\mu_i} + \frac{\delta^2\mu_i s(t_i)}{4(1 + \delta)[t_i(w_i)]^2} \\ &= s(t_i) \left[-\frac{3}{2t_i(w_i)} + \frac{(\delta + 1)}{t_i(w_i)(\delta + 1) + \delta\mu_i} - \frac{(\delta + 1)}{4\mu_i} + \frac{\delta^2\mu_i}{4(1 + \delta)[t_i(w_i)]^2} \right], \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 l_{\mathbf{w}}(\boldsymbol{\theta})}{\partial w_i \partial \beta_j} &= \frac{\partial^2 l_{\mathbf{w}}(\boldsymbol{\theta})}{\partial w_i \partial \mu_i} \frac{\mu_i}{\eta_i} \frac{\eta_i}{\beta_j} \\ &= s(t_i) \left[-\frac{\delta(\delta + 1)}{[t_i(w_i)(\delta + 1) + \delta\mu_i]^2} + \frac{(\delta + 1)}{4\mu_i^2} + \frac{\delta^2}{4(1 + \delta)[t_i(w_i)]^2} \right] \frac{1}{g'(\mu_i)^{x_{ij}}}, \end{aligned}$$

e

$$\frac{\partial l_{\mathbf{w}}(\boldsymbol{\theta})}{\partial w_i \partial \delta} = s(t_i) \left[\frac{-\mu_i}{[t_i(w_i)(\delta + 1) + \delta\mu_i]^2} - \frac{1}{4\mu_i} + \frac{\delta(\delta + 1)\mu_i}{4(1 + \delta)^2[t_i(w_i)]^2} \right].$$

Definindo os seguintes termos $\varphi_i = s(t_i)$,

$$\psi_i = -\frac{\delta(\delta + 1)}{[t_i(w_i)(\delta + 1) + \delta\mu_i]^2} + \frac{(\delta + 1)}{4\mu_i^2} + \frac{\delta^2}{4(1 + \delta)[t_i(w_i)]^2},$$

e

$$\tau_i = \frac{-\mu_i}{[t_i(w_i)(\delta + 1) + \delta\mu_i]^2} - \frac{1}{4\mu_i} + \frac{\delta(\delta + 1)\mu_i}{4(1 + \delta)^2[t_i(w_i)]^2},$$

a matriz de derivadas é dada por

$$\mathbf{\Delta} = \begin{bmatrix} \mathbf{X}^\top \mathbf{D}(\hat{\mathbf{a}}) \mathbf{D}(\hat{\boldsymbol{\psi}}) \mathbf{D}(\hat{\boldsymbol{\phi}}) \\ \hat{\boldsymbol{\tau}}^\top \mathbf{D}(\hat{\boldsymbol{\phi}}) \end{bmatrix},$$

com $\hat{\boldsymbol{\psi}} = (\hat{\psi}_1, \dots, \hat{\psi}_n)^\top$, $\hat{\boldsymbol{\tau}} = (\hat{\tau}_1, \dots, \hat{\tau}_n)^\top$ e $\hat{\boldsymbol{\phi}} = (\hat{\phi}_1, \dots, \hat{\phi}_n)^\top$ sendo os vetores com os termos estimados para o modelo adotado.

Com a análise de influência é possível verificar como está a qualidade do ajuste do modelo com a distribuição BSR, verificando quais observações estão afetando as estimativas no procedimento chamado de análise confirmatória. Uma vez que o modelo já tenha sido ajustado e analisado, mas se deseje verificar outras opções de modelagem para o MRBS, outras abordagens da distribuição podem ser consideradas. Neste trabalho foi proposta uma abordagem Bayesiana da BSR, sendo abordada tanto o ajuste da distribuição como de modelos de regressão, oferecendo uma alternativa à abordagem clássica de análise de dados.

5 ABORDAGEM BAYESIANA

5.1 Inferência Bayesiana

Na abordagem Bayesiana, assume-se que os parâmetros sobre os quais se deseja realizar inferências estão associados a uma distribuição de probabilidade, sendo a distribuição dos mesmos definidas utilizando as informações adicionais que se possa ter sobre as quantidades de interesse ou de estudos anteriores da população estudada.

Analisando possíveis modelos que possam ser construídos para a distribuição Birnbaum-Saunders reparametrizada, foram consideradas algumas distribuições *a priori* que poderiam ser assumidas para os parâmetros sendo analisado também qual a mais adequada com base em métodos de comparação de modelos.

5.1.1 Distribuições a priori gama

Como μ e δ assumem somente valores positivos, uma possível suposição seria assumir distribuições *a priori* gama independentes para ambos os parâmetros, em que a fdp *a priori* seria

$$\pi(\mu|a, b) \propto \mu^{a-1} \exp\{-b\mu\}$$

e

$$\pi(\delta|c, d) \propto \delta^{c-1} \exp\{-d\delta\},$$

com a, b, c e d sendo positivos.

Dessa forma, a distribuição *a posteriori* conjunta é dada por

$$\begin{aligned} \pi(\mu, \delta|\mathbf{y}) &\propto \prod_{i=1}^n y_i^{-3/2} \left(y_i + \frac{\delta\mu}{(1+\delta)} \right) \exp \left\{ -\frac{(1+\delta)}{4\mu} \sum_{i=1}^n y_i - \frac{\delta^2\mu}{4(1+\delta)} \sum_{i=1}^n \frac{1}{y_i} + \frac{n\delta}{2} \right\} \\ &\times \mu^{a-\frac{n}{2}-1} (\delta+1)^{\frac{n}{2}} \delta^{c-1} \exp\{-b\mu - d\delta\}. \end{aligned} \quad (5.1)$$

Observe que a distribuição *a posteriori* encontrada não é uma distribuição de probabilidade conhecida. Além disto, não existe forma fechada. Portanto, métodos de simulação, como o MCMC (Monte Carlo via Cadeias de Markov) são utilizados para gerar amostras aleatórias, e consequentemente realizar inferências sobre os parâmetros.

5.1.2 Distribuições a priori uniforme

Em algumas situações não temos informação adicional sobre a população estudada e seus parâmetros, uma possibilidade é utilizar distribuição *a priori* uniforme para μ e δ , fornecendo um suporte largo o suficiente de forma a cobrir os possíveis valores que os

parâmetros possam assumir. Dessa forma, as funções de densidade *a priori* são

$$\pi(\mu|a, b) \propto \frac{1}{b-a} \quad a < \mu < b$$

e

$$\pi(\delta|c, d) \propto \frac{1}{d-c}, \quad c < \delta < d.$$

A distribuição *a posteriori* será proporcional à função de verossimilhança, e amostras aleatórias dos parâmetros podem ser obtidas através de métodos de MCMC.

5.1.3 Distribuições *a priori* Log-Normal

Outra possibilidade de distribuição *a priori* para os parâmetros é a distribuição Log-Normal, a qual é utilizada para valores positivos e pode ser uma opção para as distribuições gama e uniforme. As funções densidade *a priori* nesse caso serão

$$\pi(\mu|a, b) \propto \frac{1}{\mu} \exp\left\{-\frac{b}{2}(\log \mu - a)^2\right\} \quad \text{e} \quad \pi(\delta|c, d) \propto \frac{1}{\delta} \exp\left\{-\frac{d}{2}(\log \delta - c)^2\right\},$$

com a e c sendo valores reais e b e d positivos. A função densidade *a posteriori* é dada por

$$\begin{aligned} \pi(\mu, \delta|\mathbf{y}) &\propto \frac{1}{\mu\delta} \prod_{i=1}^n y_i^{-3/2} \left(y_i + \frac{\delta\mu}{(1+\delta)}\right) \\ &\times \exp\left\{-\frac{(1+\delta)}{4\mu} \sum_{i=1}^n y_i - \frac{\delta^2\mu}{4(1+\delta)} \sum_{i=1}^n \frac{1}{y_i} + \frac{n\delta}{2} - \frac{b}{2}(\log \mu - a)^2 - \frac{d}{2}(\log \delta - c)^2\right\}. \end{aligned}$$

5.1.4 Simulações

Para verificar a performance de cada distribuição *a priori*, um estudo de simulação foi realizado utilizando os programas R e OpenBUGS. Amostras aleatórias da BSR foram geradas no programa R, sendo então esses valores usados em modelos escritos em códigos no OpenBUGS, através do pacote R2openBUGS.

As mesmas distribuições *a priori* foram atribuídas aos mesmos parâmetros, e os valores dos hiper-parâmetros a , b , c e d foram escolhidos com o objetivo de garantir aproximadamente valores de 10 e 100 respectivamente para a média e variância das distribuições *a priori*. Esses valores foram escolhidos de forma a representar uma situação na qual não se tem conhecimentos anteriores dos valores μ e δ , isto é, se tratam de distribuições *a priori* vagas. As distribuições *a priori* foram Uniforme(0, 50), Gama(1; 0, 1) e Log-Normal(1, 9561; 1, 4426). Amostras de tamanho 5000 foram geradas com técnicas de MCMC, descartando os primeiros 1000 valores no período de aquecimento (*burn-in*). Amostras de tamanho de 30 da distribuição BSR(1, 1) foram geradas em R.

Analisando os resultados nas Tabelas (4)-(6), pode-se observar que os modelos são

bastante similares, apesar das diferentes distribuições *a priori*. Para o parâmetro μ a diferença foi pequena, mas por outro lado, para δ o melhor modelo foi o Log-Normal, com o valor mais próximo de 1 e com o desvio padrão próximo dos demais modelos. De acordo com o componente do desvio, o melhor modelo foi o Uniforme, mas com pouca diferença dos demais.

Tabela 4: Resultados usando priores Uniforme

	real	média	DP	2,5%	25%	50%	75%	97,5%
μ	1	1,64	0,56	0,94	1,26	1,51	1,87	3,02
δ	1	0,80	0,21	0,44	0,65	0,79	0,93	1,25
desvio	-	68,24	2,11	66,14	66,72	67,61	69,09	73,94

Tabela 5: Resultados usando priores Gama

	real	média	DP	2,5%	25%	50%	75%	97,5%
μ	1	1,65	0,58	0,94	1,27	1,52	1,87	3,08
δ	1	0,80	0,21	0,44	0,65	0,78	0,93	1,26
desvio	-	68,24	2,22	66,14	66,69	67,58	69,08	74,06

Tabela 6: Resultados usando priores Log-normal

	real	média	DP	2,5%	25%	50%	75%	97,5%
μ	1	1,60	0,46	0,96	1,28	1,52	1,82	2,75
δ	1	0,87	0,21	0,51	0,72	0,85	1,00	1,30
desvio	-	68,15	2,03	66,14	66,69	67,53	69,01	73,65

5.2 Modelo de regressão

A construção de modelos de regressão Birnbaum-Saunders com enfoque Bayesiano pode ser feita de maneira semelhante à realizada em modelos lineares generalizados Bayesianos. A modelagem em MLG Bayesiano é semelhante à abordagem frequentista, mas com a diferença dos parâmetros (coeficientes de regressão e parâmetro de dispersão) possuírem uma distribuição de probabilidade associada.

