

Universidade Federal de São Carlos
Centro de Ciências Exatas e de Tecnologia
Departamento de estatística

Algumas Extensões da Distribuição
Birnbaum-Saunders: Uma Abordagem
Bayesiana

Edwin Chaiña Cahui

São Carlos/SP
Janeiro/2012

Universidade Federal de São Carlos
Centro de Ciências Exatas e de Tecnologia
Departamento de Estatística

Algumas Extensões da Distribuição
Birnbaum-Saunders: Uma Abordagem
Bayesiana

Edwin Chaiña Cahui

Orientador: Prof.Dr. Vicente Garibay Cancho

Dissertação apresentada
ao Programa de Pós-
graduação em Estatística,
para obtenção do título de
mestre em estatística

São Carlos/SP

Janeiro/2012

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

C132ae

Cahui, Edwin Chaiña.

Algumas extensões da distribuição Birnbaum-Saunders :
uma abordagem Bayesiana / Edwin Chaiña Cahui. -- São
Carlos : UFSCar, 2012.

70 f.

Dissertação (Mestrado) -- Universidade Federal de São
Carlos, 2012.

1. Estatística. 2. Inferência bayesiana. 3. Tempo de
Fadiga. 4. Kullback-Leibler. 5. Modelo de mistura. I. Título.

CDD: 519.5 (20^a)

Edwin Chaiña Cahui

**ALGUMAS EXTENSÕES DA DISTRIBUIÇÃO BIRNBAUM–SAUNDER:
UMA ABORDAGEM BAYESIANA**

Dissertação apresentada à Universidade Federal de São Carlos, como parte dos requisitos para obtenção do título de Mestre em Estatística.

Aprovada em 09 de janeiro de 2012.

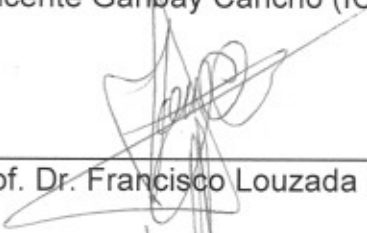
BANCA EXAMINADORA

Presidente




Prof. Dr. Vicente Garibay Cancho (ICMC-USP/Orientador)

1º Examinador



Prof. Dr. Francisco Louzada Neto (ICMC-USP)

2º Examinador



Prof. Dr. Victor Hugo Lachos Dávila (UNICAMP)

Agradecimentos

Primeiramente agradeço a Deus, por tudo que ele me deu. Ao meu orientador Prof. Dr. Vicente Garibay Cancho, pela paciente tolerância, e por estar sempre tão solícito para dirimir toda e qualquer dúvida, na elaboração, condução, sugestões oferecidas durante o processo de elaboração deste trabalho.

A meus pais Guillermo e Lucia que me deram não somente a vida, mas principalmente a minha educação, condições de estudo e apoio incondicional. A minha irmã Ana Maria meu irmão Guillermo.

Aos professores do departamento de Estatística da Universidad Federal de São Carlos que abriram as portas e me ofereceram o ambiente acolhedor e sadio, para que eu pudesse realizar o meu mestrado, em especial a professora Cristina pela atenção e ajuda nos momentos que precisava.

Ao pessoal da Secretaria Acadêmica, pela eficiência, dedicação e simpatia. Meu muito obrigado.

A todos amigos e amigas, os agradeço sinceramente pela amizade, pela preocupação e por todos os momentos divertidos. Agradeço ainda a amiga Alessandra por toda sua colaboração neste trabalho.

A todas as pessoas que, direta ou indiretamente, contribuíram para a realização dessa dissertação.

Finalmente, agradeço à CAPES pelo auxílio concedido para este trabalho.

Resumo

A distribuição Birnbaum-Saunders (BS) está baseada em um argumento físico de dano cumulativo que produz a fadiga de materiais. Esta fadiga foi identificada como uma importante causa de falhas em estruturas de engenharia. Nos últimos tempos, este modelo tem sido aplicado em outras áreas, tais como: ciências da saúde, ambientais, florestais, demográficas, financeiras, entre outras. Devido a sua importância, várias distribuições têm sido propostas para descrever o comportamento da resistência à fadiga. Entretanto não há um argumento sobre qual modelo é mais efetivo para a análise dos dados de fadiga. Um dos principais problemas para escolher uma distribuição estatística, é que frequentemente vários modelos ajustam os dados bem na parte central, porém, no entanto, os extremos da distribuição colocam em dúvida a decisão para selecionar alguns dos modelos propostos. A falta de dados nos extremos da distribuição justifica considerar outros argumentos como o uso de um modelo estatístico específico, e assim rejeitar outros modelos. Neste trabalho estudamos algumas extensões da distribuição Birnbaum-Saunders com mistura de escala normal, no qual procedimento para obtenção de inferências será considerado sob uma perspectiva Bayesiana baseada em Métodos de Monte Carlo via Cadeias de Markov (MCMC). Para detectar possíveis observações influentes nos modelos considerados, foi usado o método Bayesiano de análise de influência caso a caso, baseado na divergência de Kullback-Leibler. além disso, é proposto o modelo geométrico Birnbaum-Saunders, para dados de sobrevivência.

Palavras-chave: Inferência Bayesiana, Kullback-Leibler, Tempo de fadiga, Modelo de mistura.

Abstract

The Birnbaum-Saunders Distribution is based on an physical damage that produces the cumulative fatigue materials, This fatigue was identified as an important cause of failure in engineering structures. Recently, this model has been applied in other areas such as health sciences, environmental measures, forestry, demographic, financial, among others. Due to it's importance several distributions have been proposed to describe the behavior of fatigue resistance. However there is not an argument about which is more effective for the analysis of data from fatigue. A major problem to choose a statistical distribution, is that often several models fit the data well in the central, but, however, the extremes of distribution raise questions about the decision to select some of their models. The lack of data at the extremes distribution is justified to consider other arguments like the use of a specific statistical distribution, and thus reject other models. In this work we study some extensions of the distribution Birnbaum-Saunders with a mixture of normal scale, in which the procedure will for obtaining inferences will be considered from a Bayesian perspective based on the methods Monte Carlo Markov Chain (MCMC). to detect possible observations influential in the models considered, we used the Bayesian method of analysis influence in each case based on the Kullback-Leibler divergence. Moreover, the geometric Birnbaum-Saunders model is proposed , for data survival.

Keywords: Bayesian inference, Fatigue time, Kullback-Leibler, Mixture model.

Sumário

Sumário	iii
Lista de Tabelas	v
Lista de Figuras	vii
1 Introdução	1
1.1 Distribuição seno hiperbólico normal	5
1.2 A distribuição log-Birnbaum-Saunders	7
1.3 Critério de seleção de modelos	7
1.3.1 Critério da Densidade Preditiva Ordenada (CPO)	8
1.3.2 Critério de Informação de Desvio (DIC)	8
1.4 Diagnóstico	10
1.4.1 Análise de influência	11
1.5 Objetivos do trabalho	12
1.6 Apresentação dos capítulos	12
2 Modelo Birnbaum-Saunders com mistura de escala normal	14
2.1 Modelo Birnbaum-Saunders com mistura de escala normal	16
2.1.1 Exemplos da distribuição BSSMN	18
2.2 Distribuição log BSSMN	22
2.2.1 Exemplos da distribuição log BSSMN	23
3 Inferência Bayesiana para modelos de regressão log BSSMN	27
3.1 Distribuições a priori, a posteriori e conjunta	28
3.2 Aplicação	32

3.3	Modelos não lineares	42
3.3.1	Modelo não linear log-Birnbaum-Saunders com mistura de escala normal log BSSMN	43
3.4	Aplicação	45
3.5	Conclusão e observações	54
4	Modelo de regressão geométrico Birnbaum-Saunders	55
4.1	Formulação do problema	55
4.2	Modelo de regressão log GBS	59
4.3	Inferência Bayesiana para o modelo LGBS	61
4.4	Aplicação linear	62
5	Considerações Finais	66
	Referências Bibliográficas	67

Lista de Tabelas

3.1	Sumário <i>a posteriori</i> do modelo de regressão linear log BS para os dados simulados.	34
3.2	Sumário <i>a posteriori</i> do modelo de regressão linear log BS-t para os dados simulados.	34
3.3	Critério de comparação de modelos para os dados simulados. . .	35
3.4	Medida de divergência de Kullback-Leibler e suas calibrações para log BS e log BS-t, para os dados simulados do modelo de regressão linear.	35
3.5	Dados fadiga biaxial (Bronw & Miller (1974)).	39
3.6	Resumo <i>a posteriori</i> dos parâmetros para os dados de fadiga biaxial.	39
3.7	Sumário <i>a posteriori</i> do modelo de regressão não linear log BS para dados simulados.	46
3.8	Sumário <i>a posteriori</i> do modelo de regressão não linear log BS-t para dados simulados.	47
3.9	Critério de comparação de modelos para os dados simulados, do modelo não linear.	47
3.10	Medida de divergência de Kullback-Leibler e suas calibrações para dados simulados do modelo de regressão não linear.	50
3.11	Resumo <i>a posteriori</i> dos parâmetros para os dados de fadiga biaxial para o modelo de regressão não linear.	52
3.12	Medida de de divergência K-L, para o modelo não linear	52
4.1	Valores Esperados dos erros do modelo de regressão LGBS.	61
4.2	Resumo <i>a posteriori</i> LGBS modelo linear.	63

4.3	Medida de de divergência K-L.	63
4.4	Resumo <i>a posteriori</i> LGBS não linear.	64
4.5	Medida de de divergência K-L, não linear.	64

Lista de Figuras

1.1	Densidade da distribuição Birnbaum-Saunders com (a) β fixo; (b) α fixo.	4
1.2	Densidade seno hiperbólico normal (SHN)	7
2.1	Densidade da distribuição Birnbaum-Saunders-t com (a) $\alpha=0,1$; $0,3$; $1,0$ e β, ν fixos; (b) α e β fixos, $\nu=2$; 3 ; 6	19
2.2	Densidade da distribuição Birnbaum-Saunders-Normal Conta- minada com (a) α, β, ν fixos e $\nu=0,2$; $0,4$; $0,8$; (b) α, β, ν , fixos $\nu=0,2$; $0,5$; $0,8$	20
2.3	Densidade da distribuição Birnbaum-Saunders slash com (a) $\alpha = 0, 2; 0, 4; 0, 6$ e β, ν fixos; (b) α e β fixos e $\nu = 1; 2; 3$	21
2.4	(a) Densidades das distribuições BSSMN; (b) BSSMN com α , β fixos e com diferentes valores para ν	21
2.5	Densidade da distribuição log Birnbaum-Saunders-t com (a) $\alpha = 0,3; 0,5; 0,8$; e γ, ν fixos; (b) α, γ , fixos e $\nu=2$; 5 ; 10	24
2.6	Densidade da distribuição log Birnbaum-Saunders slash com (a) $\alpha=0,3; 0,5; 0,8$ e γ, ν fixos; (b) α, γ fixos, e $\nu=2$; 3 ; 8	26
2.7	Densidade da distribuição log BSNC com (a) α, γ, u fixos e $\nu=0,3; 0,5; 0,7$; (b) α, γ, ν fixos e $u=0,1; 0,3; 0,6$	26
3.1	Medida de divergência de Kullback-Leibler para os diferentes casos simulados para o modelo log BS.	36
3.2	Medida de divergência de Kullback-Leibler para os diferentes casos simulados para o modelo log BS-t.	37
3.3	Medida de divergência K-L para os dados de fadiga, (a) log BS, (b) log BS-t.	40

3.4	Medida de divergência K-L para os dados de fadiga, (a) log BSSL, (b) log BSNC.	40
3.5	(a) Diagrama de dispersão dos dados, (b) Diagrama de dispersão transformada com os ajustes do modelo de regressão estimada dos modelos log BSSMN.	41
3.6	Medida de divergência de Kullback-Leibler para os dados simulados para o modelo não linear, (a) log BS (SP),(b) log BS-t (SP) ;(c) e (e) log BS; (d) e (f) log BS-t	48
3.7	Medida de divergência de Kullback-Leibler para os dados simulados com as observações 1, 6, e 11 perturbadas, para o modelo não linear, (a) log BS, (b) log BS-t.	49
3.8	Medida de divergência de Kullback-Leibler para os dados simulados com as observações 1; 11; 27; 30 e 36 perturbadas, para o modelo não linear, (a) Log BS, (b) log BS-t.	49
3.9	(a) Medida de divergência K-L para os dados da aplicação usando log BS, (b) log BS-t; para o modelo de regressão não linear. . .	53
3.10	Ajuste dos dados de fadiga biaxial.	53
4.1	Densidade da distribuição GBS com $\theta=0; 0,1; 0,3; 0,6; 0,8; 0,9;$ $\beta=5; \alpha=0,2$ (gráfico (a)), $\alpha=2$ (gráfico (b)).	57
4.2	Função de risco da distribuição GBS. Os parâmetros são fixados com $\theta=0; 0,1; 0,3; 0,6; 0,8; 0,9$ e $\beta=5, \alpha=0,2$ (gráfica (a)); $\alpha=2$ (gráfica (b))	58
4.3	(a) Gráfico da curtose com $\mu = \log(5)$ fixo e $\theta=0; 0,1; 0,3; 0,6;$ $0,8; 0,9;$ (b) Skewness com parâmetro $\mu=\log 5$ fixo.	59
4.4	Densidade da distribuição LGBS.	60

Capítulo 1

Introdução

O desgaste de material é de grande importância em diferentes áreas de engenharia e um dos grandes problemas na área industrial é a ruptura ou falha dos materiais. Tais falhas são podem ser ocasionadas por diversas causas e uma destas causas é a fadiga.