No caso da distribuição Birnbaum-Saunders reparametrizada, considerando uma amostra aleatória $\mathbf{t} = (t_1, \dots, t_n)^\top$ como variável resposta, um vetor \mathbf{x} de variáveis explicativas e $\boldsymbol{\beta}$ um vetor dos seus coeficientes de regressão associados, uma opção pode ser adotar um modelo em que para cada indivíduo se tem a seguinte estrutura hierárquica

$$\begin{aligned} t_i | \boldsymbol{\beta}, \delta &\sim BSR(g^{-1}\{\mathbf{x}_i^\top \boldsymbol{\beta}\}, \delta), i = 1, \dots, n \\ \beta_j &\sim \mathcal{N}(0, \sigma^2), j = 1, \dots, p \\ \delta &\sim \text{Uniforme}(a, b), \end{aligned}$$

com n sendo o tamanho da amostra e p o número de coeficientes de regressão. Nesse modelo, os coeficientes de regressão são considerados identicamente distribuídos com distribuição normal, sendo utilizado distribuições *a priori* independentes, podendo ser atribuída uma variância σ^2

alta no caso de se ter pouca informação a respeito da relação entre as variáveis preditoras e a variável resposta. Nesse modelo também foi considerado que a distribuição de probabilidade do vetor $\boldsymbol{\beta}$ é independente da distribuição do parâmetro de dispersão δ , e foi utilizada uma distribuição *a priori* não informativa para este, em que não se tendo informações adicionais sobre tal parâmetro na população de estudo, um largo comprimento para o intervalo (a, b) de uma distribuição uniforme pode ser adotado a fim de cobrir os possíveis valores que δ pode assumir.

Em posse das distribuições *a priori* dos coeficientes de regressão, assim como no MLG frequentista, a média é obtida por $\mu_i = g^{-1}(\mathbf{x}_i^\top \boldsymbol{\beta})$, sendo $g(\cdot)$ uma função de ligação estritamente monótona e positiva. Com isso, as distribuições *a posteriori* dos parâmetros do modelo são obtidas por

$$\begin{aligned}\pi(\boldsymbol{\beta}|\delta, \mathbf{t}) &\propto L(\mathbf{t}, \boldsymbol{\beta}, \delta)\pi(\delta) \\ \pi(\delta|\boldsymbol{\beta}, \mathbf{t}) &\propto L(\mathbf{t}, \boldsymbol{\beta}, \delta)\pi(\boldsymbol{\beta})\end{aligned}$$

sendo $L(\mathbf{t}, \boldsymbol{\beta}, \delta)$ a função de verossimilhança dos dados, $\pi(\delta)$ e $\pi(\boldsymbol{\beta})$ as respectivas distribuições *a priori* de δ e $\boldsymbol{\beta}$, e também $\pi(\boldsymbol{\beta}|\delta, \mathbf{t})$ e $\pi(\delta|\boldsymbol{\beta}, \mathbf{t})$, suas respectivas distribuições *a posteriori*.

5.2.1 Simulações

Para verificar a performance do MRBS Bayesiano foram realizadas simulações utilizando os programas R e OpenBUGS. A situação considerada foi a mesma da simulação para o MRBS frequentista apresentado no Capítulo 4, utilizando a função de ligação logarítmica. Assim como o MLG é apresentado em GELMAN *et al* (2004), o modelo da distribuição dos dados na BSR é da forma

$$L(t_i|\boldsymbol{\beta}, \delta) = \prod_{i=1}^n \frac{\exp\{\delta/2\}\sqrt{1+\delta}}{4\sqrt{\pi\exp(\eta_i)}} t_i^{-3/2} \left(t_i + \frac{\delta \exp(\eta_i)}{(1+\delta)} \right) \exp \left\{ -\frac{\delta}{4} \left(\frac{(1+\delta)t_i}{\delta \exp(\eta_i)} + \frac{\delta \exp(\eta_i)}{(1+\delta)t_i} \right) \right\},$$

em que n é o tamanho de amostra, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^\top$ é o vetor de parâmetros e $\eta_i = \mathbf{x}_i^\top \boldsymbol{\beta}$ é o preditor linear da i -ésima observação. Os valores da covariáveis x_{1i} e x_{2i} foram gerados das distribuições Uniforme(0, 1) e $\mathcal{N}(0, 1)$, respectivamente.

Novamente, foram tomados os valores de parâmetro $\beta_0 = 2$, $\beta_1 = 1,5$, $\beta_2 = 0,5$ e $\delta = 10$. Os tamanhos de amostra considerados foram $n \in \{15, 30, 50\}$, e para cada tamanho de amostra os valores das covariáveis e variável resposta foram gerados, sendo tomado $t_i|\boldsymbol{\beta}, \delta \sim BS(\exp\{\mathbf{x}_i^\top \boldsymbol{\beta}\}, \delta)$. Com os dados gerados, os valores foram usados no programa OpenBUGS para construir os modelos de regressão Bayesianos. As estimativas obtidas das distribuições *a posteriori* dos parâmetros foram a média, desvio padrão, mediana e percentis de ordem 2,5 ($P_{2,5}$) e 97,5 ($P_{97,5}$). Os resultados do OpenBUGS também fornecem a medida MC error, ou erro de Monte Carlo, que se trata de uma medida de incerteza da média *a posteriori*, sendo utilizado para indicar que a média *a posteriori* é precisa em um intervalo de mais ou menos o MC error (COWLES, 2013).

Tabela 7: Estimativas dos parâmetros do MRBS Bayesiano simulado.

n	Parâmetro	Real	Média	DP	MC error	$P_{2,5}$	Mediana	$P_{97,5}$
$n = 15$	β_0	2	1,891	0,6009	0,0587	0,05	2,068	2,37
	β_1	1,5	1,344	0,844	0,08048	0,51	1,164	4,29
	β_2	0,5	0,365	0,1499	0,0114	0,11	0,3535	0,73
	δ	10	18,83	8,823	0,8276	0,08	19,38	38,56
$n = 30$	β_0	2	2,001	0,4279	0,0407	0,01	2,071	2,43
	β_1	1,5	1,324	0,4235	0,0384	0,014	1,379	2,02
	β_2	0,5	0,582	0,2844	0,0258	0,026	0,5407	1,44
	δ	10	8,332	2,983	0,2642	0,01	8,568	13,79
$n = 50$	β_0	2	2,057	0,3696	0,0341	1,72	2,044	2,82
	β_1	1,5	1,411	0,4991	0,0471	-0,002	1,536	2,01
	β_2	0,5	0,527	0,1122	0,01006	0,03	0,5375	0,68
	δ	10	8,632	2,782	0,2544	1,179	8,84	13,24

Considerando os valores utilizados para gerar as amostras como $\theta_s = (\beta_0 = 2, \beta_1 = 1,5, \beta_2 = 0,5, \delta = 10)^\top$ e analisando os dados da Tabela (7), pode-se observar que o estimador de Bayes (média) se torna mais próximo dos valores de θ_s à medida que n aumenta, ocorrendo também uma diminuição dos valores de DP e MC error na maioria dos casos, sendo estes indícios de melhora no ajuste para tamanho de amostra maior. Também se pode verificar que todos os intervalos formados pelos percentis $P_{2,5}$ e $P_{97,5}$ contêm seus respectivos valores de θ_s .

5.3 Comparação de modelos

Uma vez que uma distribuição *a priori* precisa ser escolhida para construir o modelo, se não houver informações adicionais sobre a distribuição dos parâmetros, uma forma de escolher a melhor distribuição *a priori* é usar algum critério de comparação entre os ajustes feitos com cada modelo. Um desses critérios é o CPO (*Conditional Predictive Ordinate*), que avalia o ajuste dos modelos com base na qualidade da predição de novos valores. O CPO considera a probabilidade do modelo prever o valor de uma das observações usando todos os dados exceto a observação predita, o que de acordo com Christensen et al. (2011) pode ser calculado de uma amostra de parâmetros $\theta^1, \dots, \theta^s$ por

$$CPO_i^{-1} = \frac{1}{s} \sum_{i=1}^s \frac{1}{f_i(y_i | \theta^k, M^{[i]})},$$

onde y_i é o valor observado e $M^{[i]}$ é o valor atual do modelo, construído com todos os dados exceto a i -ésima observação. Quanto maior for o valor de CPO_i , melhor é o indício de que o modelo está ajustando bem a observação considerada.

Dados dois modelos M_1 e M_2 , segundo Paulino *et al.* (2003) e Geisser e Eddy (1979), os modelos podem ser comparados através do Pseudo Fator de Bayes (PFB), no qual os valores de $CPO(M_1)$ e $CPO(M_2)$ sob os respectivos modelos são utilizados na verificação

de qual realizou um melhor ajuste dos dados. Denotando-se por $PFB(M_1, M_2)$, essa medida é calculada da forma

$$PFB(M_1, M_2) = \frac{\prod_{i=1}^n C\hat{P}O_i(M_1)}{\prod_{i=1}^n C\hat{P}O_i(M_2)},$$

conforme é apresentado em Farias e Lemonte (2011), sendo $C\hat{P}O_i(M_j)$ o valor de CPO obtido da i -ésima observação sob o modelo j . Caso $PFB(M_1, M_2) < 1$ o modelo M_2 é preferível em relação a M_1 , e caso $PFB(M_1, M_2) > 1$, o modelo M_1 é preferível em relação a M_2 .

5.4 Diagnóstico de influência Bayesiano

Com o objetivo de verificar a presença de observações influentes no modelo de regressão Bayesiano, Cho *et al* (2009) propuseram um método utilizando a medida de divergência de Kullback-Leibler (K-L) para identificar tais pontos.