A fadiga consiste na falha estrutural que resulta quando o material é submetido a diferentes esforços dinâmicos repetidos ou oscilações dinâmicas. Existem vários modelos estatísticos que podem ser usados para descrever a variação aleatória dos tempos de falhas associados aos materiais expostos à fadiga sob diferentes padrões. Distribuições que se ajustam bem na região central da distribuição de tempos de vida como gama, normal inversa, log normal e Weibull são bastantes usadas para descrever o comportamento dos tempos de vida por fadiga. No entanto, essas distribuições não apresentam bom desempenho quando o interesse está nos percentis extremos.

Motivados por problemas de vibração encontrados em aviões comerciais novos e falhas de material, Birnbaum & Saunders (1969) derivaram uma nova família de distribuições que modela o tempo de vida de materiais e equipamentos submetidos a cargas dinâmicas.

Considere um material que é sujeito a um padrão cíclico de tensão e força. A falha do material ocorre devido ao desenvolvimento e ao crescimento de

uma rachadura dominante dentro do material. Um ciclo é definido como "m" oscilações e cada aplicação da i -ésima oscilação em um ciclo, resulta em uma extensão aleatória da rachadura X_i . A distribuição desta variável aleatória depende somente da rachadura atual causada pela oscilação de carga neste ciclo. A extensão da rachadura devido ao j -ésimo ciclo é dada por,

$$Y_j = \sum_{i=1}^m X_i,$$

em que Y_j representa a extensão da rachadura devido ao j -ésimo ciclo, sendo uma variável aleatória independente e identicamente distribuída com média μ e variância σ^2 para $j = 1, 2, \dots, P$ e X_i é a extensão aleatória da rachadura na i -ésima oscilação. Considere a variável aleatória W_z como uma extensão total da rachadura após z ciclos, dada por

$$W_z = \sum_{j=1}^z Y_j,$$

com função de distribuição acumulada dada por,

$$H_z(w) = P(W_z \leq w), \quad \text{para } z = 1, 2, 3, \dots$$

Seja C o número de ciclos até que seja observada a falha com w o comprimento crítico da rachadura. A falha ocorre quando o comprimento da rachadura dominante excede o comprimento crítico w , isto é,

$$P(C \leq z) = P\left(\sum_{j=1}^z Y_j \geq w\right) = 1 - H_z(w).$$

A função de distribuição acumulada (f.d.a.) de C pode ser aproximada usando o Teorema Central do Limite, isto é,

$$\begin{aligned} P(C \leq z) &= P\left(\sum_{j=1}^z Y_j \geq w\right) \\ &= 1 - P\left(\sum_{j=1}^z \frac{Y_j - \mu}{\sigma\sqrt{z}} < \frac{w - z\mu}{\sigma\sqrt{z}}\right) \\ &= \Phi\left(\frac{\mu\sqrt{z}}{\sigma} - \frac{w}{\sigma\sqrt{z}}\right) \end{aligned}$$

$$= \Phi \left(\frac{\sqrt{w\mu}}{\sigma} \left[\frac{\sqrt{\mu}\sqrt{z}}{\sqrt{w}} - \frac{\sqrt{w}}{\sqrt{\mu}\sqrt{z}} \right] \right)$$

em (1.1) se z é substituído por uma variável real não negativa T , e se definimos $\alpha = \frac{\sigma}{\sqrt{\mu w}}$ e $\beta = \frac{w}{\mu}$, a variável aleatória T é a extensão contínua da variável aleatória discreta C . Então T representa o tempo total até que ocorra a falha. Birnbaum & Saunders (1969) mostraram, que sob certas condições, a função de distribuição acumulada da variável aleatória T é dada por:

$$F_T(t) = P(T \leq t) = \Phi \left(\frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right] \right), \quad (1.1)$$

no qual dizemos que T tem uma distribuição Birnbaum-Saunders, com parâmetros α e β e denotamos esta distribuição por $T \sim BS(\alpha, \beta)$. Uma propriedade importante na construção desta distribuição é que a variável aleatória

$$Z = \alpha^{-1} \left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right) \sim N(0, 1).$$

Assim, a distribuição Birnbaum-Saunders, está estritamente relacionada com a distribuição normal pela equação

$$T = \frac{\beta}{4} \left(\alpha Z + \sqrt{\alpha^2 Z^2 + 4} \right)^2, \quad (1.2)$$

sendo que a variável Z tem distribuição normal-padrão. A função de distribuição de probabilidade (f.d.p.) da variável aleatória T é dada por

$$f_T(t) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2\alpha^2} \left[\frac{t}{\beta} + \frac{\beta}{t} - 2 \right] \right) \frac{t^{-\frac{3}{2}} [t + \beta]}{2\alpha\sqrt{\beta}}, \quad t > 0, \quad (1.3)$$

em que $\alpha > 0$ é o parâmetro de forma, $\beta > 0$ é o parâmetro de escala. A Figura 1.1 apresenta algumas formas da f.d.p da distribuição BS. Nota-se na Figura 1.1 (a), que com o aumento no valor do parâmetro de α a forma da distribuição $BS(\alpha, \beta)$ torna-se assimétrica.

Algumas propriedades da distribuição BS são dadas na seguinte proposição:

Proposição 1.1. *Se $T \sim BS(\alpha, \beta)$, então,*

- $1/T \sim BS(\alpha, \beta^{-1})$;

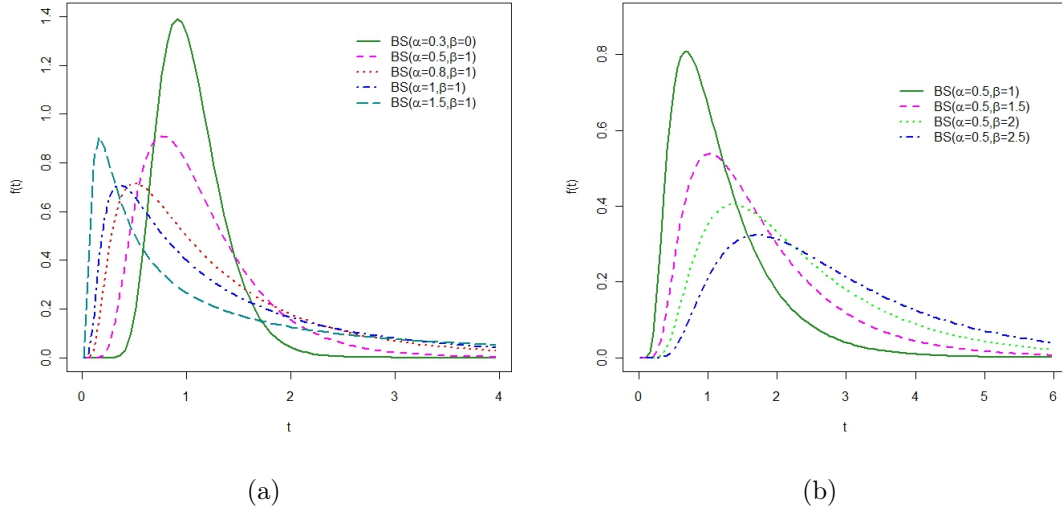


Figura 1.1: Densidade da distribuição Birnbaum-Saunders com (a) β fixo; (b) α fixo.

- $aT \sim BS(\alpha, a\beta)$, para todo $a > 0$.
- O r -ésimo momento da distribuição BS é dado por,

$$E(T^r) = \beta \sum_{j=0}^r \binom{2r}{2j} \sum_{i=0}^j \binom{j}{i} \frac{(2(r-j+i))!}{2^{r-j+i}(r-j+i)!} \left(\frac{\alpha}{2}\right)^{2(r-j+i)}. \quad (1.4)$$

A prova da Proposição 1.1 pode ser encontrada em Balakrishnan *et al.* (2009).

Da equação (1.4), temos que a esperança e a variância da distribuição BS com parâmetros α e β são dadas por,

$$E(T) = \beta \left(1 + \frac{\alpha^2}{2}\right) \text{ e } Var(T) = \beta^2 \alpha^2 \left(\frac{5}{4}\alpha^2 + 1\right).$$

Nos últimos 30 anos, surgiram algumas generalizações e aplicações da distribuição Birnbaum-Saunders. Desmond (1985) aprimorou a justificativa física usada para a construção dessa distribuição, flexibilizando algumas suposições feitas por Birnbaum & Saunders (1969), além de apresentar uma derivação

mais geral baseado em um modelo biológico. Desmond (1986) estabeleceu uma conexão entre a distribuição Gaussiana inversa e a distribuição Birnbaum Saunders. Rieck & Nedelman (1991) propuseram um modelo log-linear para a distribuição Birnbaum-Saunders e desenvolveram métodos de estimação. Achar & Espinosa (1992) desenvolveram métodos bayesianos em teste de vida acelerados considerando um modelo log-linear para a distribuição Birnbaum-Saunders. Owen & Padgett (1999) desenvolveram a distribuição Birnbaum Saunders com três parâmetros. Diaz-Garcia & Leiva-Sanchez (2005) propuseram uma classe de modelos de distribuição de vida generalizando a distribuição Birnbaum-Saunders a partir da distribuição de contornos elípticos. Os autores argumentam que essa generalização baseia-se na busca de distribuições de tempo de vida que cresçam mais rapidamente e que possuam caudas com menor ou maior curtose do que a distribuição Birnbaum-Saunders. Além disso, essa distribuição Birnbaum-Saunders generalizada considera como caso particular a distribuição Birnbaum Saunders e aplicações da distribuição Birnbaum Saunders generalizada são apresentadas por Leiva *et al.* (2008a,b). Vilca-Labra & Leiva (2006) baseados em argumentos semelhantes aos considerados por Diaz-Garcia & Leiva-Sanchez (2005), obtiveram uma outra generalização ao desenvolver a distribuição Birnbaum-Saunders mediante distribuições elípticas assimétricas. Assim, a distribuição Birnbaum-Saunders assimétrica é mais flexível na curtose e assimetria do que todas as versões anteriores da distribuição Birnbaum-Saunders. Recentemente, Balakrishnan *et al.* (2009) consideram estimação na distribuição Birnbaum-Saunders baseada em misturas de escala normal, usando o algoritmo EM.

Rieck & Nedelman (1991) propuseram a família de distribuição seno hiperbólico, que está relacionada com a distribuição Birnbaum-Saunders e tem um papel fundamental em modelos log-lineares, que detalhamos a seguir.

1.1 Distribuição seno hiperbólico normal

Uma variável aleatória Y tem distribuição seno hiperbólico com parâmetros $\alpha > 0$, $\gamma \in \mathfrak{R}$ e $\sigma > 0$ se sua f.d.p. é dada por (Rieck & Nedelman, 1991)

$$f_Y(y) = \phi\left(\frac{2}{\alpha} \sinh\left(\frac{y-\gamma}{\sigma}\right)\right) \frac{2}{\alpha\sigma} \cosh\left(\frac{y-\gamma}{\sigma}\right), \quad (1.5)$$

sendo $\phi(\cdot)$ a função de densidade da distribuição de probabilidade da normal padrão. A notação $Y \sim SHN(\alpha, \gamma, \sigma)$ indica que a variável aleatória Y tem distribuição seno hiperbólico normal com parâmetros $\gamma \in \mathfrak{R}$, $\sigma > 0$ e $\alpha > 0$.

A seguinte proposição mostra uma forma alternativa de obter a distribuição seno hiperbólico.

Proposição 1.2. *Se $Z \sim N(0, 1)$, então a variável aleatória*

$$Y = \gamma + \sigma \operatorname{arcsinh}\left(\frac{\alpha Z}{2}\right) \quad (1.6)$$

tem distribuição seno hiperbólico com parâmetros $\gamma \in \mathfrak{R}$, $\sigma > 0$ e $\alpha > 0$. Isto é, $Y \sim SHN(\alpha, \gamma, \sigma)$.

Prova: Consideremos a f.d.a. da variável aleatória dada em (1.6)

$$\begin{aligned} F_Y(y) = P(Y \leq y) &= P\left(\gamma + \sigma \operatorname{arcsinh}\left(\frac{\alpha Z}{2}\right) \leq y\right) \\ &= P\left(Z \leq \frac{2}{\alpha} \sinh\left(\frac{y-\gamma}{\sigma}\right)\right) \\ &= \Phi\left(\frac{2}{\alpha} \sinh\left(\frac{y-\gamma}{\sigma}\right)\right). \end{aligned}$$

Assim $Y \sim SHN(\alpha, \gamma, \sigma)$.

□

Algumas propriedades da distribuição seno hiperbólico são enumeradas na Proposição 1.3.

Proposição 1.3. *Se $Y \sim SHN(\alpha, \gamma, \sigma)$, então*

- $E(Y) = \gamma$ e $Var(Y) = \sigma^2 w(\alpha)$, onde $w(\alpha)$ é definido com

$$w(\alpha) = E(S^2) = \int_{\mathfrak{R}} s^2 \frac{2 \cosh(s)}{\sqrt{2\pi\alpha^2}} \exp\left\{-\frac{2}{\alpha^2} \sinh^2(s)\right\} ds,$$

sendo $S \sim SHN(\alpha, 0, 1)$;

- simétrico em torno de γ ;

- *unimodal para $\alpha < 2$ e bimodal para $\alpha > 2$.*

Para outras propriedades da distribuição seno hiperbólico normal se pode encontrar em Rieck & Nedelman (1991). Algumas formas da distribuição seno hiperbólico normal são apresentadas na Figura 1.2. Observe que o parâmetro α modifica a forma da distribuição e quando ela aumenta a curtose também aumenta.

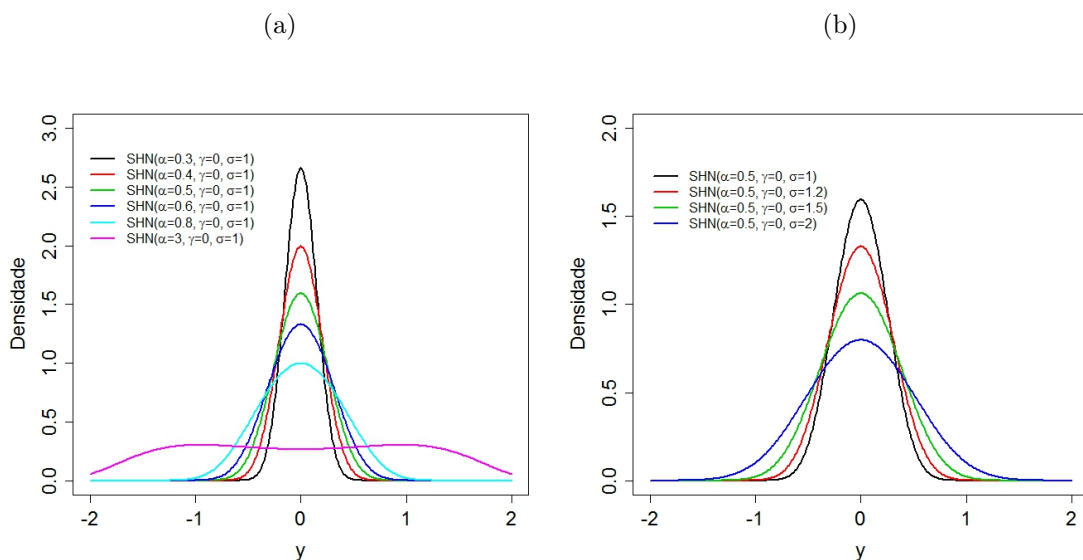


Figura 1.2: Densidade seno hiperbólico normal (SHN)

1.2 A distribuição log-Birnbaum-Saunders

Rieck & Nedelman (1991) provaram que se $T \sim BS(\alpha, \beta)$, então $Y = \log T \sim SHN(\alpha, \gamma, 2)$, para $\gamma = \log \beta$. Por esta razão, a distribuição seno hiperbólico normal é também denominada de distribuição log Birnbaum-Saunders (log BS) com parâmetros α e γ . A f.d.p. de $Y \sim \text{log-BS}(\alpha, \gamma)$ é dada por

$$f_Y(y) = \phi\left(\frac{2}{\alpha} \sinh\left(\frac{y-\gamma}{2}\right)\right) \frac{1}{\alpha} \cosh\left(\frac{y-\gamma}{2}\right). \quad (1.7)$$

1.3 Critério de seleção de modelos

Uma questão importante consiste na avaliação e escolha do modelo que melhor represente a situação em estudo. Neste trabalho utilizamos dois crité-

rios de seleção de modelos, os quais especificamente são usados na metodologia Bayesiana nos quais amostras das distribuições *a posteriori* para os parâmetros do modelo são obtidas usando métodos MCMC. A seguir descrevemos cada um deles.

1.3.1 Critério da Densidade Preditiva Ordenada (CPO)

O critério da Ordenada Preditiva Condicional (CPO) (Gelfand & Dey (1994)) é uma ferramenta de avaliação do modelo muito útil e extensamente usada na literatura estatística sob vários contextos. Seja \mathbf{D} os dados completos e \mathbf{D}_{-i} os dados com a i -ésima observação excluída. Denotamos a densidade *a posteriori* de γ dado \mathbf{D}_{-i} por $\pi(\gamma|\mathbf{D}_{-i})$, $i = 1, 2, \dots, n$ e γ é um vetor de parâmetros. Assim podemos escrever a CPO_i para a i -ésima observação

$$CPO_i = \int_{\theta} f_Y(y_i|\gamma)\pi(\gamma|\mathbf{D}_{-i})d\gamma = \left\{ \int_{\theta} \frac{\pi(\gamma|\mathbf{D})}{f_Y(y_i|\gamma)} d\gamma \right\}^{-1},$$

em que $f(y_i|\gamma)$ é a função densidade de probabilidade. Um resultado que se utiliza neste método é que para valores altos de CPO_i temos um melhor ajuste do modelo. Uma estimativa de Monte Carlo para CPO_i , considerando uma amostra de tamanho S da distribuição *a posteriori* $\pi(\gamma|\mathbf{D})$ é dada por,

$$\widehat{CPO}_i = \left\{ \frac{1}{S} \sum_{s=1}^S \frac{1}{f_Y(y_i|\gamma_s)} \right\}^{-1}. \quad (1.8)$$

Como em Cancho *et al.* (2010) utilizamos a estatística $B = \sum_{i=1}^S \log(\widehat{CPO}_i)$ na seleção dos modelos sendo que valores grandes de B indicam o melhor modelo.

1.3.2 Critério de Informação de Desvio (DIC)

Entre os métodos de seleção de modelos, existem alguns problemas em especificar o número de parâmetros. Por exemplo considere o seguinte modelo

hierárquico,

$$t_{ij}|\theta_i, \sigma^2 \sim \text{log-normal}(\mu + \theta_i, \sigma^2), i = 1, \dots, n, j = 1, \dots, m,$$

$$\theta_i|\tau^2 \sim N(0, \tau^2), i = 1, \dots, n,$$

$$\mu \sim N(0, 10^6),$$

$$\tau^2 \sim \text{Gama Inversa}(0, 001, 0, 001),$$

$$\sigma^2 \sim \text{Gama Inversa}(0, 001, 0, 001).$$

Uma visão muito rigorosa sugere que a dimensão do espaço parâmetro é $n + 3$ (uma dimensão para $\theta_1, \dots, \theta_n, \mu, \tau^2$ e σ^2), outra visão sugere que os parâmetros atuais sejam $(\mu, \tau^2$ e $\sigma^2)$, e assim sua dimensão seria 3. Uma mais realista sugere que a dimensão está entre 3 e $n+3$.

O Critério de Informação de Desvio (DIC) foi introduzido por Spiegelhalter *et al.* (2002) como uma medida de comparação e adequabilidade, resolve o problema da pobre definição do espaço parâmetro. Este critério é definido por.

$$\begin{aligned} DIC(m) &= 2\overline{D(\theta_m, m)} - D(\bar{\theta}_m, m) \\ &= D(\bar{\theta}_m, m) + 2\rho_m, \end{aligned}$$

em que $D(\theta_m, m)$ é a medida usual da *deviance* o qual é dado por,

$$D(\theta_m, m) = -2\log f(t|\theta_m, m),$$

e $\overline{D(\theta_m, m)}$ é a média *a posteriori*, ρ_m que pode ser interpretado como o número de parâmetros apropriados para o modelo m e é dado por,

$$\rho_m = \overline{D(\theta_m, m)} - D(\bar{\theta}_m, m),$$

$\bar{\theta}_m$ sendo a média *a posteriori* dos parâmetros envolvidos no modelo m ; pequenos valores do DIC indicam um melhor ajuste do modelo.

A *deviance* média *a posteriori* pode ser aproximada por

$$\overline{D} = \frac{1}{S} \sum_{s=1}^S D(\theta_s) \quad \text{com} \quad D(\theta) = -2 \sum_{i=1}^n \log [f(t_i|\theta_s)].$$

O critério DIC também pode ser aproximado considerando amostras MCMC, por $\widehat{DIC} = 2\bar{D} - \widehat{D}$, no qual

$$\widehat{D} = D \left(\frac{1}{S} \sum_{s=1}^S \theta_s \right).$$

1.4 Diagnóstico

Uma maneira de avaliar as suposições feitas sobre o modelo e de detectar possíveis pontos influentes, pode ser efetuada pela análise de diagnóstico. Esta análise teve início com a análise de resíduos para detectar a presença de pontos extremos e avaliar a adequação da distribuição proposta para a variável resposta. Uma referência importante neste assunto é o artigo de Cox & Snell (1968), que apresenta uma forma bastante geral de definir resíduos usada até os dias atuais. Uma das propostas mais inovadoras nesta área foi apresentada por Cook (1986), que propôs avaliar a influência conjunta das observações sob pequenas perturbações no modelo, ao invés da avaliação pela retirada individual ou conjunta de pontos. Se essas perturbações causam efeitos desproporcionais, pode ser indício de que o modelo está mal ajustado ou que podem existir afastamentos sérios das suposições feitas para o mesmo (Farfán Carrasco, 2007). Embora a metodologia proposta por Cook (1986) venha sendo aplicada com sucesso em diferentes áreas da estatística, observamos que dependendo da complexidade do modelo a aplicação da metodologia envolve extensas manipulações algébricas e em alguns casos um duro trabalho computacional. Vários autores têm aplicado técnicas de influência local. Peng & Dey (1995) apresentam duas distintas abordagens Bayesianas para detectar observações influentes no ajuste de modelos de regressão, uma é baseada na distribuição *a posteriori* e a outra baseada na distribuição preditiva. Uma medida que podemos ressaltar foi apresentada por Cho *et al.* (2009), no qual propuseram um método Bayesiano de análise de influência caso a caso para dados de sobrevivência, baseado na divergência de Kullback-Leibler, no qual desenvolveram medidas de diagnóstico para avaliar a influência de um caso nas distribuições *a posteriori* conjuntas e marginais fundamentadas na divergência Kullback-Leibler (K-L).

1.4.1 Análise de influência

Uma maneira de avaliar a influência de uma observação no ajuste de um modelo é por meio da deleção de causas (Cook & Weisberg (1982)).