Denotando D como o conjunto de dados completo e $D^{[i]}$ o conjunto de dados sem a i -ésima observação, a divergência K-L é calculada por

$$K(P, P^{[i]}) = \int \pi(\boldsymbol{\theta}|D) \log \left\{ \frac{\pi(\boldsymbol{\theta}|D)}{\pi(\boldsymbol{\theta}|D^{[i]})} \right\} d\boldsymbol{\theta},$$

em que $\boldsymbol{\theta}$ é o vetor de parâmetros do modelo, P a distribuição *a posteriori* de $\boldsymbol{\theta}$ para os dados completos e $P^{[i]}$ a distribuição *a posteriori* para os dados sem a i -ésima observação. Os autores Cho *et al* (2009) também mostraram que esta medida pode ser obtida através da expressão

$$K(P, P^{[i]}) = -\log(CPO_i) + \mathbb{E}_{\boldsymbol{\theta}} [\log \{f(y_i|\boldsymbol{\theta})\} | D], \quad (5.2)$$

em que $\mathbb{E}_{\boldsymbol{\theta}} [\cdot | D]$ representa a esperança da distribuição *a posteriori* de $\boldsymbol{\theta}$ dado D . A expressão (5.2) possui grande utilidade, pois permite calcular a divergência K-L utilizando as amostras de MCMC da distribuição *a posteriori* de $\boldsymbol{\theta}$ para os dados completos, assim como foi feito no cálculo do CPO.

Uma medida utilizada com a divergência K-L é a de calibração, cuja definição pode ser encontrada em McCulloch (1989). Dado que $K(Q_1, Q_2) = k$ para duas distribuições *posteriori* Q_1 e Q_2 , o objetivo consiste em encontrar um valor $q(k)$ tal que $K[B(1/2), B(q(k))] = k$, em que $B(p)$ representa uma distribuição de Bernoulli de parâmetro p . Segundo McCulloch (1989) o valor de $q(k)$ é a medida de calibração de k , e permite observar a semelhança das divergências K-L entre as situações Q_1 e Q_2 , e $B(1/2)$ e $B(q(k))$. Em um exemplo apresentado por McCulloch (1989), se $q(k) \approx 1$, os modelos Q_1 e Q_2 são notoriamente diferentes, enquanto se $q(k) \approx 0,5$, os modelos são bastante similares. Tomando $q(k)$ como p_i , segundo Cho *et al* (2009), $K(P, P^{[i]})$ pode ser medida através de

$$K(P, P^{[i]}) = K(B(1/2), B(p_i)) = \frac{-\log \{4p_i(1-p_i)\}}{2},$$

no qual utilizando a medida de divergência para calcular p_i , tem-se que

$$p_i = \frac{1 + \sqrt{1 - \exp\{-2K(P, P^{[i]}\)}}}{2},$$

sendo neste caso $0,5 \leq p_i \leq 1$.

Com as técnicas de diagnóstico de influência e comparação de modelos, a verificação de qualidade do ajuste no MRBS através da abordagem Bayesiana pode ser realizada, o que complementa os métodos de modelagem apresentados no capítulo. Conjuntamente com os conceitos na metodologia clássica, a distribuição RBS pode ser utilizada em aplicações com dados reais, o que será feito no próximo capítulo, verificando como é realizada a modelagem através das técnicas apresentadas e sua performance.

6 APLICAÇÃO

6.1 Pacientes com leucemia

6.1.1 Análise descritiva

Estes dados foram analisados inicialmente por Feigl e Zelen (1965) utilizando modelos com distribuição exponencial, e consistem em dados sobre o tempo (em semanas) de sobrevivência de 33 pacientes que faleceram devido a uma forma de leucemia aguda. Os pacientes formavam dois grupos do fator AG, em relação à presença de bastão de Auer nas células da medula (FEIGL e ZELLEN, 1965). Além de AG, a outra variável explicativa do modelo era o logaritmo do número de células brancas no sangue do paciente no momento do diagnóstico da doença, em que segundo Feigl e Zelen (1965), quanto maior o número de células brancas, menos tempo de vida o paciente tende a ter.

Denotando-se as variáveis tempo de vida e logaritmo do número de células brancas por *semanas* e *IWBC*, respectivamente, na Tabela 8 estão presentes algumas estatísticas descritivas referentes às mesmas. Pode ser observado pela mediana de *semanas* que 50% dos pacientes sobreviveram até 22 semanas, porém, o tempo médio de vida foi de 40,87 semanas, um número maior que a mediana devido à natureza assimétrica dos dados. Para a covariável *IWBC*, a assimetria das observações foi de $-0,08$, revelando uma distribuição consideravelmente simétrica da mesma. Em relação à covariável *AG*, teve-se 16 pacientes com ausência deste fator (tomados como $AG = 0$, chamados de *AG* negativos), e 17 pacientes que apresentaram o fator *AG* nas análises da medula (tomados como $AG = 1$, chamados de *AG* positivos).

Tabela 8: Análise descritiva.

Variável	Mínimo	Mediana	Média	Máximo	CV	Assimetria	Excesso de Curtose
<i>semanas</i>	1	22	40,87	156	1,1424	1,11	-0,06
<i>IWBC</i>	6,62	9,25	9,52	11,51	0,1411	-0,08	-0,96

Na Figura 5a, pode-se notar que os pacientes *AG* positivos aparentam ter um tempo de sobrevivência maior que os *AG* negativos, por outro lado, pela Figura 5b, não parece haver relação entre as covariáveis *AG* e *IWBC*.

Na Figura 6, pode ser observado que à medida que *IWBC* aumenta (maior número de células brancas), o tempo de sobrevivência dos pacientes aparenta diminuir, sendo este tempo de vida geralmente menor para os pacientes *AG* negativos, como observado em 5b.

6.1.2 Modelos de regressão clássicos

Utilizando a distribuição Birnbaum-Saunders para modelar a variável *semanas*, foi

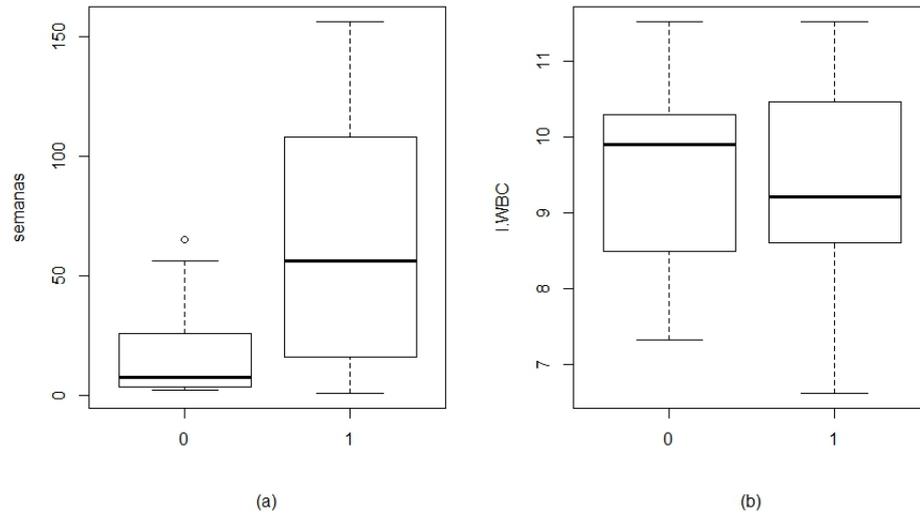


Figura 5: Boxplot das variáveis *semanas* e *IWBC* para cada nível de *AG*.

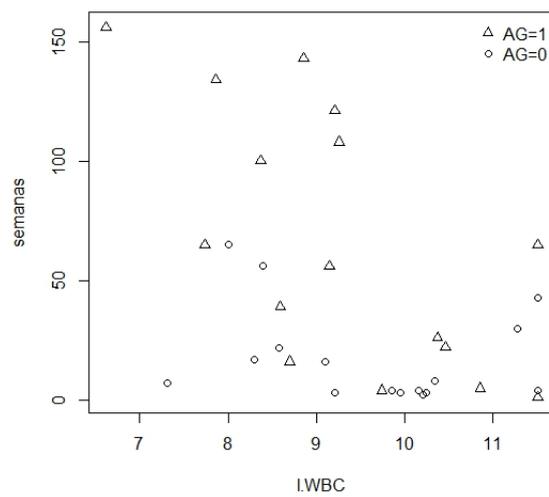


Figura 6: Gráfico de dispersão das variáveis *semanas* e *IWBC* para cada nível de *AG*.

ajustado um modelo da forma

$$g(\mu_i) = \eta_i = \beta_0 + \beta_1 IWBC_i + \beta_2 AG_i, \quad i = 1, \dots, 33,$$

em que neste caso $semanas_i \sim BSR(\mu_i, \delta)$ e foi utilizada uma função de ligação logarítmica, $g(\mu_i) = \log(\mu_i)$. Na Tabela 9 são apresentadas as estimativas do modelo.

Tabela 9: Estimativas do MRBS para dados de leucemia.

Parâmetro	Estimativa	Erro padrão
β_0	7,9516	1,4124
β_1	-0,5058	0,1435
β_2	0,6736	0,3800
δ	1,0727	0,2641

Pelos dados da Tabela 9, pode-se observar que a média da variável *semanas* tende a decrescer à medida que a variável *IWBC* cresce, pois o tempo esperado de vida em *semanas* de um paciente decresce 39,69% ($e^{\hat{\beta}_1} = 0,3969$) quando o logaritmo do número de células brancas no sangue aumenta em uma unidade, mantendo a variável *AG* fixa. Em relação à variável *AG*, a presença de deste fator aumenta em 96,13% ($e^{\hat{\beta}_2} = 1,9613$) o tempo de vida esperado do paciente, em semanas. Esses resultados vão ao encontro das observações feitas por Feigl e Zelen (1965) sobre os dados e pelas análises das Figuras 6 e 5.

Na Figura 7 estão dispostos os resíduos versus os valores ajustados, que permite verificar como está a adequação do modelo ajustado. Pode-se observar que a grande maioria dos resíduos se encontram em torno de zero, na região de ± 2 , mas existem dois pontos que se destacam dos demais, as observações #33 e #17, que podem se tratar de pontos influentes no modelo.

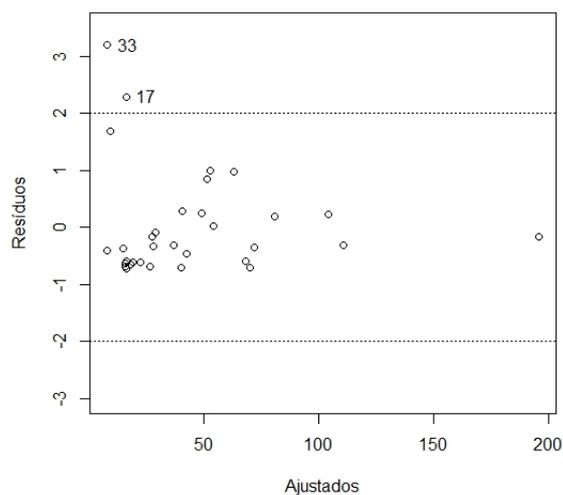


Figura 7: Gráfico dos resíduos vs. os valores ajustados do MRBS para dados de leucemia.

Outra forma de verificar a adequabilidade do modelo é através do gráfico de envelope simulado (ver Cordeiro e Demétrio (2007) e Atkinson (1985), por exemplo). Neste trabalho, a análise de envelope simulado para o MRBS foi implementada no programa R, sendo obtida a Figura 8. Pode-se notar que todos os pontos se encontram dentro das bandas de confiança de 95%, o que é um indício de que o modelo ajustado está adequado.

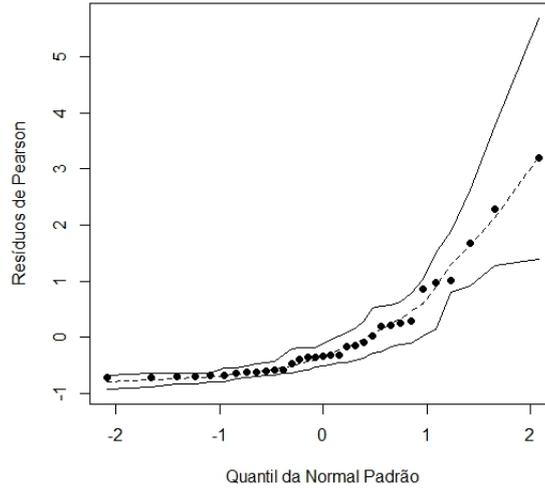


Figura 8: Gráfico de envelope simulado para o MRBS dos dados de leucemia.

Na Figura 9 de d_{max} versus os índices das observações, pode-se observar quais pontos são influentes para o modelo sob o método de ponderação de casos, sendo a linha tracejada dada por $2\bar{d}_{max}$. Observa-se que os pontos #33 e #17 são os mais influentes para o modelo, assim como foi observado no gráfico de resíduos, porém outros pontos também se destacam, as observações #31, #14 e #15.

6.1.3 Modelos de regressão Bayesianos

Utilizando os métodos apresentados no Capítulo 5, foi ajustado um modelo de regressão com abordagem Bayesiana para os dados, com as seguintes distribuições *a priori* para as quantidades de interesse

$$\begin{aligned} semana_i | \boldsymbol{\beta}, \delta &\sim BSR(\mu_i, \delta), i = 1, \dots, 33 \\ \beta_j &\sim \mathcal{N}(0; 0, 001), j = 1, \dots, 3 \\ \delta &\sim \text{Log-Normal}(1, 9561; 1, 4426), \end{aligned}$$

em que $\log(\mu_i) = \eta_i = \beta_1 + \beta_2 lWBC_i + \beta_3 AG_i$, sendo utilizada uma função de ligação logarítmica para a média. As distribuições *a priori* escolhidas foram as utilizadas nas simulações, sendo feita a opção pela distribuição Log-Normal para δ por sua performance ter sido melhor

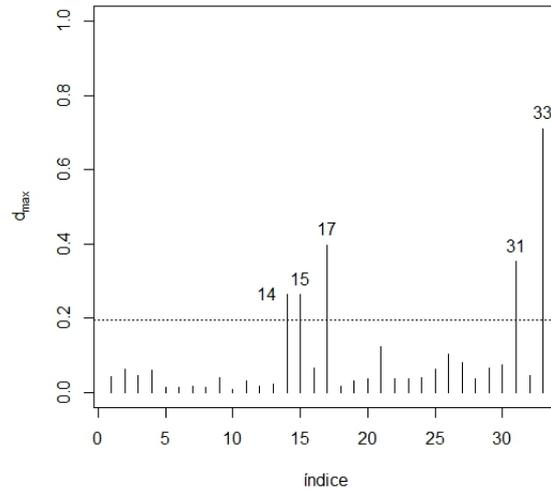


Figura 9: Gráfico de influência local para o MRBS dos dados de leucemia.

do que as demais nas simulações, sendo os hiper-parâmetros escolhidos de forma a ter-se uma distribuição *a priori* vaga, com alta variância para δ . As estimativas obtidas para o modelo estão apresentadas na Tabela 10.

Tabela 10: Predições dos parâmetros do MRBS Bayesiano para os dados de leucemia.

Parâmetro	Média	DP	$P_{2,5}$	Mediana	$P_{97,5}$
β_1	8,009	1,2968	5,518	8,004	10,550
β_2	-0,511	0,1291	-0,763	-0,512	-0,262
β_3	0,669	0,3659	-0,062	0,671	1,369
δ	1,132	0,2523	0,682	1,116	1,662

Analisando os dados da Tabela 10, nota-se que as predições são próximas das estimativas obtidas no modelo clássico, sendo as interpretações semelhantes. Nota-se que, a medida que a variável *IWBC* cresce, o tempo médio de sobrevivência dos pacientes tende a diminuir, enquanto que a medida que a presença do fator AG no paciente tende a aumentar o tempo de vida do paciente. Observa-se também que o zero está incluído no intervalo de 95% de credibilidade para β_3 , sendo um indício de que a variável AG não teria tanta importância para explicar *semanas*, porém, este valor se encontra perto do limite do intervalo.

Analisando a Figura 10, tem-se as medidas de divergência K-L das observações, indicando quais são as influentes para o modelo. Pode-se ver que as observações mais influentes são as #33, #14 e #15, se destacando das demais. As observações #17, #21 e #31 também se destacam no gráfico de divergência K-L.

Na Tabela 11 estão os valores de calibração para o modelo ajustado, em ordem decrescente. Pode-se ver que a observação #33 tem um valor próximo de 1, indicando que esta observação é bastante influente para o modelo assim como foi observado na Figura 10. Outros valores altos também foram das observações #14 e #15, com $k = 0,8464$.

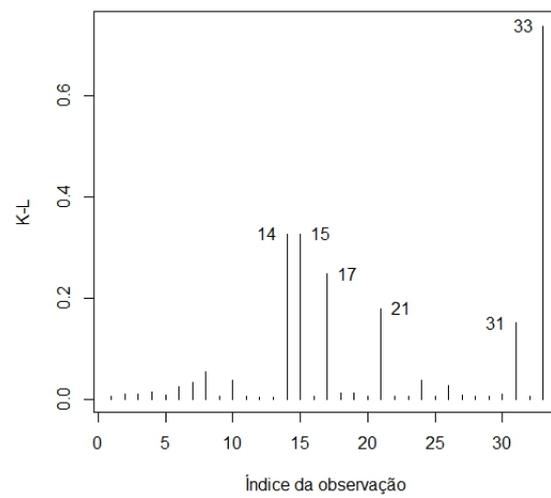


Figura 10: Gráfico da divergência K-L para o MRBS Bayesiano dos dados de leucemia.

Tabela 11: Calibração da divergência K-L dados de leucemia.

Observação	K-L	Calibração
#33	0,7357	0,9388
#14	0,3271	0,8464
#15	0,3271	0,8464
#17	0,2478	0,8126
#21	0,1804	0,7752
#31	0,1512	0,7554

6.1.4 Análise confirmatória

As observações identificadas como influentes no modelo podem estar afetando o valor das estimativas dos parâmetros (modelo clássico) ou das predições *a posteriori* (modelo Bayesiano). Dessa forma, é necessária uma análise confirmatória para verificar o impacto causado nas estimativas pela permanência de alguma dessas observações influentes, com o objetivo de estudar a influência das mesmas e possivelmente obter um modelo menos influenciado por observações atípicas.

Neste trabalho esta análise será feita conforme as abordagens clássica e Bayesiana, verificando e comparando os resultados de ambas, e para isso, as medidas de influência d_{max} e K-L foram centralizadas no intervalo $[0, 1]$ para que os valores obtidos em ambas pudessem ser comparados. Nas Figuras 11a e 11b, pode-se observar que a observação #33 é dada como influente em ambas abordagens. Por outro lado, na abordagem clássica as duas observações mais influentes foram #33, #17 e #31, enquanto que na abordagem Bayesiana foram as observações #14 e #15.

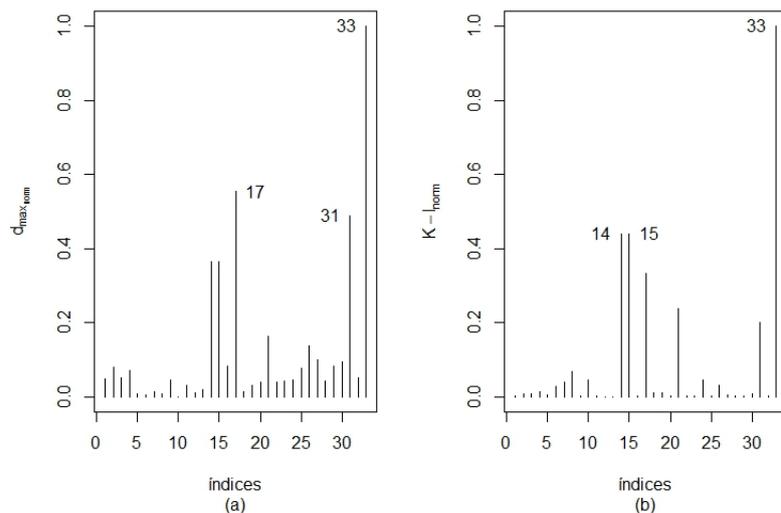


Figura 11: Comparação dos gráficos de influência local e da divergência K-L.

Na Tabela 12 são apresentados os dados das observações diagnosticadas como influentes na Figura 11. Nas interpretações feitas nos modelos, a presença do fator *AG* geralmente leva a um tempo esperado de vida maior para os pacientes, porém, os pacientes #14 e #15 são *AG* positivos e apresentaram o número mínimo de semanas de vida observados, o que pode ter levado essas observações a serem apontadas como influentes. Pode-se observar que os valores da Tabela 12 possuem um valor próximo de *IWBC*, mas a observação #17 possui um número de semanas bastante alto quando comparados com as demais observações. O mesmo ocorre para as observações #31 e #33, que também apresentaram um alto número de semanas de vida, porém, estas observações possuem $AG = 0$, e segundo as interpretações dos modelos, pacientes *AG* negativos tendem a ter tempo de vida menor, e essa ligeira contradição pode ter levado essas observações a serem tomadas como influentes nas análises feitas.