Suponha que $K(P, P_{(-i)})$ denota a divergência K-L entre P e $P_{(-i)}$, em que P denota a distribuição a *posteriori* de θ para os dados completos e $P_{(-i)}$ é a distribuição a *posteriori* de θ sem o i -ésimo caso, assim

$$K(P, P_{(-i)}) = \int \pi(\theta|\mathbf{D}) \log \left\{ \frac{\pi(\theta|\mathbf{D})}{\pi(\theta|\mathbf{D}_{(-i)})} \right\} d\theta, \quad (1.9)$$

usaremos uma expressão equivalente para a equação (1.9), consideremos a seguinte relação,

$$\begin{aligned} \frac{\pi(\theta|\mathbf{D})}{\pi(\theta|\mathbf{D}_{(-i)})} &= \frac{\prod_{j \in D} f(y_j|\theta) \pi(\theta)}{\int_{\theta \in \Theta} \prod_{j \in D} f(y_j|\theta) \pi(\theta) d\theta} \times \frac{\int_{\theta \in \Theta} \prod_{j \in D_{(-i)}} f(y_j|\theta) \pi(\theta) d\theta}{\prod_{j \in D_{(-i)}} f(y_j|\theta) \pi(\theta)} \\ &= \frac{f(y_i|\theta)}{\int_{\theta \in \Theta} \prod_{j \in D} f(y_j|\theta) \pi(\theta) d\theta} \times \int_{\theta \in \Theta} \prod_{j \in D_{(-i)}} f(y_j|\theta) \pi(\theta) d\theta \\ &= \frac{f(y_i|\theta)}{\int_{\theta \in \Theta} f(y_i|\theta) \prod_{j \in D_{(-i)}} f(y_j|\theta) \pi(\theta) d\theta} \times \int_{\theta \in \Theta} \prod_{j \in D_{(-i)}} f(y_j|\theta) \pi(\theta) d\theta \\ &= \frac{f(y_i|\theta)}{\int_{\theta \in \Theta} f(y_i|\theta) \pi(\theta|\mathbf{D}_{(-i)})} = \frac{f(y_i|\theta)}{CPO_i}, \end{aligned}$$

assim usando esta última relação temos,

$$K(P, P_{(-i)}) = -\log(CPO_i) + E_{\theta|\mathbf{D}} \{ \log [f(y_i|\theta)] \}. \quad (1.10)$$

Uma aproximação de Monte Carlo para a equação (1.10) é expressada por,

$$K(\widehat{P}, \widehat{P}_{(-i)}) = -\log(\widehat{CPO}_i) + \frac{1}{S} \sum_{s=1}^S \log [f(y_i|\theta)], \quad (1.11)$$

sendo $K(P, P_{(-i)})$ mede o efeito de omitir o i -ésimo caso dos dados completos na distribuição a *posteriori* de θ , como é indicado por McCulloch (1989), Peng & Dey (1995) e Cho *et al.* (2009). A calibração de $K(P, P_{(-i)})$ pode ser efetuada resolvendo em p_i a equação (1.12) dada por

$$K(P, P_{(-i)}) = K(B(0.5), B(p_i)) = -\frac{\log \{4p_i(1-p_i)\}}{2}, \quad (1.12)$$

em que $B(p)$ denota uma distribuição de Bernoulli com probabilidade de sucesso p . Isto implica que descrever resultados usando $\pi(\theta|D_{(-i)})$ ao invés de $p(\theta|D)$ é equivalente com a descrição de um evento não observado com probabilidade p_i quando a probabilidade correta é 0,5. Após o cálculo de $K(P, P_{(-i)})$ em (1.10) a solução do cálculo (1.12) para p_i é dada por

$$p_i = 0.5 \left[1 + \sqrt{1 - \exp\{-2K(P, P_{(-i)})\}} \right].$$

Isto implica que $0,5 \leq p_i \leq 1$, além disso, se $p \gg 0,5$ implica que o i -ésimo caso é influente.

1.5 Objetivos do trabalho

O objetivo deste trabalho é o desenvolvimento de procedimentos inferências e de diagnóstico na classe de modelos lineares e não lineares, com erros independentes com distribuição log BS com mistura de escala normal (Balakrishnan *et al.*, 2009). O procedimento inferencial é realizado em uma perspectiva Bayesiana e é baseado em métodos de Monte Carlo via Cadeia de Markov (MCMC). O estudo de diagnóstico para os modelos considerados é baseado na medida de divergência de K-L (Cancho *et al.*, 2010). Outro objetivo deste trabalho é o desenvolvimento de procedimentos inferenciais e de diagnóstico para uma nova proposta de família de distribuição baseada na distribuição Birnbaum-Saunders, para dados de tempos de vida. Essa nova família surge em base a problemas de riscos latentes (Cancho *et al.*, 2011), no sentido de que não há informação sobre qual fator foi o responsável pela falha do componente e somente o tempo de vida observado quando uns dos fatores de risco é ativado.

1.6 Apresentação dos capítulos

A presente dissertação encontra-se dividida em cinco capítulos, no Capítulo 1 é apresentada a distribuição Birnbaum-Saunders e algumas de suas propriedades assim como sua relação com a distribuição seno hiperbólico normal. No Capítulo 2 são abordados os modelos: mistura de escala normal

(SMN), Birnbaum-Saunders com mistura de escala normal(SMNBS) e log-Birnbaum-Saunders com mistura de escala normal (log SMNBS). Na Capítulo 3 é apresentado o modelo linear e não linear com erros independentes com distribuição log SMNBS e para os modelos considerados desenvolvemos procedimentos inferenciais e de diagnósticos. No Capítulo 4 é apresentado o modelo geométrico Birnbaum-Saunders(GBS), e o modelo log geométrico Birnbaum-Saunders (LGBS). Finalmente, no Capítulo 5, algumas considerações finais e propostas de continuidade do trabalho.

Capítulo 2

Modelo Birnbaum-Saunders com mistura de escala normal

A família de distribuições mistura de escala normal (Kelker, 1970) fornece um grupo de distribuições com caudas mais pesadas do que a distribuição normal. Essas famílias de distribuições são frequentemente usados para estimações robustas de parâmetros nos quais as distribuições associadas são simétricas (Lange *et al.*, 1989, Lucas, 1997). A classe de distribuição de escala normal (SMN) inclui como casos particulares a distribuição normal, t-Student, slash, normal contaminada, entre outras.

Dizemos que uma variável aleatória Y tem distribuição SMN com parâmetro de locação μ e parâmetro de escala $\sigma^2 > 0$ e sua f.d.p. é dado por,

$$f_Y(y) = \int_0^\infty \phi(y|\mu, \kappa(u)\sigma^2) dH(u, \boldsymbol{\nu}), \quad (2.1)$$

sendo $H(\cdot, \boldsymbol{\nu})$ a função de distribuição acumulada (f.d.a.) de uma variável aleatória positiva U indexada pelo vetor de parâmetros $\boldsymbol{\nu}$ (suposto conhecido), $\kappa(\cdot)$ é uma função estritamente positiva (neste trabalho consideraremos $k(u) = \frac{1}{u}$) e $\phi(\cdot|\mu, \kappa(u)\sigma^2)$ é a função de densidade normal com média μ e variância $\kappa(u)\sigma^2$. A variável aleatória Y com f.d.p. dada por (2.1) é denotada por $Y \sim SMN(\mu, \sigma^2, H)$. Para $\mu = 0$ e $\sigma^2 = 1$ usamos a notação $Y \sim SMN(H)$.

A representação estocástica de uma variável aleatória $Y \sim SMN(\mu, \sigma^2, H)$ é dada por,

$$Y = \mu + k^{1/2}(U) Z, \quad (2.2)$$

sendo $Z \sim N(0, \sigma^2)$ e U é uma variável aleatória independente de Z . Considere a notação $\phi_{SMN}(\cdot; H)$ e $\Phi_{SMN}(\cdot; H)$ para indicar a f.d.p. e f.d.a. da distribuição SMN, respectivamente.

A Proposição 2.1 mostra a versão multivariada da distribuição SMN (Lange & Sinsheimer, 1993).

Proposição 2.1. *Considere $\mathbf{Y} = \boldsymbol{\mu} + U^{-1/2}\mathbf{Z}$, em que $\boldsymbol{\mu}$ é um vetor constante $k \times 1$, U uma variável aleatória positiva com função de distribuição $H(u)$, \mathbf{Z} um vetor aleatório $k \times 1$ independente de U com distribuição normal com média 0 e matriz de covariância $\boldsymbol{\Omega}$*

- A densidade de \mathbf{Y} é dada por

$$\int_0^\infty \frac{u^{k/2}}{(2\pi)^{k/2} |\boldsymbol{\Omega}|^{1/2}} e^{-u\delta^2/2} dH(u, \boldsymbol{\nu}),$$

em que $\delta^2 = (x - \boldsymbol{\mu})^T \boldsymbol{\Omega}^{-1} (x - \boldsymbol{\mu})$.

- A densidade de \mathbf{Y} é infinitamente diferenciável em y , $\boldsymbol{\mu}$, $\boldsymbol{\Omega}$ exceto onde $y = \boldsymbol{\mu}$, considerando $E(U^{(k/2)+j}) < \infty$.
- Se o vetor de médias $E(\mathbf{Y})$ existe, então $E(\mathbf{Y}) = \boldsymbol{\mu}$
- Se a covariância $cov(\mathbf{Y})$ existe, então $cov(\mathbf{Y}) = E(U^{-1})\boldsymbol{\Omega}$
- A função característica de \mathbf{Y} é dada por,

$$E\left(e^{i\theta^t Y}\right) = e^{i\theta^t \boldsymbol{\mu}} \int_0^\infty e^{-1/(2u)\theta^t \boldsymbol{\Omega} \theta} dH(u, \boldsymbol{\nu}).$$

A prova desses resultados e outras propriedades da versão multivariada da distribuição de mistura escala normal pode ser consultada em Lange & Sinsheimer (1993).

Para aplicação desta família em modelos lineares mistos da SMN, veja por exemplo, Rosa *et al.* (2003).

Na seguinte Seção é apresentado uma extensão da distribuição Birnbaum-Saunders baseada na versão univariada da distribuição SMN.

2.1 Modelo Birnbaum-Saunders com mistura de escala normal

A generalização da distribuição BS baseada na distribuição SMN apresentada por Balakrishnan *et al.* (2009) é mostrada na Proposição 2.2.

Proposição 2.2. *Se $Y \sim SMN(H)$, então a variável aleatória*

$$T = \frac{\beta}{4} \left(\alpha Y + \sqrt{\alpha^2 Y^2 + 4} \right)^2, \quad (2.3)$$

tem distribuição Birnbaum-Saunders com mistura de escala normal (BSSMN), com f.d.a. dada por

$$F_T(t) = \Phi_{SMN}(a(t); H), \quad t > 0, \quad (2.4)$$

com $a(t) = \frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right]$, $\alpha > 0, \beta > 0$, e $\Phi_{SMN}(\cdot; H)$ é a f.d.a. de uma variável aleatória SMN com parâmetros $\mu = 0$ e $\sigma^2 = 1$.

Prova: Seja T a variável aleatória dada em (2.3). A f.d.a. de T é

$$F_T(t) = P(T \leq t) = P\left(\frac{\beta}{4} \left(\alpha Y + \sqrt{\alpha^2 Y^2 + 4} \right)^2 \leq t\right).$$

De (2.3) temos $\sqrt{\frac{T}{\beta}} = \frac{1}{2} (\alpha Y + \sqrt{\alpha^2 Y^2 + 4})$, e $\sqrt{\frac{\beta}{T}} = -\left(\frac{\alpha Y - \sqrt{\alpha^2 Y^2 + 4}}{2}\right)$; assim Y pode ser expressado por,

$$Y = \frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right].$$

Portanto,

$$\begin{aligned} P(T \leq t) &= P\left(\frac{\beta}{4} \left(\alpha Y + \sqrt{\alpha^2 Y^2 + 4} \right)^2 \leq t\right) \\ &= P\left(Y \leq \frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right]\right) \\ &= \Phi_{SMN}\left(\frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right]\right). \end{aligned}$$

□

A notação $T \sim BSSMN(\alpha, \beta; H)$ indica que a variável aleatória T tem distribuição BS com mistura de escala normal.

Da Proposição 2.2, tem-se que a f.d.p. da variável aleatória T , pode ser obtida, derivando a equação (2.4) com respeito a t , e é dada por

$$f_T(t) = \phi_{SMN}(a(t); H)A(t), \quad t > 0, \quad (2.5)$$

em que $A(t) = t^{-3/2} [t + \beta] / [2\alpha\beta^{1/2}]$ é a primeira derivada de $a(t)$ com respeito a t .

Os resultados apresentados nas Proposições 2.3 e 2.4 obtidos por Balakrishnan *et al.* (2009), estão relacionados com a função de risco e os momentos da distribuição BSSMN, que são úteis para a caracterização desta distribuição.

Proposição 2.3. *Se $T \sim BSSMN(\alpha, \beta; H)$ então a função de risco de T é*

$$r_T(t) = \frac{\phi_{SMN}(a(t); H)A(t)}{\Phi_{SMN}(-a(t); H)}, \quad t > 0,$$

em que $\phi_{SMN}(\cdot; H)$, $\Phi_{SMN}(\cdot; H)$ são f.d.p. e f.d.a. de SMN com $\mu = 0$ e $\sigma = 1$. $a(t) = \frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right]$ e $A(t) = t^{-3/2} [t + \beta] / [2\alpha\beta^{1/2}]$.

A prova é direta usando a definição de função de risco.

Proposição 2.4. *Se $T \sim BSSMN(\alpha, \beta; H)$, então o k -ésimo momento de T , é dado por*

$$E [T^k] = \beta^k \sum_{i=0}^k \binom{2k}{2i} \sum_{j=0}^i \binom{i}{j} w_{k+j-i} \left[\frac{\alpha}{2} \right]^{2(k+j-i)}, \quad k = 1, 2, \dots, \quad (2.6)$$

para $w_r = E [U^{-r}]$.

A prova da Proposição 2.4 pode ser encontrada em Balakrishnan *et al.* (2009).