Tabela 12: Observações influentes da Figura 11.

Observação	<i>semanas</i>	<i>IWBC</i>	<i>AG</i>
#14	1	11,5129	1
#15	1	11,5129	1
#17	65	11,5129	1
#31	30	11,2772	0
#33	43	11,5129	0

Para verificar o efeito de cada observação, foram ajustados modelos com algumas destas observações removidas e calculada a porcentagem da mudança relativa na estimativa do parâmetro (no caso do modelo Bayesiano, o Estimador de Bayes) através de $|(\hat{\theta} - \hat{\theta}^{[I]})/\hat{\theta}| \times 100\%$, em que $\hat{\theta}$ representa as estimativas com os dados completos e $\hat{\theta}^{[I]}$ sem um conjunto I de observações, denotando-se por MR(%).

Tabela 13: MR(%) para as observações da Tabela 12 no MRBS clássico.

Removida	β_0	β_1	β_2	δ
#33	8,15	19,69	43,12	18,09
#17	6,56	11,25	18,29	8,47
#31	2,95	7,64	19,44	6,15
#14	8,15	11,83	23,04	11,23
#15	8,15	11,83	23,04	11,23
#33, #17, #31	31,81	70,04	79,96	87,74
#33, #14, #15	16,36	17,72	94,03	62,99

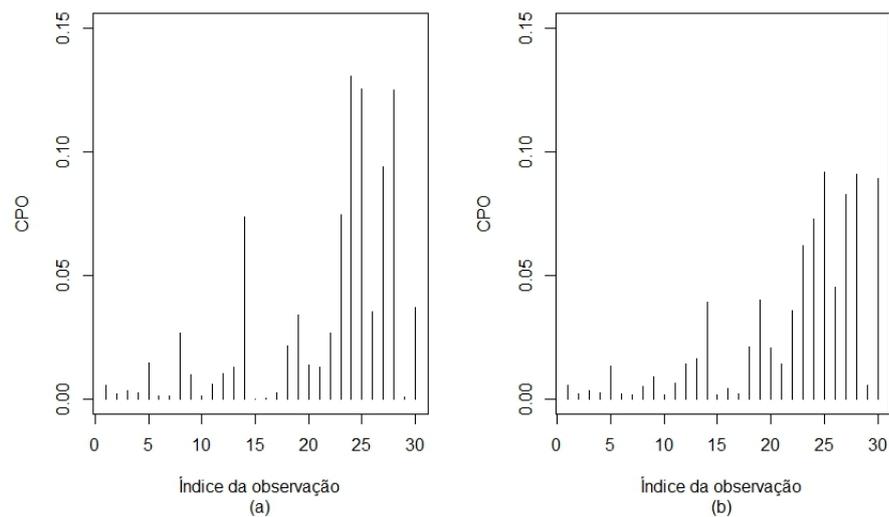
Pode-se observar na Tabela 13 que a retirada da observação #33 é a que, individualmente, causa as maiores mudanças nos parâmetros, chegando a 43% para β_2 . Outro fato interessante é que as observações #14 e #15 individualmente possuem um efeito maior na mudança das estimativas do que as observações #17 e #31, porém quando estes pares são retirados conjuntamente com #33, tem-se um resultado contrário, em que a retirada de #33, #17 e #31 tem um efeito consideravelmente maior do a retirada de #33, #14 e #15. Tal resultado pode explicar porque as observações #14 e #15 são consideradas mais influentes do que #17 e #31 na divergência K-L, pois de fato individualmente o são, porém, em conjunto com com #33, o efeito de maior influência fica como indicado para a análise de influência local, com o d_{max} , em que #17 e #31 são consideradas mais influentes. A mesma situação ocorre para os modelos Bayesianos que foram ajustados, como mostrado nos dados da Tabela 14.

Denotando como M_1 o modelo sob o caso C_1 , sem as observações #33, #17 e #31, e M_2 o modelo sob o caso C_2 , sem as observações #33, #14 e #15, os valores de CPO para cada observação sob cada modelo pode ser visto na Figura 12. Utilizando esses valores para calcular o Pseudo Fator de Bayes para compara os modelos M_1 e M_2 , tem-se que $PFB(M_1, M_2) < 0,0001$, indicando que o modelo M_2 é melhor que o M_1 , mesmo tendo sido percebido na Tabela 12 que tal modelo possui observações consideradas influentes.

Analisando os dados das Tabelas 15 e 16, é possível notar algumas diferenças entre

Tabela 14: MR(%) para as observações da Tabela 12 no MRBS Bayesiano.

Removida	β_1	β_2	β_3	δ
#33	7,30	18,04	41,62	18,01
#17	5,45	9,28	20,27	10,29
#31	2,59	7,40	24,04	7,13
#14	8,60	12,07	28,70	10,66
#15	8,60	12,07	28,70	10,66
#33,#17,#31	30,48	67,15	78,82	80,39
#33,#14,#15	16,75	18,52	95,62	60,19

Figura 12: Comparação dos valores de CPO nos modelos M_1 (a) e M_2 (b)Tabela 15: Estimativas dos casos C_1 e C_2 sob enfoque clássico

Parâmetros	C_1		C_2	
	Estimativas	DP	Estimativas	DP
β_0	10,481	1,243	6,651	1,352
β_1	-0,860	0,130	-0,416	0,140
β_2	1,212	0,322	1,307	0,341
δ	2,014	0,520	1,748	0,451

Tabela 16: Estimativas dos casos C_1 e C_2 sob enfoque Bayesiano

Caso Parâmetro	C_1				
	Média	DP	$P_{2,5}$	Mediana	$P_{97,5}$
β_1	10,451	1,179	8,189	10,460	12,750
β_2	-0,856	0,126	-1,095	-0,858	-0,608
β_3	1,198	0,323	0,573	1,195	1,813
δ	2,042	0,495	1,178	2,004	3,116
Caso Parâmetro	C_2				
	Média	DP	$P_{2,5}$	Mediana	$P_{97,5}$
β_1	6,668	1,332	4,144	6,640	9,361
β_2	-0,417	0,138	-0,688	-0,414	-0,149
β_1	1,310	0,326	0,663	1,311	1,945
δ	1,814	0,449	1,041	1,779	2,804

os dois casos. Para C_1 tem-se o intercepto maior dos obtidos com C_2 , indicando que sob a primeira situação os pacientes iniciam o estudo com um tempo de vida esperado maior do que estimado sob C_2 . Outra diferença notável entre ambos os casos é em relação ao efeito da variável $lWBC$ no valor esperado de *semanas*. Para o caso C_1 esse efeito é considerado maior que no caso C_2 .

Calculando os valores de AIC e BIC para os modelos comparados sob o enfoque clássico, o valor de AIC para os casos C_1 e C_2 são, respectivamente 253.7471 e 277.1419, enquanto que os valores de BIC para M_1 e M_2 são, respectivamente 259.3519 e 282.7467. Esses resultados indicam que, embora o modelo sob o caso C_1 tenha valores menores de AIC e BIC, a diferença entre deste para o caso C_2 não é tão expressiva nos modelos clássicos como nos modelos Bayesianos pelo PFB. Dessa forma, o modelo a ser adotado poderia ser o M_2 , em que as observações #33, #14 e #15 são retiradas, uma vez que este poderia ser considerado equivalente ao M_1 , no caso de se utilizar a abordagem clássica.

Com o resultado apresentado, os pontos a serem retirados diferem em relação às estimativas nas abordagens clássica e Bayesiana, no caso de serem retirados as 3 observações mais influentes. Porém, ambas metodologias mostraram uma relação de equivalência como observado na Figura 12, em que o conjunto de observações influentes é bastante semelhante (no caso dos 6 pontos mais influentes o resultado foi o mesmo, sendo destacados os pontos #33, #31, #17, #21, #14 e #15).

7 CONSIDERAÇÕES FINAIS

Neste trabalho foi apresentada a distribuição Birnbaum-Saunders e uma extensão dela para modelos de regressão. Uma forma reparametrizada da distribuição, foi discutida, bem como métodos de estimação dos parâmetros e propriedades, e posteriormente uma abordagem Bayesiana da distribuição foi proposta. Modelos de regressão criados utilizando a distribuição BSR foram apresentados, sendo discutidos seu desenvolvimento, métodos de estimação e análise de diagnóstico. Uma abordagem Bayesiana do MRBS também foi proposta, sendo desenvolvida uma abordagem diferente para modelar dados com o MRBS, discutindo métodos de diagnóstico e comparação de diferentes modelos sob o aspecto Bayesiano.

Um exemplo de aplicação foi apresentado, analisando a influência da presença do fator AG (bastão de Auer que pode estar contido em células da medula) e do valor de *IWBC* (logaritmo do número de células brancas) no tempo de vida de pacientes com um determinado tipo de leucemia. Diferentes modelos foram feitos sob as metodologias clássica e Bayesiana, verificando uma relação negativa entre o tempo esperado de vida dos pacientes e *IWBC*, enquanto que a presença do fator AG tende a aumentar este tempo. Neste exemplo de aplicação, uma comparação foi feita entre os métodos de diagnóstico de influência clássica e Bayesiana, influência local e divergência de K-L, respectivamente. Observou-se que ambas metodologias são equivalentes, porém, diferentes conclusões podem ser obtidas dependendo da abordagem adotada. Dessa forma, este trabalho apresentou duas diferentes formas de se utilizar a distribuição Birnbaum-Saunders reparametrizada em modelos de regressão, sendo mais uma ferramenta disponível na modelagem de dados.

Como propostas de trabalho futuro, um modelo semiparamétrico poderia ser utilizado no modelo, no qual a flexibilidade de modelos livres de distribuição poderia potencializar o MRBS. A abordagem Bayesiana para modelos de regressão não-paramétricos também poderia ser utilizada, sendo discutida métodos de inferência e diagnóstico para tal modelo.

REFERÊNCIAS BIBLIOGRÁFICAS

- ACHCAR, J. A. Inferences for the Birnbaum-Saunders fatigue life model using Bayesian methods. **Computational Statistics & Data Analysis**, v. 15, p.367-380, 1993.
- AHMED, S. E.; BUDSABA, K.; LISAWADI, S.; VOLODIN, A. Parametric Estimation for the Birnbaum-saunders Lifetime Distribution Based on a New Parametrization . **Thailand Statistician** , v. 6, p.213-240, 2008.
- ATKINSON, A. C. **Plots, Transformations and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis**. New York: Oxford University Press, 1985.
- BARROS, M.; PAULA, G. A.; LEIVA, V. An R implementation for generalized Birnbaum-Saunders distributions. **Computational Statistics and Data Analysis**, v. 53, p.1511-1528, 2009.
- BIRNBAUM, Z. W.; SAUNDERS, S. C. A new family of life distributions. **Journal of Applied Probability**, v. 6, p.319-327, 1969a.
- BIRNBAUM, Z. W.; SAUNDERS, S. C. Estimation for a Family of Life Distributions with Applications to Fatigue. **Journal of Applied Probability**, v. 6, p.328-347, 1969b.
- CANCHO, V. G.; ORTEGA, E. M. M.; PAULA, G. A. On estimation and influence diagnostics for log-Birnbaum-Saunders Student-t regression models: Full Bayesian analysis. **Journal of Statistical Planning and Inference**, v. 140, p.2486-2496, 2010.
- CHRISTENSEN, R.; JOHNSON, W. O.; BRANSCUM, A. J.; TIMOTHY, E. **Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians**. New York: Chapman and Hall, 2011.
- COOK, R. D. Assessment of local influence. **Journal of the Royal Statistical Society B**, v. 48, p.133-169, 1986.
- CORDEIRO, G. M.; DEMÉTRIO, C. G. B. **Modelos Lineares Generalizados**. Santa Maria: UFSM, 2007. (Minicurso para o 12^a SEAGRO e 52^a Reunião Anual da RBRAS)
- CHO, H.; IBRAHIM, J. G.; SINHA, D.; ZHU, H. Bayesian Case Influence Diagnostics for Survival Models. **Biometrics**, v. 65, p.116-124, 2009.
- COWLES, M. K. **Applied Bayesian Statistics**. New York: Springer, 2013.
- DÍAZ-GARCÍA, J. A.; LEIVA, V. A new family of life distributions based on the elliptically contoured distributions. **Journal of Statistical Planning and Inference**, v. 128, p.445-457, 2005.
- FARIAS, R. B. A.; LEMONTE, A. J. Bayesian inference for the Birnbaum-Saunders nonlinear regression model. **Statistical Methods & Applications**, v. 20, p.423-438, 2011.
- FEIGL, P.; ZELEN, M. Estimation of Exponential Survival Probabilities with Concomitant Information. **Biometrics**, v. 21,(4), p.826-838, 1965.

- GEISSER, S.; EDDY, W. F. A Predictive Approach to Model Selection. **Journal of the American Statistical Association**, v. 74,(365), p.153-160, 1979.
- GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. **Bayesian Data Analysis**. 2.ed. New York: Chapman & Hall, 2004.
- LEIVA, V.; SANTOS-NETO, M.; CYSNEIROS, F. J. A.; BARROS, M. E. Birnbaum-Saunders statistical modelling: a new approach. **Statistical Modelling**, v. 14,(1), p.21-48, 2014.
- LEITE, J. G.; SINGER, J. M. **Métodos assintóticos em Estatística**. São Paulo: USP, 1990. (Minicurso para o 9^a Simpósio Nacional de Probabilidade e Estatística)
- LEMONTE, A. J. **Inferência sobre os parâmetros da distribuição Birnbaum-Saunders bi-paramétrica**. 2006. Dissertação (Mestrado em Estatística)-Centro de Ciências Exatas e da Natureza, Universidade Federal de Pernambuco, 2006.
- MAGNUS, J. R.; NEUDECKER, H. **Matrix differential calculus with applications in statistics and econometrics**. 3.ed. New York: John Wiley & Sons, 2007.
- McCULLOCH, R. E. Local Model Influence. **Journal of the American Statistical Association**, v. 84,(406), p.473-478, 1989.
- MORGADO, A. C.; CARVALHO, J. B. P.; CARVALHO, P. C. P.; FERNANDEZ, P. **Análise combinatória e probabilidade**. 9. ed. Rio de Janeiro: SBM, 2006.
- NG, H. K. T.; KUNDU, D.; BALAKRISHNAN, N. Modified moment estimation for the two parameters Birnbaum-Saunders distribution. **Computational Statistics and Data Analysis**, v. 43, p.283-298, 2003.
- OWEN, W. J.; PADGETT, W. J. A Birnbaum-Saunders accelerated life model. **IEEE Transactions on Reliability**, v. 49, p.224-229, 2000.
- PATRIOTA, A. G. On scale-mixture Birnbaum Saunders distributions. **Journal of Statistical Planning and Inference**, v. 142, p.2221-2226, 2012.
- PAULA, G. A. **Modelos de Regressão com Apoio Computacional**. São Paulo: IME/USP, 2013.
- PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesianana**. Lisboa: Fundação Calouste Gulbenkian, 2003.
- THOMAS, A.; O'HARA, B.; LIGGES, U.; STURTZ, S. Making BUGS Open. **R News**, v.6, p.12-17, 2006. Disponível em: <<http://www.openbugs.net/w/Downloads>> Acesso em: julho de 2014.
- R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria, 2014. Disponível em: <<http://www.R-project.org/>> Acesso em: junho de 2014.
- RIECK, J. R.; NEDELMAN, J. R. A Log-Linear Model for the Birnbaum-Saunders Distribution. **Technometrics**, v. 33, p.51-60, 1991.
- ROSS, S. M. **A first course in probability**. 8. ed. New York: Prentice Hall, 2010.

SANTOS-NETO, M.; CYSNEIROS, F. J. A.; LEIVA, V.; AHMED, S. E. On New Parametrizations Of The Birnbaum-Saunders Distribution. **Pakistan Journal of Statistics**, v. 28,(1), p.1-26, 2012.

APÊNDICE A – FUNÇÃO DENSIDADE DE PROBABILIDADE

Através da relação entre a distribuição BS e a distribuição Normal, tem-se que

$$F(t) = \Phi \left[\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right] = \int_{-\infty}^{\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right)} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{z^2}{2} \right\} dz,$$

em que $Z \sim \mathcal{N}(0, 1)$, e derivando tal equação em relação a t obtém-se a função densidade de probabilidade de T ,

$$\begin{aligned} f(t) &= \frac{d}{dt} \int_{-\infty}^{\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right)} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{z^2}{2} \right\} dz \\ &= \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left[\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right]^2 \right\} \left[\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right]' \\ &= \frac{1}{\alpha\sqrt{2\pi}} \left[\frac{1}{2\sqrt{t\beta}} - \frac{\sqrt{\beta}}{2t^{3/2}} \right] \exp \left\{ -\frac{1}{2\alpha^2} \left(\frac{t}{\beta} + \frac{\beta}{t} - 2 \right) \right\} \\ &= \frac{t^{-3/2}(t+\beta)}{2\alpha\sqrt{2\pi\beta}} \exp \left\{ -\frac{1}{2\alpha^2} \left(\frac{t}{\beta} + \frac{\beta}{t} - 2 \right) \right\}, \quad t > 0 \end{aligned}$$

em que t' representa a primeira derivada de t .

APÊNDICE B – RELAÇÃO ENTRE AS DISTRIBUIÇÕES BS E NORMAL

Definindo a função

$$\zeta(t) = t^{1/2} - t^{-1/2}, \quad t > 0,$$

tem-se que $F_T(t) = \Phi\left(\frac{1}{\alpha}\zeta(T/\beta)\right)$ e $\frac{1}{\alpha}\zeta(T/\beta) \sim \mathcal{N}(0, 1)$.

Seja $X \sim \mathbb{N}\left(0, \frac{\alpha^2}{4}\right)$, pode-se ver que, em distribuição,

$$2X = \zeta(T/\beta).$$

Demonstração. Sabe-se que

$$f(x) = \frac{2}{\alpha\sqrt{2\pi}} \exp\left\{-\frac{2x^2}{\alpha^2}\right\}, \quad -\infty < x < \infty$$

Para $Y = 2X$,

$$X = \frac{Y}{2} = h(X) \quad \text{e} \quad dx = \frac{dy}{2}$$

Pelo método do jacobiano, pode-se obter

$$g(y) = f(h^{-1}(x)) \left| \frac{dx}{dy} \right| = \frac{1}{\alpha\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2\alpha^2}\right\} \rightarrow Y \sim \mathcal{N}(0, \alpha^2).$$

Assim, tem-se que $\frac{1}{\alpha}\zeta(t/\beta) \sim \mathcal{N}(0, 1)$, mas o interesse é em $\zeta(t/\beta)$. Tomando então $V = \zeta(t/\beta)$, tem-se

$$\frac{\zeta(t/\beta)}{\alpha} = \frac{V}{\alpha} \quad \text{e} \quad \left[\frac{\zeta(t/\beta)}{\alpha} \right]^{(1)} = \frac{dv}{\alpha}$$

Pelo jacobiano obtém-se

$$g(v) = \frac{1}{\alpha\sqrt{2\pi}} \exp\left\{-\frac{v^2}{2\alpha^2}\right\}, \quad -\infty < v < \infty$$

Portanto,

$$2X \stackrel{D}{=} \zeta(t/\beta)$$

□

Pode ser definida também uma função ρ , tal que

$$\rho(X) = X + \sqrt{1 + X^2}, \quad X > 0,$$

em que, para $y > 0$,

$$\begin{aligned} 2x &= \zeta(y) \\ 2x &= y^{1/2} - y^{-1/2} \\ 2xy^{1/2} &= y - 1 \\ y - 2xy^{1/2} - 1 &= 0, \end{aligned}$$

e resolvendo a equação do segundo grau,

$$\begin{aligned} \Delta &= (2x)^2 - 4(-1) = 4(x^2 + 1) \\ \sqrt{y} &= \frac{2x \pm \sqrt{4(x^2 + 1)}}{2} = x \pm \sqrt{x^2 + 1}, \end{aligned}$$

como $\sqrt{y} > 0$ e $x < \sqrt{1 + x^2}$, tem-se

$$\sqrt{y} = x + \sqrt{x^2 + 1} = \rho(x)$$

Com isso, definindo uma outra função Ψ que está relacionada com as duas anteriores, de forma que

$$\Psi(x) = [\rho(x)]^2 = y = \zeta^{-1}(2x).$$

Com esses resultados, tem-se que

$$2X = \zeta(T/\beta) \Rightarrow \zeta^{-1}(2X) = \frac{T}{\beta} \Rightarrow T = \beta \zeta^{-1}(2X) = \beta \Psi(X),$$

logo

$$T = \beta \Psi(X) = \beta [\rho(X)]^2 = \beta \left[1 + 2X^2 + 2X\sqrt{1 + X^2} \right].$$

em que $X \sim \mathcal{N}\left(0, \frac{\alpha^2}{4}\right)$.