Da Proposição 2.4 temos que a média, a variância, o coeficiente de assimetria e curtoses da variável aleatória T com distribuição BSSMN($\alpha, \beta; H$) são respectivamente, dadas por

$$\begin{aligned} E[T] &= \frac{\beta}{2} [2 + \omega_1 \alpha^2], \quad Var[T] = \frac{\beta^2 \alpha^2}{4} [\omega_1 + (2\omega_2 - \omega_1^2) \alpha^2], \\ \alpha_3 &= \frac{4\alpha [(3\omega_2 - 3\omega_1^2) + 0.5\alpha^2 (2\omega_3 - 3\omega_1\omega_2 + \omega_1^3)]}{[4\omega_1 + (2\omega_2 - \omega_1^2) \alpha^2]^{1.5}}, \end{aligned}$$

e

$$\alpha_4 = \frac{16\omega_2 + \alpha^2(32\omega_3 - 48\omega_1\omega_2 + 24\omega_1^3) + \alpha^4(8\omega_4 - 16\omega_1\omega_3 + 12\omega_1^2\omega_2 - \omega_1^4)}{[4\omega_1 + (2\omega_2 - \omega_1^2)\alpha^2]^2}.$$

Note que α_3 e α_4 são controlados pelo parâmetro de forma α .

2.1.1 Exemplos da distribuição BSSMN

Nesta seção consideramos alguns casos particulares da distribuição BS-SMN, os quais são obtidos diretamente da representação estocástica dada na Proposição 2.2.

A distribuição Birnbaum-Saunders t com ν graus de liberdade

Neste caso consideramos que a variável aleatória U , em (2.1), tem distribuição gama com f.d.p. dada por

$$h(u; \nu) = \frac{[\frac{\nu}{2}]^{\frac{\nu}{2}} u^{\frac{\nu}{2}-1} \exp\left(-\frac{\nu u}{2}\right)}{\Gamma(\frac{\nu}{2})}, \quad u > 0, \quad \nu > 0.$$

Da proposição 2.2, a f.d.p. da variável aleatória T tem a seguinte forma

$$f_T(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{(\pi\nu)\Gamma(\nu/2)}} \left[1 + \frac{1}{\nu\alpha^2} \left\{\frac{t}{\beta} + \frac{\beta}{t} - 2\right\}\right]^{-\frac{(\nu+1)}{2}} \frac{t^{-3/2}[t + \beta]}{2\alpha\sqrt{\beta}}, \quad t > 0, \quad (2.7)$$

com $\alpha > 0$, $\beta > 0$ e $\nu > 0$. Consideramos a notação BS-t(α, β, ν) para indicar a distribuição Birnbaum-Saunders t-Student com parâmetros $\alpha > 0$, $\beta > 0$ e $\nu > 0$. Na Figura 2.1 apresentamos algumas formas da f.d.p. desta distribuição. Um caso particular da distribuição BS-t é a distribuição BS Cauchy, obtida para $\nu = 1$. Quando $\nu \rightarrow \infty$, obtemos a distribuição BS como uma distribuição limite BS-t.

Na Figura 2.1 apresentamos o gráfico da f.d.p. BS-t para diferentes valores de α , a medida que o valor α aumenta a assimetria é mais notável, além disso quando o grau de liberdade aumenta a distribuição BS-t converge para a distribuição BS.

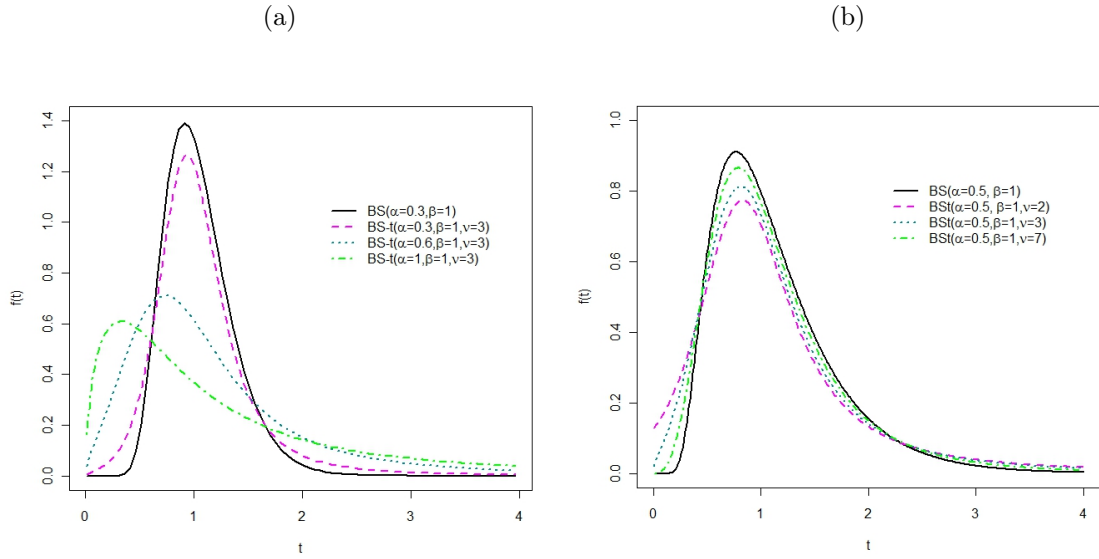


Figura 2.1: Densidade da distribuição Birnbaum-Saunders-t com (a) $\alpha=0,1,3$; β, ν fixos; (b) α e β fixos, $\nu=2; 3; 6$.

Distribuição Birnbaum-Saunders normal contaminada

Neste caso a variável aleatória U é discreta e assume dois valores. A função de probabilidade de U é dada por

$$h(u; \boldsymbol{\nu}) = \nu \mathbb{I}_{(u=v)} + (1 - \nu) \mathbb{I}_{(u=1)}, \quad 0 < \nu < 1, \quad 0 < v \leq 1,$$

para $\boldsymbol{\nu} = (\nu, v)^\top$ e $\mathbb{I}_A(\cdot)$ denota a função indicadora do conjunto A . Da Proposição 2.2, a f.d.p. da variável aleatória T tem a seguinte forma

$$f_T(t) = \left[\nu v^{1/2} \phi(v^{1/2} a(t)) + (1 - \nu) \phi(a(t)) \right] \frac{t^{-3/2} [t + \beta]}{2\alpha\sqrt{\beta}}, \quad t > 0, \quad (2.8)$$

onde $\alpha > 0$, $\beta > 0$, e $0 < \nu < 1$, $0 < v < 1$ sendo $\phi(\cdot)$ a f.d.p. da distribuição normal padrão. Consideremos a notação $\text{BSNC}(\alpha, \beta, \boldsymbol{\nu})$ para indicar a distribuição birnbaum Saunder Normal Contaminada com parâmetros $\alpha, \beta > 0$ e $\boldsymbol{\nu}$. Essa distribuição inclui o caso distribuição BS, $v = 1$ e $\nu = 1$ ou $v = 1/2$. Algumas formas dessa distribuição são apresentadas na Figura 2.2.

Na Figura 2.2 (a) pode-se observar o gráfico da f.d.p. da distribuição BSNC para diferentes valores de v , a medida que aumenta o parâmetro v a curtose aumenta. Na Figura 2.2 (b) se observa a f.d.p. da distribuição BSNC para diferentes valores de ν , note que quando o parâmetro ν aumenta a curtose diminui.

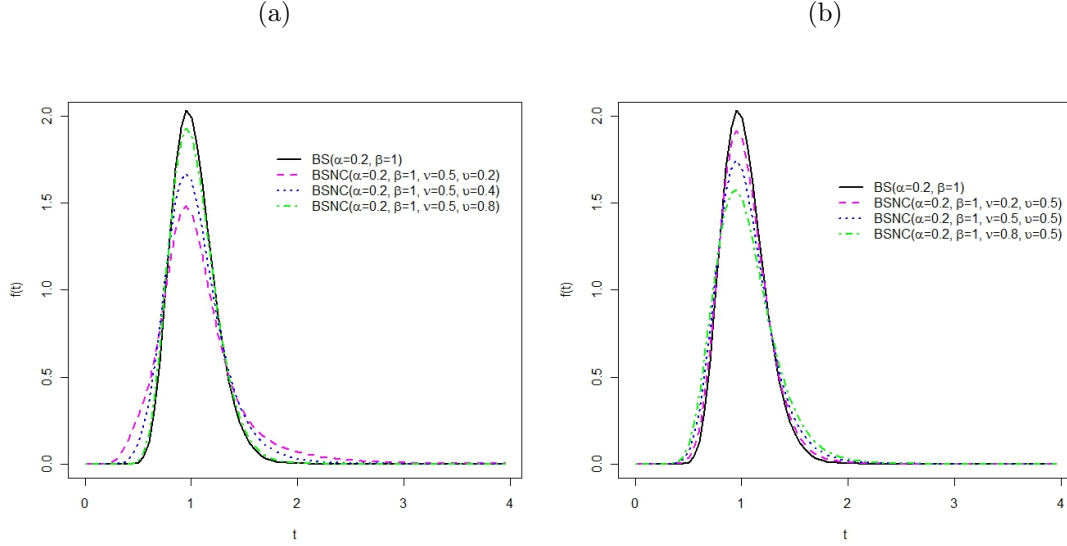


Figura 2.2: Densidade da distribuição Birnbaum-Saunders-Normal Contaminada com (a) α, β, ν fixos e $\nu=0,2; 0,4; 0,8$; (b) α, β, ν , fixos $\nu=0,2; 0,5; 0,8$.

Distribuição Birnbaum-Saunders slash

Neste caso temos que $U \sim \text{beta}(\nu, 1)$ com parâmetro de forma $\nu > 0$ e usamos a notação $T \sim \text{BSSL}(\mu, \sigma^2, \nu)$. A f.d.p. da variável T é dada por

$$f_T(t) = \left[\nu \int_0^1 u^{\nu-1/2} \phi(a(t)u^{1/2}) du \right] \frac{t^{-3/2}[t + \beta]}{2\alpha\sqrt{\beta}}, \quad t > 0, \alpha > 0, \beta > 0, \nu > 0, \quad (2.9)$$

sendo que $a(t)$ é dada em (2.4). A distribuição BSSL apresenta caudas mais pesadas que a distribuição BS (Gómez *et al.*, 2009). Além disso, quando $\nu \rightarrow \infty$ a distribuição BSSL se reduz a distribuição BS. Na Figura 2.3 (a) é apresentada o gráfico da f.d.p. BSSL, em que a medida que o parâmetro α aumenta a curtose diminui. Na Figura 2.3 (b) apresentamos o gráfico da f.d.p. da distribuição BSSL; a medida que ν aumenta a f.d.p. BSSL se aproxima para a f.d.p. BS.

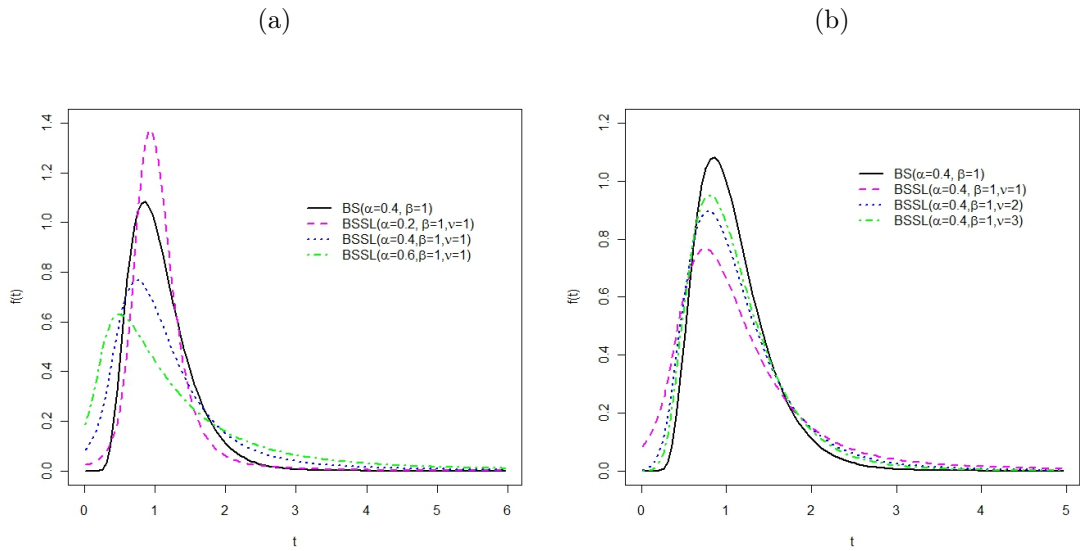


Figura 2.3: Densidade da distribuição Birnbaum-Saunders slash com (a) $\alpha = 0, 2; 0, 4; 0, 6$ e β, ν fixos; (b) α e β fixos e $\nu = 1; 2; 3$