A partir deste resultado pode ser obtida a equação (2.3),

$$T = \beta \left[\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1} \right]^2, \quad Z \sim \mathcal{N}(0, 1).$$

Essa relação se deve ao fato de que, se a distribuição de X for padronizada, será obtida a seguinte igualdade

$$Z = \frac{X - 0}{\alpha/2} = \frac{2X}{\alpha} \Rightarrow X = \frac{Z\alpha}{2}.$$

APÊNDICE C – MOMENTO DE ORDEM R DA DISTRIBUIÇÃO BS

Através da equação (2.3) é possível obter a expressão do r -ésimo momento de $T \sim BS(\alpha, \beta)$,

$$\mathbb{E}[T^r] = \mathbb{E} \left[\left(\beta \left[\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2} \right)^2 + 1} \right]^2 \right)^r \right] = \beta^r \mathbb{E} \left[\left(\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2} \right)^2 + 1} \right)^{2r} \right],$$

Utilizando o polinômio de Leibniz (MORGADO *et al*, 2006), tem-se que

$$\begin{aligned} \mathbb{E}[T^r] &= \beta^r \mathbb{E} \left[\sum_{k+j=2r} \frac{(2r)!}{k!j!} \left(\frac{\alpha Z}{2} \right)^k \left(\left(\frac{\alpha Z}{2} \right)^2 + 1 \right)^{\frac{j}{2}} \right] \\ &= \beta^r \mathbb{E} \left[\sum_{k=2r-j} \frac{(2r)!}{(2r-j)!j!} \left(\frac{\alpha Z}{2} \right)^{2r-j} \left(\left(\frac{\alpha Z}{2} \right)^2 + 1 \right)^{\frac{j}{2}} \right] \\ &= \beta^r \mathbb{E} \left[\sum_{j=0}^{2r} \binom{2r}{j} \left(\frac{\alpha Z}{2} \right)^{2r-j} \left(\left(\frac{\alpha Z}{2} \right)^2 + 1 \right)^{\frac{j}{2}} \right], \end{aligned}$$

substituindo j na expressão por $2m$ e trabalhando com o somatório, obtém-se

$$\mathbb{E}[T^r] = \beta^r \sum_{m=0}^r \binom{2r}{2m} \mathbb{E} \left[\left(\frac{\alpha Z}{2} \right)^{2(r-m)} \left(\left(\frac{\alpha Z}{2} \right)^2 + 1 \right)^m \right],$$

utilizando o polinômio de Leibniz novamente obtém-se

$$\begin{aligned} \mathbb{E}[T^r] &= \beta^r \sum_{m=0}^r \binom{2r}{2m} \mathbb{E} \left[\left(\frac{\alpha Z}{2} \right)^{2(r-m)} \sum_{n=0}^m \binom{m}{n} \left(\frac{\alpha Z}{2} \right)^{2n} \right] \\ &= \beta^r \sum_{m=0}^r \binom{2r}{2m} \sum_{n=0}^m \binom{m}{n} \mathbb{E} \left[Z^{2(r-m+n)} \right] \left(\frac{\alpha}{2} \right)^{2(r-m+n)}. \end{aligned}$$

Existe uma propriedade que permite calcular os r -ésimos momentos de uma distribuição normal, para $r > -1$, utilizando o fatorial duplo (ver Ross (2010)). Para a variável $Z \sim \mathcal{N}(0, 1)$ essa propriedade é da seguinte forma

$$\mathbb{E}[Z^r] = \begin{cases} 0 & \text{se } r \text{ for ímpar} \\ (r-1)!! & \text{se } r \text{ for par} \end{cases}$$

em que o duplo fatorial, denotado por $n!!$.

O fatorial duplo possui uma relação com o fatorial que é útil na definição da fórmula para o cálculo dos r -ésimos momentos de uma distribuição Birnbaum-Saunders. Sendo $n = 2k - 1$, essa relação é

$$n!! = (2k - 1)!! = \frac{(2k)!}{2^k k!}.$$

Dessa forma, com as propriedades anteriores tem-se que $E(T^r)$ pode ser reescrito da seguinte forma

$$\begin{aligned} \mathbb{E}[T^r] &= \beta^r \sum_{m=0}^r \binom{2r}{2m} \sum_{n=0}^m \binom{m}{n} \left(\frac{\alpha}{2}\right)^{2(r-m+n)} \mathbb{E}\left[Z^{2(r-m+n)}\right] \\ &= \beta^r \sum_{m=0}^r \binom{2r}{2m} \sum_{n=0}^m \binom{m}{n} \left(\frac{\alpha}{2}\right)^{2(r-m+n)} [2(r-m+n) - 1]!! \\ &= \beta^r \sum_{m=0}^r \binom{2r}{2m} \sum_{n=0}^m \binom{m}{n} \left(\frac{\alpha}{2}\right)^{2(r-m+n)} \frac{[2(r-m+n)]!}{2^{(r-m+n)}(r-m+n)!}. \end{aligned}$$

APÊNDICE D – CÓDIGOS DO PROGRAMA R PARA A DISTRIBUIÇÃO BSR

D.1 Modelos usando somente o R

```
#####
# Função para gerar valores aleatórios da BSR #####

rRBS=function(t,m=1,d=1){
  if (any(m <= 0)) stop(paste("mu deve ser positivo", "\n", ""))
  if (any(d <= 0)) stop(paste("delta deve ser positivo", "\n", ""))
  Z = rnorm(t)
  r = (m*d/(d+1))*(Z/sqrt(2*d) + sqrt((Z/sqrt(2*d))^2 + 1))^2
  r
}

#####
# Função acumulada da dist. BSR #####

pRBS=function(t,m=1,d=1){
  cdf = pnorm( sqrt(d/2)*( sqrt((t*(d+1))/(d*m))-sqrt((d*m)/(t*(d+1))) ) )
  cdf
}

#####
# Função densidade da dist. BSR #####

dRBS=function(t,m=1,d=1){
  if (any(m <= 0)) stop(paste("mu deve ser positivo", "\n", ""))
  if (any(d <= 0)) stop(paste("delta deve ser positivo", "\n", ""))
  fy = exp(d/2)*
    sqrt((1+d)/(16*pi*m))*
    (t^(-.5) + (t^(-1.5)*d*m)/(d+1))*
    exp( -(d/4)*( t*(d+1)/(d*m) + (d*m)/(t*(d+1)) ) )
  fy
}

#####
# Função quantil da dist. BSR #####

qRBS=function(q,m=1,d=1){
  if (any(m <= 0)) stop(paste("mu deve ser positivo", "\n", ""))
  if (any(d <= 0)) stop(paste("delta deve ser positivo", "\n", ""))
  if(q < 10^(-250)){resp = 0}
  else{
    q=(d*m/(d+1)) * (qnorm(q)/sqrt(2*d) + sqrt(1+(qnorm(q)/sqrt(2*d))^2))^2
  }
}
```

```

    }
  q
}

#####
# Função para o Estimador pelo método dos momentos modificados
# da dist. BSR

modmeRBS = function(t){

  th = mean(1/t)

  mu = mean(t)

  delta = 1/(sqrt(mu*th) - 1)
  list(mu=mu,delta=delta)
}

# t = rRBS(30,m=10,d=10)
# modmeRBS(t)

#####
# Função para o EMM da dist. BSR #####

mmeRBS = function(t){
  mu = mean(t)

  n = length(t)
  s2 = (n-1)*var(t)/n

  delta = ( mu^2 - s2 + sqrt(mu^4 + 3*s2*mu^2) )/s2
  list(mu=mu,delta=delta)
}

# t = rRBS(30,m=10,d=10)
# mmeRBS(t)

#####
# Função para a matriz de informação da dist. BSR #####

# Função para calcular integral da matriz de informação
Idt=function(m,d){
  if (any(m <= 0)) stop(paste("mu must be positive", "\n", ""))

```

```

if (any(d <= 0)) stop(paste("delta must be positive", "\n", ""))
f.Idt=function(t){
  exp(d/2)*
  sqrt((1+d)/(16*pi*m))*(t^(-1.5))*
  (t + (d*m)/(d+1))^-1) *
  exp( -(d/4)*( t*(d+1)/(d*m) + d*m/(t*(d+1)) ) )
}
integrate(f.Idt,0,Inf)$value
}

#####
# Função para o EMV da dist. Birnbaum-Saunders #####

mleRBS = function(y, n.iter=10){

# Derivadas da matriz de informação
dldm = function(mu,delta,y) sum( -1/(2*mu) + delta/(delta*y+y+delta*mu) +
(delta+1)*y/(4*mu^2) - (delta^2)/(4*y*(delta+1)) )
d2ldm2 = function(mu,delta,y) length(y)*( -delta/(2*mu^2) -
(delta^2)*Idt(mu,delta)/(delta+1)^2 )
dlld = function(mu,delta,y) sum( delta/(2*(delta+1)) + (y+mu)/(delta*y+y+
delta*mu) - y/(4*mu) - mu*delta*(delta+2)/(4*y*(delta+1)^2) )
d2ldd2 = function(mu,delta,y) length(y)*( -(delta^2 + 3*delta + 1)/(2*
(delta+1)^2 * delta^2) - (mu^2)*Idt(mu,delta)/((delta+1)^4) )
d2ldmdd = function(mu,delta,y) length(y)*( -1/(2*mu*(delta+1)) -
delta*mu*Idt(mu,delta)/((delta+1)^3) )

# Vetor escore
VetEsc = function(mu,delta,y) c(dldm(mu,delta,y), dlld(mu,delta,y))

# Matriz de informação de Fisher
MatInf = function(mu, delta, y) - matrix(c(d2ldm2(mu,delta,y),
d2ldmdd(mu,delta,y),d2ldd2(mu,delta,y)),2,2)

# Palpite inicial
B0 = c(modmeRBS(y)$mu,modmeRBS(y)$delta)

for(i in 1:n.iter){
B = B0 + solve(MatInf(B0[1],B0[2],y))%*%VetEsc(B0[1],B0[2],y)
B0 = B
}
list(mu=B0[1],delta=B0[2])
}

# t = rRBS(30,m=1,d=200)

```

```

# mleRBS(t,n.iter=20)

#####
# Modelos de regressão Birnbaum-Saunders #####

regRBS = function(y,Xev,n.it=20){

# Função de ligação logarítmica
g = function(x) log(x); dg = function(x) 1/x; gi = function(x) exp(x);
d2g = function(x) -1/(x^2)

# Função de ligação raiz
# g = function(x) sqrt(x); dg = function(x) 1/(2*sqrt(x)); gi = function(x) x^2

# Função de ligação inversa
# g = function(x) 1/x; dg = function(x) -1/(x^2); gi = function(x) 1/x

# Pacote necessário para o ginv
require(MASS)

# Valores iniciais para beta
beta = lm(g(y)~Xev)$coef

# Valores iniciais para delta
delta = mleRBS(y)$delta

# Vetor de parâmetros
theta = c(beta, delta)

# Matriz de modelo
X = cbind(rep(1,dim(as.matrix(Xev))[1]),as.matrix(Xev))

# Matriz de modelo aumentada
Xt = rbind(
cbind(X,rep(0,dim(X)[1])),
cbind(t(rep(0,dim(X)[2])),1)
)

# Iteração de estimação
for(i in 1:n.it){

```

```

# Vetor da média
mu = gi(X%%beta)

#####
z = -1/(2*mu) + delta/(delta*y+y+delta*mu) + (delta+1)*y/(4*mu^2) -
(delta^2)/(4*y*(delta+1))
#####
b = delta/(2*(delta+1)) + (y+mu)/(delta*y+y+delta*mu) - y/(4*mu) -
mu*delta*(delta+2)/(4*y*(delta+1)^2)
Db = diag(as.vector(b))
#####
a = 1/dg(mu)
Da = diag(as.vector(a))
#####
RI dt = sapply(mu, function(x) Idt(x,delta))
#####
v = delta/(2*(mu*dg(mu))^2) + (RI dt*delta^2)/(((delta+1)*dg(mu))^2)
Dv = diag(as.vector(v))
#####
s = 1/(2*mu*(delta+1)) + delta*mu*RI dt/(delta+1)^3
#####
u = (delta^2 + 3*delta + 1)/(2*(delta*(delta+1))^2) + (mu^2)*RI dt/(delta+1)^4
Du = diag(as.vector(u))
#####
Dab = rbind(
  cbind(Da,rep(0,dim(Da)[2])),
  cbind(t(rep(0,dim(Da)[1])),sum(diag(Db)))
)

# Matriz de pesos aumentada
W = rbind(
  cbind(Dv, Da%%s),
  cbind(t(s)%%Da, sum(diag(Du)))
)

##### Estimação

# Variável dependente
Z = Xt%%theta + ginv(W)%%Dab%%rbind(z,1)
theta.n = ginv(t(Xt)%%W%%Xt)%%t(Xt)%%W%%Z

# theta.n = theta + ginv(t(Xt)%%W%%Xt)%%t(Xt)%%Dab%%rbind(z,1)
beta = theta.n[1:length(theta)-1]
delta = theta.n[length(theta)]
theta = theta.n

```

```

}

#####
dmu2 = (1/(2*mu^2) - (delta^2)/(delta*y+y+delta*mu)^2 - y*(delta+1)/(2*mu^3))
cv = dmu2*(a^2) + z*a
Dc = diag(as.vector(cv))
#####
m = y/(delta*y+y+delta*mu)^2 + y/(4*mu^2) - delta*(delta+2)/(4*y*(delta+1)^2)
#####
d = 1/(2*(delta+1)^2) - ((y+mu)^2)/(delta*y+y+delta*mu)^2 - mu/(2*y*(delta+1)^3)
Dd = diag(as.vector(d))

# Matriz hessiana
L = rbind(
  cbind(t(X)%*%Dc%*%X, t(X)%*%Da%*%m),
  cbind(t(m)%*%Da%*%X, sum(diag(Dd)))
)

# Matriz de perturbação (ponderação de casos)
Mat.Pert = rbind(
  t(X)%*%Da%*%diag(as.vector(z)),
  t(b)
)

# Matriz de variância-covariância
K = ginv(t(Xt)%*%W%*%Xt)

# Resíduo do tipo Pearson padronizado
Vy = (mu^2)*(2*delta+5)/(delta+1)^2
resid = (y-mu)/sqrt(Vy)

# Critérios de informação de Akaike e Bayesiano
logLik.i = function(mu,delta,y){ delta/2 - .5*log(delta+1) -
.5*log(16*pi*mu) - 1.5*log(y) +
log(delta*y+y+delta*mu) - y*(delta+1)/(4*mu) -
(mu*delta^2)/(4*(1+delta)*y) }
logLik = sum(sapply(1:length(y), function(x) logLik.i(mu[x],delta,t[x])))
AIC = -2*logLik + 2*length(theta)
BIC = -2*logLik + log(length(y))*length(theta)

return(list(beta = beta,delta = delta, vcov = K,
resid = resid, ajustados = mu, Mat.Hes = L,

```

```

Mat.Pert = Mat.Pert, AIC=AIC, BIC=BIC))

}

#####
##### Envelope simulado para o MRBS #####

envelope.MRBS = function(y,Xev, mod.BSR){

n = length(y)
d = sort(mod.BSR$resid) # resíduos ordenados do modelo
z = qnorm((c(1:n)-.375)/(n+.25)) # quantis da normal padrão

# Matriz para guardar amostras geradas sob o modelo ajustado
Amostras = matrix(NA, 19, n)
for(i in 1:19){
Amostras[i,] = rRBS(n,m=mod.BSR$ajustados[i],d=mod.BSR$delta)
}

# Resíduos dos modelos das amostras geradas
Res = matrix(NA,n,19)
for(i in 1:19){
ajust.envel = regRBS(Amostras[i,], Xev)
Res[,i] = ajust.envel$resid
}

Res2 = apply(Res,2,sort) # resíduos ordenados dos modelos

# Construção do gráfico de envelope simulado
Env = cbind(apply(Res2,1,max),apply(Res2,1,mean),d,apply(Res2,1,min))
matplot(z,Env,lty=c(1,2,0,1),type=c("l","l","b","l"),
col=c(1,1,1,1),pch=19,ylab="Resíduos de Pearson",xlab="Quantil da Normal Padrão")
}

```

D.2 Modelo usando conjuntamente o OpenBUGS pelo R2OpenBUGS

```

# MBSR Bayesiano

# Pacote para utilizar o OpenBUGs com o R
library(R2OpenBUGS)

# Observações
y = semanas; x1 = 1.WBC; x2 = AG

```

```

N = length(y)
data = list("N","x1","x2","y")
# valores iniciais
inits = list(list(delta = 1.072715,beta = c(7.9516005,-0.5058312,0.6736365)))
# parâmetros do modelo
parameters = c("beta","delta")

# Modelo no OpenBUGs
modBayes = bugs(data = data, inits = inits, parameters.to.save = parameters,
model.file = "MRBS-sobrev.odc", n.chains = 1, n.iter = 10000,n.burnin = 1000)

# Resultados
modBayes$summary
modBayes$summary[,c(1,2,3,5,7)]

names(modBayes)
dim(modBayes$sims.matrix)
head(modBayes$sims.matrix)
p = dim(modBayes$sims.matrix)[2] - 1 # número de parâmetros

# Cálculo do CPO

# matriz para guardar as amostras dos parâmetros
sim.par.lnorm = modBayes$sims.matrix[,1:p]
dim(sim.par.lnorm)

# Função para calcular a média estimada
mu.estim = function(beta,x) exp(beta%*%x)

#matriz para guardar inversos das probabilidades
sim.inv.lnorm = matrix(NA,dim(sim.par.lnorm)[1],N)
X = cbind(rep(1,length(1.WBC)),1.WBC,AG) #matriz de modelo
for(j in 1:N){
sim.inv.lnorm[,j] = mapply(function(x,m,d) 1/dRBS(x,m,d),
y[j],mu.estim(sim.par.lnorm[,1:(p-1)],X[j,]),sim.par.lnorm[,p])
}
dim(sim.inv.lnorm)

# cálculo da média harmônica para obter o CPO
cpo.lnorm <- 1/apply(sim.inv.lnorm, 2, mean)

# Gráfico: CPO x índice (verifica qualidade de ajuste de cada observações)
plot(1:N, cpo.lnorm, type = "h", xlab = "Índice da observação", ylab = "CPO")

# divergência de Kullback-Leibler
KL.lnorm = -log(cpo.lnorm) + colMeans(log(1/sim.inv.lnorm))

```

```
plot(1:N, KL.lnorm, type = "h", xlab = "Índice da observação", ylab = "K-L")
identify(1:N, KL.lnorm, n=6)

# Calibração
calib.lnorm = .5*(1+sqrt( 1-exp(-2*KL.lnorm) ))
plot(1:N, calib.lnorm, type = "h", xlab = "Índice da observação", ylab = "calibração")
identify(1:N, calib.lnorm, n=6)
```

APÊNDICE E – CÓDIGOS DO PROGRAMA OPENBUGS PARA A DISTRIBUIÇÃO BSR

```

# Modelos de regressão da BSR

model {
  for (i in 1:N) {
    dummy[i] <- 0
    dummy[i] ~ dloglik(logLike[i])
    # log(likelihood)
    logLike[i]<- 0.5 * (delta + log((delta+1) /
    (16*PI*mu[i]))) ) - 1.5*log(y[i]) + log( y[i] +
    delta*mu[i]/(delta+1) ) - (delta + 1)*y[i]/(4*mu[i]) -
    delta*delta*mu[i]/(4*(delta+1)*y[i])

    # função de ligação logarítmica para a média
    mu[i] <- exp(beta[1] + beta[2]*x1[i] + beta[3]*x2[i])
  }

  # Priors normal para os coeficientes de regressão
  for(j in 1:3){
    beta[j] ~ dnorm(0, 0.001)
  }
  # delta ~ dunif(0, 50)
  # delta ~ dgamma(0, 0.1)
  delta ~ dlnorm(1.9561, 1.4426)

  PI <- 3.141592

}

```