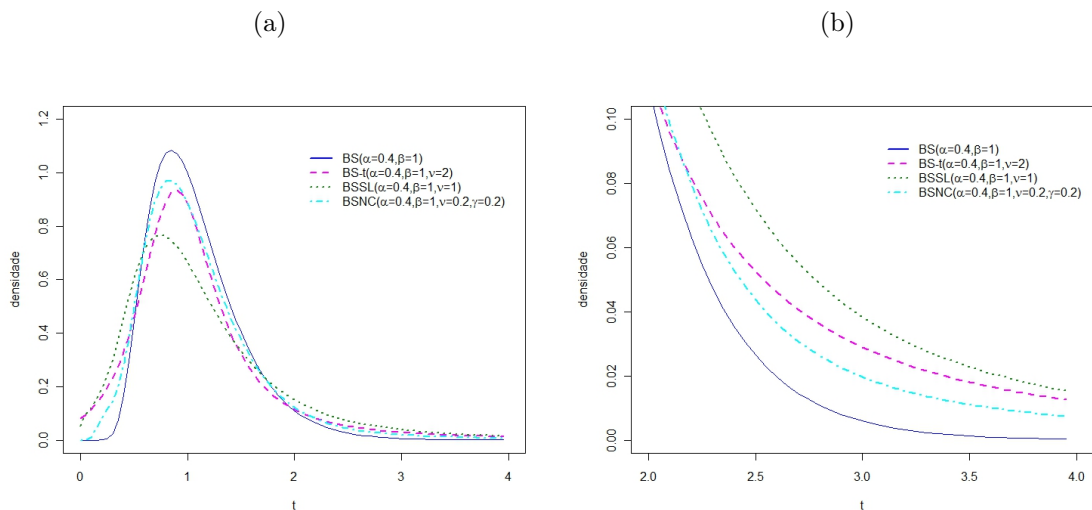


Figura 2.4: (a) Densidades das distribuições BSSMN; (b) BSSMN com α, β fixos e com diferentes valores para ν

2.2 Distribuição log BSSMN

Nesta seção apresentamos a distribuição log BSSMS como um caso particular da distribuição seno hiperbólico com mistura de escala normal (SHSMN). A família de distribuição SHSMN é obtido diretamente de Diaz-Garcia & Dominguez-Molina (2006), a qual é apresentada na seguinte proposição.

Proposição 2.5. *Se $Z \sim \Phi_{SMN}(\cdot; H)$ (2.2), então a variável aleatória*

$$Y = \gamma + \sigma \operatorname{arcsinh} \left(\frac{\alpha Z}{2} \right) \quad (2.10)$$

tem função de densidade dada por

$$f_Y(y) = \frac{2}{\alpha\sigma} \cosh \left(\frac{y - \gamma}{\sigma} \right) \phi_{SMN} \left(\frac{2}{\alpha} \sinh \left(\frac{y - \gamma}{\sigma} \right); H \right), \quad y \in \mathfrak{R}, \alpha > 0, \quad (2.11)$$

para $\alpha > 0$, $\gamma \in \mathfrak{R}$.

Prova: Considere a f.d.a. da variável aleatória definida em (2.10)

$$\begin{aligned} F_Y(y) &= P(Y < y) = P \left(\gamma + \sigma \operatorname{arcsinh} \left(\frac{\alpha Z}{2} \right) < y \right) \\ &= P \left(Z < \frac{2}{\alpha} \sinh \left(\frac{y - \gamma}{\sigma} \right) \right) \\ &= \Phi_{SMN} \left(\frac{2}{\alpha} \sinh \left(\frac{y - \gamma}{\sigma} \right); H \right), \quad y \in \mathfrak{R}, \end{aligned}$$

derivando a $F_Y(y)$ com respeito a y , temos

$$f_Y(y) = \frac{2}{\alpha\sigma} \cosh \left(\frac{y - \gamma}{\sigma} \right) \phi_{SMN} \left(\frac{2}{\alpha} \sinh \left(\frac{y - \gamma}{\sigma} \right); H \right).$$

□

A notação $Y \sim SHSMN(\alpha, \gamma, \sigma^2, H)$ indica que a variável aleatória Y tem distribuição seno hiperbólico com mistura de escala normal com parâmetros α , γ , σ^2 e distribuição de mistura H . Como no caso da distribuição seno hiperbólico normal, a f.d.p. de Y dada em (2.11) é simétrica ao redor de γ , o parâmetro σ^2 modifica a escala, enquanto os parâmetros ν e α controlam as caudas e a curtose da distribuição respectivamente. Uma relação importante entre a família log BSSMN e SHSMN é baseada na Proposição 2.6.

Proposição 2.6. *Uma variável aleatória $T \sim BSSMN(\alpha, \beta; H)$ se e somente se $Y = \log T \sim SHSMN(\alpha, \gamma, \sigma^2; H)$, onde $\gamma = \log \beta$ e $\sigma = 2$ o qual será denotado por $Y \sim \log BSSMN(\alpha, \gamma; H)$.*

A prova da Proposição 2.6 pode ser encontrado em Balakrishnan *et al.* (2009).

Proposição 2.7. *Uma variável aleatória $T \sim BSSMN(\alpha, \beta; H)$, então a variável aleatória $Y = \log(T)$ pode ser expressado por*

$$Y = \gamma + 2 \log \left(\frac{\alpha Z}{2} + \sqrt{\frac{\alpha^2 Z^2}{4} + 1} \right), \quad (2.12)$$

para $Z \sim SMN(H)$ e $\gamma = \log(\beta)$.

A prova da Proposição 2.7 é imediato usando a relação (2.3). A equação (2.12) pode ser usada para gerar variáveis aleatórias log BSSMN.

2.2.1 Exemplos da distribuição log BSSMN

Nesta seção consideramos alguns casos particulares da distribuição log BS-SMN.

Log Birnbaum-Saunders-t (log BS-t)

Usando as Proposições 2.5 e 2.6 e a equação (2.7), a f.d.p. para o modelo log BS-t é dada por

$$f_Y(y) = \frac{1}{\alpha} \cosh\left(\frac{y - \gamma}{2}\right) \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{\left(\frac{2}{\alpha} \sinh\left(\frac{y-\gamma}{2}\right)\right)^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}, \quad (2.13)$$

onde $\alpha > 0$ parâmetro de forma, γ parâmetro de locação, e o parâmetro ν que está relacionado com a forma da distribuição.

Algumas formas da f.d.p. do modelo log BS-t são apresentadas na Na Figura 2.5. Na Figura 2.5 (a) é apresentada a f.d.p. log BS-t com o parâmetro α assumindo diferentes valores, se observa que ela é simétrica e a medida que aumenta α a curtose diminui. Na Figura 2.5 (b) apresentamos a f.d.p. log BS-t, para α , fixado e ν assumindo diferentes valores. Podemos observar para valores grandes a distribuição log BS-t se aproxima para a distribuição log BS.

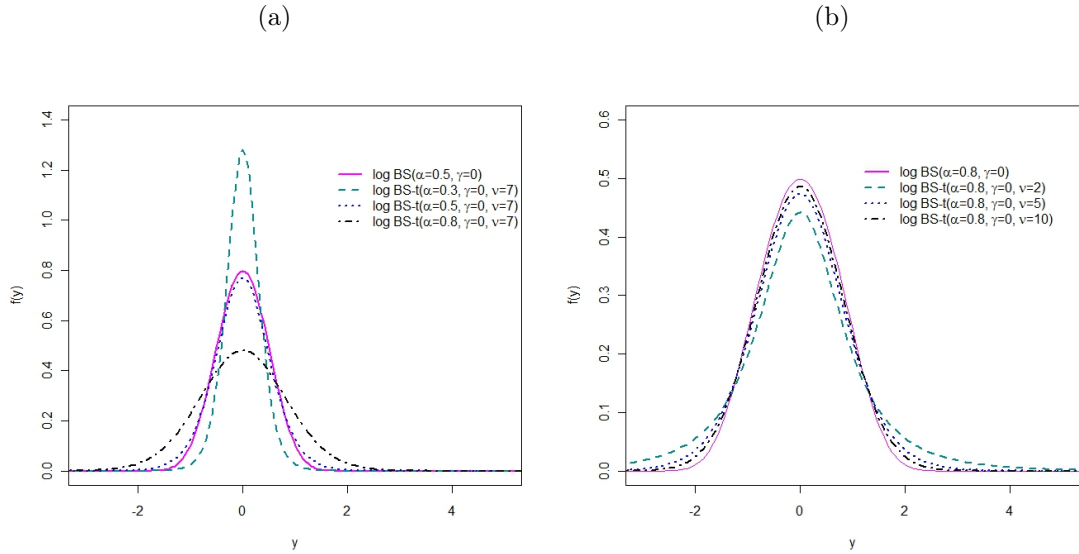


Figura 2.5: Densidade da distribuição log Birnbaum-Saunders-t com (a) $\alpha = 0,3; 0,5; 0,8$; e γ, ν fixos; (b) α, γ , fixos e $\nu = 2; 5; 10$.

Log Birnbaum-Saunders slash(log BSSL)

Usando a Proposição 2.5 e 2.6, e a equação (2.9), a f.d.p. para o modelo log BSSL é dada por

$$f_Y(y) = \frac{1}{\alpha} \cosh\left(\frac{y - \gamma}{2}\right) \int_0^\infty \nu u^{\nu-1} \phi\left(\frac{2}{\alpha} \sinh\left(\frac{y - \gamma}{2}\right) \mid 0, \frac{1}{u}\right) du, \quad (2.14)$$

sendo ϕ é a distribuição normal padrão, $\nu > 0$ e $\alpha > 0$.

Na Figura 2.6 (a) é apresentada a f.d.p. log BSSL com o parâmetro α assumindo diferentes valores. Observamos que ela é simétrica e a medida que α aumenta, a curtose diminui. Na Figura 2.6 (b) é apresentada a f.d.p. log BSSL, fixando o parâmetro α , e o parâmetro ν assumindo diferentes valores. A medida que aumenta ν se pode observar que ela converge para a distribuição BS.

Log Birnbaum-Saunders Normal Contaminada(log BSNC)

Usando as Proposições 2.5 e 2.6 e a equação (2.8), a f.d.p. para o modelo log BSNC é dada por

$$f_Y(y) = \frac{1}{\alpha} \cosh\left(\frac{y-\gamma}{2}\right) \left[\nu \phi\left(\frac{2}{\alpha} \sinh\left(\frac{y-\gamma}{2}\right) \mid 0, 1/\gamma\right) + (1-\nu) \phi\left(\frac{2}{\alpha} \sinh\left(\frac{y-\gamma}{2}\right)\right) \right], \quad (2.15)$$

em que o parâmetro α modifica a forma, o parâmetro γ é a escala e os parâmetros ν e v permitem modificar a curtose. Na Figura 2.7 (a) é apresentada a densidade da distribuição log BSNC, no qual observamos que a medida que o parâmetro ν cresce a curtose diminui, (b) a medida que o parâmetro v cresce a distribuição log BSNC aproxima-se para a distribuição log BS.

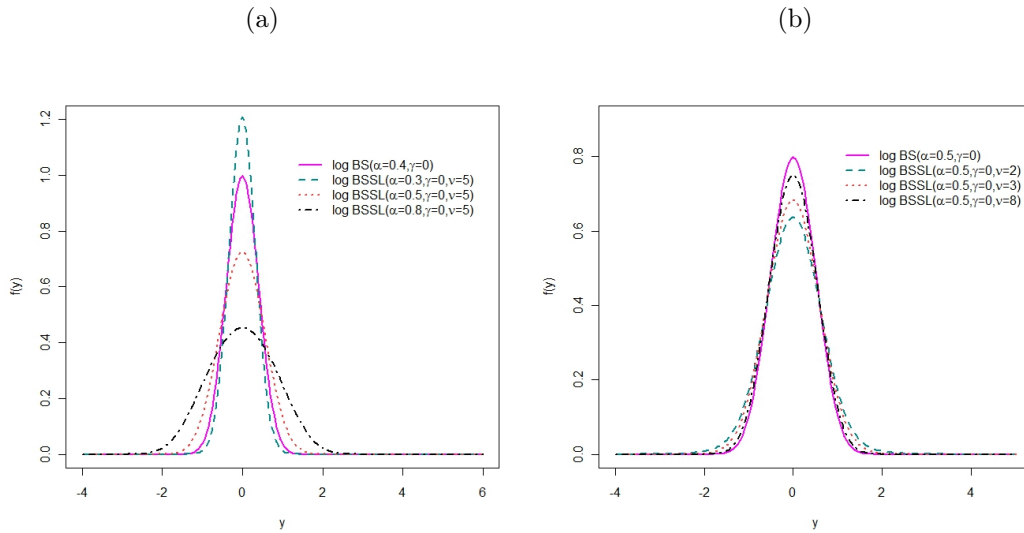


Figura 2.6: Densidade da distribuição log Birnbaum-Saunders slash com (a) $\alpha=0,3; 0,5; 0,8$ e γ, ν fixos; (b) α, γ fixos, e $\nu=2; 3; 8$.

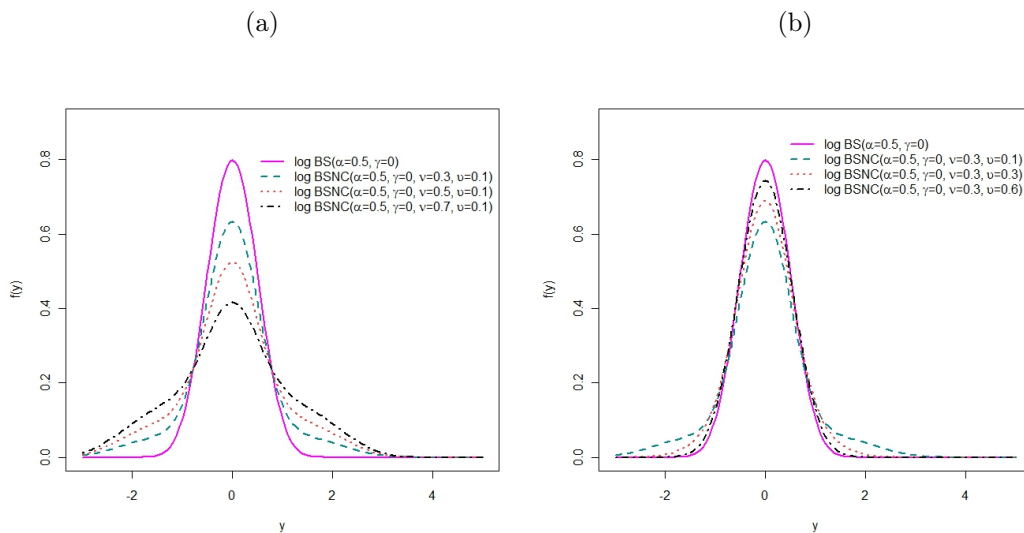


Figura 2.7: Densidade da distribuição log BSNC com (a) α, γ, u fixos e $\nu=0,3; 0,5; 0,7$; (b) α, γ, ν fixos e $u=0,1; 0,3; 0,6$.

Capítulo 3

Inferência Bayesiana para modelos de regressão log BSSMN

Na prática, existem situações em que uma ou mais covariáveis estão associadas aos tempos de vida. Por exemplo, na indústria o tempo de vida de um determinado equipamento pode ser influenciado pelo nível de voltagem a que o equipamento é submetido; na área médica o tempo de vida de um paciente pode estar relacionado com o tipo e tamanho do tumor, quantidade de hemoglobina no sangue, raça e idade do paciente, entre outros factores.

Consideramos agora o modelo de regressão baseado na distribuição log BSSMN dada pela Proposição 2.6, assumindo uma relação linear entre log do tempo de fadiga de um material, denotado por Y e o vetor de covariáveis $\mathbf{x}=(x_1, \dots, x_p)^\top$. Temos que Y dado \mathbf{x} tem distribuição log BSSMN e a relação entre \mathbf{x} e y pode ser representada por

$$y = \mathbf{x}^\top \boldsymbol{\beta} + \epsilon, \quad (3.1)$$

onde $\boldsymbol{\beta}=(\beta_1, \dots, \beta_p)^\top$, é um vetor de parâmetros desconhecido para ser estimados e ϵ tem distribuição log BSSMN, pode ser demonstrado que a f.d.p. de Y dado \mathbf{x} é dada por

$$f_Y(y) = \frac{1}{\alpha} \cosh\left(\frac{y - \mathbf{x}^\top \boldsymbol{\beta}}{2}\right) \phi_{SMN}\left(\frac{2}{\alpha} \sinh\left(\frac{y - \mathbf{x}^\top \boldsymbol{\beta}}{2}\right); H\right), \quad y \in \Re, \alpha > 0.$$

Como caso particular do modelo dado em (3.1), temos os modelos de regressão log BS, log BS-t, log BSSL e log BSNC.

Considere um conjunto de n observações $(x_1, y_1), \dots, (x_n, y_n)$ sendo que $x_i = (x_{i,1}, \dots, x_{i,p})^\top$, $\epsilon_i \sim \log BSSMN(\alpha, 0; H)$, e a função de verossimilhança para os parâmetros α, β, ν é dada por,

$$L(\alpha, \beta, \nu | \mathbf{x}, \mathbf{y}) \propto \prod_{i=1}^n \frac{1}{\alpha} \cosh\left(\frac{y_i - \mathbf{x}_i^T \beta}{2}\right) \phi_{SMN}\left[\frac{2}{\alpha} \sinh\left(\frac{y_i - \mathbf{x}_i^T \beta}{2}\right); H\right], \quad (3.2)$$

para $i = 1, \dots, n$. Note que não existe soluções explícitas para o problema de maximização da função de verossimilhança (3.2); no entanto, esta pode ser maximizada numericamente usando métodos numéricos, tais métodos já estão implementados como por exemplo, o "optim" do aplicativo R (R Development Core Team, 2011). Além disso, as distribuições amostrais do estimador de máxima verossimilhança (EMV) são conhecidas assintoticamente, o que pode resultar em um problema sério, uma vez que em geral, trabalha-se com amostras não muito grandes. Desta forma, inferências baseadas nessas estatísticas podem não ser confiáveis. Um procedimento alternativo que pode contornar esse problema é a utilização de métodos Bayesianos.

Neste capítulo apresentamos uma análise Bayesiana do modelo de regressão log BSSMN, usando Métodos de Monte Carlo via Cadeias de Markov (MCMC), tais como o amostrador de Gibbs (Gilks & Wild, 1992) e Metropolis-Hastings (Chib & Greenberg, 1995).

3.1 Distribuições a priori, a posteriori e conjunta

Uma característica importante destes modelos (log BSSMN) é que permitem a implementação computacional. Considerando as equações (2.2) e (2.10) temos uma representação estocástica para os erros, dada por

$$\epsilon_i = 2 \operatorname{arcsinh}\left(\frac{\alpha U^{-1/2} Z_i}{2}\right), \quad i = 1, \dots, n, \quad (3.3)$$

em que $U_i \sim H(\cdot, \nu)$, Z_i tem uma distribuição normal padrão e $\mathbf{u} = (u_1, \dots, u_n)$. Note que de (3.1) e de (3.3) temos $\epsilon | U \sim \log BS(\alpha u^{1/2}, 0)$ que é equivalente a dizer $y_i | u_i \sim \log BS(\alpha u_i^{-1/2}, x_i^T \beta)$ ou $y_i | u_i \sim SHN(\alpha u_i^{-1/2}, x_i^T \beta, 2)$, portanto considerando a variável de mistura temos que a função de verossimilhança ampliada para

os parâmetros α , $\boldsymbol{\beta}$, $\boldsymbol{\nu}$, é dada por,

$$L(\alpha, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{u} | \mathbf{y}, \mathbf{x}) \propto \alpha^{-n} \prod_{i=1}^n u_i^{\frac{1}{2}} \cosh\left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2}\right) \times \exp\left(-\frac{2}{\alpha^2} \sum_{i=1}^n u_i \sinh^2\left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2}\right)\right) \times h(u_i; \boldsymbol{\nu}), \quad (3.4)$$

em que $h(u_i; \boldsymbol{\nu})$ é a função de distribuição de probabilidade da variável de mistura U_i .

Considere a seguinte reparametrização $\lambda = \alpha^{-2}$. Esta reparametrização é útil porque ela nos permite obter a condicional *a posteriori* para λ de forma fechada; assim a função de verossimilhança pode ser escrita como,

$$L(\lambda, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{u} | \mathbf{y}, \mathbf{x}) \propto \lambda^{\frac{n}{2}} \prod_{i=1}^n u_i^{\frac{1}{2}} \cosh\left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2}\right) \times \exp\left(-2\lambda \sum_{i=1}^n u_i \sinh^2\left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2}\right)\right) \times h(u_i; \boldsymbol{\nu}). \quad (3.5)$$

Consideramos uma distribuição *a priori* conjunta própria para os parâmetros do modelo com o intuito de garantir que a distribuição *a posteriori* conjunta seja própria. Consideramos uma distribuição *a priori* para todos os parâmetros desconhecidos de acordo ao espaço parâmetro, e também consideramos os parâmetros independentes. A densidade da distribuição *a priori* dos parâmetros desconhecidos é dada por,

$$\pi(\lambda, \boldsymbol{\beta}, \boldsymbol{\nu},) \propto \pi(\lambda) \left[\prod_{j=1}^p \pi(\beta_j) \right] \pi(\boldsymbol{\nu}), \quad (3.6)$$

para $\beta_j \stackrel{ind}{\sim} N(\mu_{\beta_j}, \sigma_{\beta_j}^2)$, $j = 1, \dots, p$, $\lambda \sim G(a, b)$, sendo que G denota a distribuição gama com média a/b e variância a/b^2 . A distribuição *a priori* de $\boldsymbol{\nu}$ com densidade $\pi(\boldsymbol{\nu})$ depende de cada caso particular da distribuição SMN, portanto considerando *a priori* correspondente, junto com a equação (3.5), a distribuição *a posteriori* conjunta para λ , $\boldsymbol{\beta}$ e $\boldsymbol{\nu}$ é dada por,

$$\begin{aligned} \pi(\lambda, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{u} | \mathbf{x}, \mathbf{y}) &\propto \lambda^{\frac{n}{2}} \prod_{i=1}^n \left[u_i^{\frac{1}{2}} \cosh\left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2}\right) \right] \\ &\times \exp\left\{-2\lambda \sum_{i=1}^n u_i \sinh^2\left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2}\right)\right\} \times h(u_i; \boldsymbol{\nu}) \\ &\times \pi(\lambda, \boldsymbol{\beta}, \boldsymbol{\nu}). \end{aligned} \quad (3.7)$$

A densidade *a posteriori* conjunta dada em (3.7), é analiticamente intratável, porque a integração da densidade *a posteriori* conjunta não é fácil de ser obtida. Portanto consideramos Métodos de Monte Carlo via Cadeias de Markov(MCMC), tais como Gibbs sampling e o algoritmo Metropolis Hasting, os quais podem ser usados para obter amostras; para isso, as condicionais *a posteriori* são necessárias. Dado \mathbf{u} , as condicionais das distribuições *a posteriori* são como no modelo log BS padrão.

As distribuição condicional *a posteriori* para o parâmetro λ é dada por

$$\begin{aligned}\pi(\lambda|\boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{u}, \mathbf{y}, \mathbf{x}) &\propto \lambda^{\frac{n}{2}} \times \exp \left\{ -2\lambda \sum_{i=1}^n u_i \sinh^2 \left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2} \right) \right\} \pi(\lambda) \\ &\propto \lambda^{\frac{n}{2}+a-1} \times \exp \left\{ -\lambda \left(2 \sum_{i=1}^n u_i \sinh^2 \left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2} \right) + b \right) \right\}.\end{aligned}$$

Portanto

$$\lambda|\boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{u}, \mathbf{y}, \mathbf{x} \sim G \left(\frac{n+2a}{2}, 2 \sum_{i=1}^n u_i \sinh^2 \left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2} \right) + b \right). \quad (3.8)$$

A distribuição condicional *a posteriori* para $\boldsymbol{\beta}$ é dada por,

$$\begin{aligned}\pi(\boldsymbol{\beta}|\lambda, \boldsymbol{\nu}, \mathbf{u}, \mathbf{y}, \mathbf{x}) &\propto \exp \left\{ -2\lambda \sum_{i=1}^n u_i \sinh^2 \left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2} \right) \right\} \\ &\times \prod_{i=1}^n \cosh \left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2} \right) \pi(\boldsymbol{\beta}) \quad i = 1, \dots, n.\end{aligned} \quad (3.9)$$

Para completar as especificações para o esquema de amostragem Gibbs, precisamos da distribuição condicional *a posteriori* para \mathbf{u} e o parâmetro $\boldsymbol{\nu}$; assim para cada elemento de \mathbf{u} a distribuição condicional *a posteriori* é dada por,

$$\begin{aligned}\pi(u_i|\lambda, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{u}, \mathbf{y}, \mathbf{x}) &\propto u_i^{\frac{1}{2}} \times \exp \left\{ -2\lambda u_i \sinh^2 \left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2} \right) \right\} \times h(u_i, \boldsymbol{\nu}) \\ &\propto G \left(\frac{3}{2}, 2\lambda \sinh^2 \left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{2} \right) \right) \times h(u_i; \boldsymbol{\nu}),\end{aligned} \quad (3.10)$$

para $i = 1, \dots, n$, e a distribuição condicional *a posteriori* para $\boldsymbol{\nu}$ é,

$$\pi(\boldsymbol{\nu}|\lambda, \boldsymbol{\beta}, \mathbf{u}, \mathbf{y}, \mathbf{x}) \propto \pi(\boldsymbol{\nu}) \prod_{i=1}^n h(u_i, \boldsymbol{\nu}). \quad (3.11)$$

Podemos observar que a forma das distribuições *a posteriori* dadas em (3.10) e (3.11) depende da distribuição de u .

Vamos agora apresentar alguns exemplos.

Student-t

Como vimos na Seção 2.1.1, consideramos uma distribuição gama para a variável aleatória u . Assim, a distribuição condicional *a posteriori* para u_i é dada por,

$$\begin{aligned}\pi(u_i|\lambda, \beta, \nu, u, y, x) &\propto G\left(\frac{3}{2}, 2\lambda \sinh^2\left(\frac{y_i - x_i^T \beta}{2}\right)\right) \times h(u_i; \nu) \\ &\propto G\left(\frac{3}{2}, 2\lambda \sinh^2\left(\frac{y_i - x_i^T \beta}{2}\right)\right) \times u_i^{\frac{\nu}{2}-1} \exp\left\{-\frac{\nu}{2}u\right\},\end{aligned}$$

logo

$$u_i \sim G\left(\frac{\nu+1}{2}, -0.5\nu + 2\lambda \sinh^2\left(\frac{y_i - x_i^T \beta}{2}\right)\right), \quad i = 1, \dots, n.$$

Para o parâmetro ν consideramos uma *priori* especificada pela distribuição exponencial truncada; assim a distribuição condicional *a posteriori* é dada por,

$$\begin{aligned}\pi(\nu|\lambda, \beta, u, y, x) &\propto \pi(\nu) \prod_{i=1}^n h(u_i, \nu) \\ \pi(\nu|\lambda, \beta, u, y, x) &\propto \pi(\nu) \prod_{i=1}^n \frac{\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)} u_i^{\frac{\nu}{2}-1} \exp\left\{-\frac{\nu}{2}u_i\right\}.\end{aligned}\tag{3.12}$$

Slash

Considere uma distribuição $G(a, b)$ nos quais os valores a e b são pequenos e positivos com $a \ll b$, para adotar uma *priori* para ν , neste caso a distribuição condicional *a posteriori* para cada u_i é dada em (3.10). Lembre-se neste caso a distribuição para u é uma $beta(\nu, 1)$ e portanto, a distribuição condicional *a posteriori* para u é dada por,

$$\pi(u_i|\lambda, \beta, \nu, u, y, x) \propto G\left(\frac{3}{2}, 2\lambda \sinh^2\left(\frac{y_i - x_i^T \beta}{2}\right)\right) \times h(u_i, \nu),$$

e assim,

$$u_i \sim G\left(\frac{3}{2} + \nu, 2\lambda \sinh^2\left(\frac{y_i - x_i^T \beta}{2}\right)\right) \times I_{(0,1)},\tag{3.13}$$

e a distribuição condicional *a posteriori* para ν é dada por,

$$\pi(\nu|\lambda, \beta, u, y, x) \propto \pi(\nu) \prod_{i=1}^n h(u_i, \nu).$$

Portanto

$$\nu \sim G\left(a, b - \sum_{i=1}^n \log U_i\right).\tag{3.14}$$

Normal Contaminada

Os possíveis estados dos pesos para u_i são v ou 1 , com $\boldsymbol{\nu} = (\nu, v)^T$. Uma distribuição uniforme $U(0, 1)$ é usada como *priori* para ν , e uma distribuição *beta*(a, b) é adaptada para v . A distribuição condicional *a posteriori* para u , é dada por

$$\pi(u_i | \lambda, \boldsymbol{\beta}, \boldsymbol{\nu}, u, y, x) \propto G\left(\frac{3}{2}, 2\lambda \sinh^2\left(\frac{y_i - x_i^T \boldsymbol{\beta}}{2}\right)\right) \times h(u_i, \boldsymbol{\nu}),$$

Ou seja,

$$\pi(u_i | \lambda, \boldsymbol{\beta}, \boldsymbol{\nu}, u, y, x) \propto \begin{cases} G\left(\frac{3}{2}, 2\lambda \sinh^2\left(\frac{y_i - x_i^T \boldsymbol{\beta}}{2}\right)\right) \nu, & \text{se } u_i = v, \\ (1 - \nu) G\left(\frac{3}{2}, 2\lambda \sinh^2\left(\frac{y_i - x_i^T \boldsymbol{\beta}}{2}\right)\right), & \text{se } u_i = 1. \end{cases} \quad (3.15)$$

A distribuição condicional *a posteriori* para ν é dada por,

$$\nu | \lambda, \boldsymbol{\beta}, v, u, y, x \sim \text{beta}\left(a + \frac{n - \sum_{i=1}^n u_i}{1 - v}; b + \frac{\sum_{i=1}^n u_i - n\nu}{1 - v}\right). \quad (3.16)$$

A distribuição condicional *a posteriori* para v é dada por,

$$\pi(v | \lambda, \boldsymbol{\beta}, \nu, u, y, x) \propto \nu^{\left(\frac{n - \sum u_i}{1 - \nu}\right)} \times (1 - \nu)^{\left(\frac{\sum u_i - n\nu}{1 - \nu}\right)}. \quad (3.17)$$

3.2 Aplicação

Dados artificiais

Para examinar o desempenho das medidas de diagnóstico dos modelos propostos, consideramos para o problema (3.1) um conjunto de dados simulados. Sem perda de generalidade, neste caso consideramos um modelo com uma covariável, $\epsilon_i \sim \text{logBS}(0.4, 0)$, consideramos $\beta_0 = 12$, $\beta_1 = -0.5$; x_1 , foi gerada a partir de uma gama com parâmetros (1, 0.001), e com tamanho da amostra $n = 40$, e para mostrar os resultados de nosso modelo foram perturbadas as observações 8, 15, 22, 36, 37 e algumas combinações delas, as perturbações foram com relação à variável resposta $y_i^* = y_i + 5S_y$, $i=8, 15, 22, 36, 37$, sendo S_y o desvio padrão dos y_i 's.

Para analisarmos os dados gerados sob uma perspectiva Bayesiana, inicialmente é considerado o modelo log BS, uma *priori* para $\lambda \sim G(1, 0.0001)$. Também assumimos as *priori*, $\beta_0 \sim N(0, 20)$ e $\beta_1 \sim N(0, 20)$. Geramos uma cadeia com 100000

iterações cada. As 5000 primeiras foram descartadas para reduzir a possível influência do ponto inicial da cadeia. As restantes foram selecionadas de 5 em 5 para evitar correlações nas séries, resultando em uma amostra de 19000 observações. Para verificar a convergência do algoritmo, usamos técnicas gráficas e o método de Geweke (Cowles & Carlin, 1996).

Com a finalidade de mostrar a robustez do modelo log BSSMN, foi considerado o modelo log BS-t, e adotamos uma *priori* para $\nu \sim \text{Ex}(0.1)\mathbb{I}_{(2,\infty)}(\nu)$, sendo que Ex denota a distribuição exponencial, as *priori* para os parâmetros β_0, β_1 , foram os mesmos que o caso log BS. As amostras obtidas para distribuição *a posteriori* dos parâmetros foram semelhantes ao caso log BS.

Na Tabela 3.1 mostramos o sumário *a posteriori* do modelo de regressão linear cujos erros têm distribuição log BS, com a média e o desvio padrão (DP) dos parâmetros. Observamos na primeira linha que quando não temos observações perturbadas, a estimativa dos parâmetros do modelos log BS são próximas aos verdadeiros valores, mas quando temos observações influentes em cada um dos casos essas estimativas são distantes dos verdadeiros valores. Na Tabela 3.2 apresentamos o sumário *a posteriori* do modelo de regressão linear cujos erros têm distribuição log BS-t, e observamos que, na primeira linha, quando não temos observações perturbadas a estimativa dos parâmetros é parecida aos valores estimados usando o modelo log BS, os graus de liberdade estimados (ν) são elevados, o que é uma indicativa que o modelo log BS-t se aproxima ao modelo log BS(gerado por uma normal); mas quando temos observações influentes, as estimativas para o parâmetro β_0 para cada um dos casos perturbados são próximas aos verdadeiros valores, o mesmo acontece com a estimativa do parâmetro β_1 .

Na Tabela 3.3 apresentamos o critério de comparação de modelos para os dados simulados para os dois modelos. Observamos que os valores B e DIC para o caso sem perturbação (primeira linha) são próximos nos dois modelos, mas nos casos perturbados em cada um deles os valores DIC e B, dos modelos log BS e log BS-t apresentam uma grande diferença, em ambos casos indicaram que o modelo log BS-t se ajusto melhor aos dados, o que mostra que o modelo log BS-t é robusto na presença de observações influentes.

Finalmente na Tabela 3.4 apresentamos as medidas de divergência Kullback-Leibler (as quais foram obtidas a partir da amostra *a posteriori* dos parâmetros do modelo) e suas respectivas calibrações para o modelo log BS e log BS-t, no qual notamos que essas medidas, usando o modelo log BS são elevadas mas, são reduzidos consideravelmente quando usamos o modelo log BS-t. Além disso apresentamos na Figura 3.1 e 3.2 os respectivos gráficos de K-L, para os diferentes casos nos quais foram perturbadas.

Tabela 3.1: Sumário *a posteriori* do modelo de regressão linear log BS para os dados simulados.

Casos perturbados	β_0		β_1		λ	
	média	DP	média	DP	média	DP
Nenhum	11,7	0,3	-0,5	0,04	5,4	1,2
8	10,1	1,1	-0,1	0,2	0,4	0,1
22	13,5	0,3	-0,7	0,1	1,8	0,4
36	10,8	0,6	-0,3	0,1	1,2	0,3
8;36	9,9	1,2	0,0	0,2	0,4	0,1
8;22;36	12,6	0,7	-0,4	0,1	0,3	0,1
8;15;22;36;37	13,0	0,8	-0,4	0,2	0,2	0,1

Tabela 3.2: Sumário *a posteriori* do modelo de regressão linear log BS-t para os dados simulados.

Casos perturbados	β_0		β_1		λ		ν	
	média	DP	média	DP	média	DP	média	DP
Nenhum	11,7	0,3	-0,5	0,04	5,8	1,4	18,3	11,3
8	11,7	0,3	-0,5	0,1	7,1	2,3	2,9	0,7
22	12,1	0,4	-0,5	0,1	6,5	2,1	3,6	1,4
36	11,7	0,3	-0,4	0,1	6,4	2,1	3,6	1,3
8;36	11,7	0,3	-0,5	0,1	6,8	2,3	2,5	0,4
8;22;36	12,0	0,3	-0,5	0,1	6,6	2,3	2,3	0,3
8;15;22;36;37	11,9	0,4	-0,5	0,1	5,5	2,0	2,2	0,2

Tabela 3.3: Critério de comparação de modelos para os dados simulados.

Casos perturbados	Log-BS		LogBS-t	
	B	DIC	B	DIC
SP	-25,13	50,77	-26,25	53,48
8	-91,41	145,2	-40,65	76,08
22	-60,68	92,28	-39,84	72,77
36	-59,69	106,38	-39,00	72,82
8;36	-84,5	146,6	-43,91	86,02
8;22;36	-81,49	149,99	-48,34	95,45
8,15;22;36;37	-80,59	156,45	-58,91	117,28

Tabela 3.4: Medida de divergência de Kullback-Leibler e suas calibrações para log BS e log BS-t, para os dados simulados do modelo de regressão linear.

Pert.	Id.	Log BS		Log BS-t	
		casos	$K(P, P_i)$	Cal	$K(P, P_i)$
SP	8	0,11	0,72	0,09	0,71
SP	22	0,25	0,81	0,23	0,80
SP	36	0,09	0,70	0,07	0,68
	8	9,37	1,00	2,70	0,99
	22	6,91	1,00	3,52	0,99
	36	8,36	1,00	2,98	0,99
8;36	8	6,99	1,00	0,78	0,95
	36	0,09	0,69	0,29	0,83
8;22;36	8	4,81	0,99	0,31	0,84
	22	3,24	0,99	0,20	0,78
	36	0,08	0,69	0,16	0,76
8;15;22	8	1,59	0,99	0,14	0,74
36;37	15	0,91	0,96	0,13	0,73
	22	1,13	0,97	0,13	0,74
	36	0,01	0,58	0,10	0,71
	37	0,01	0,58	0,09	0,71

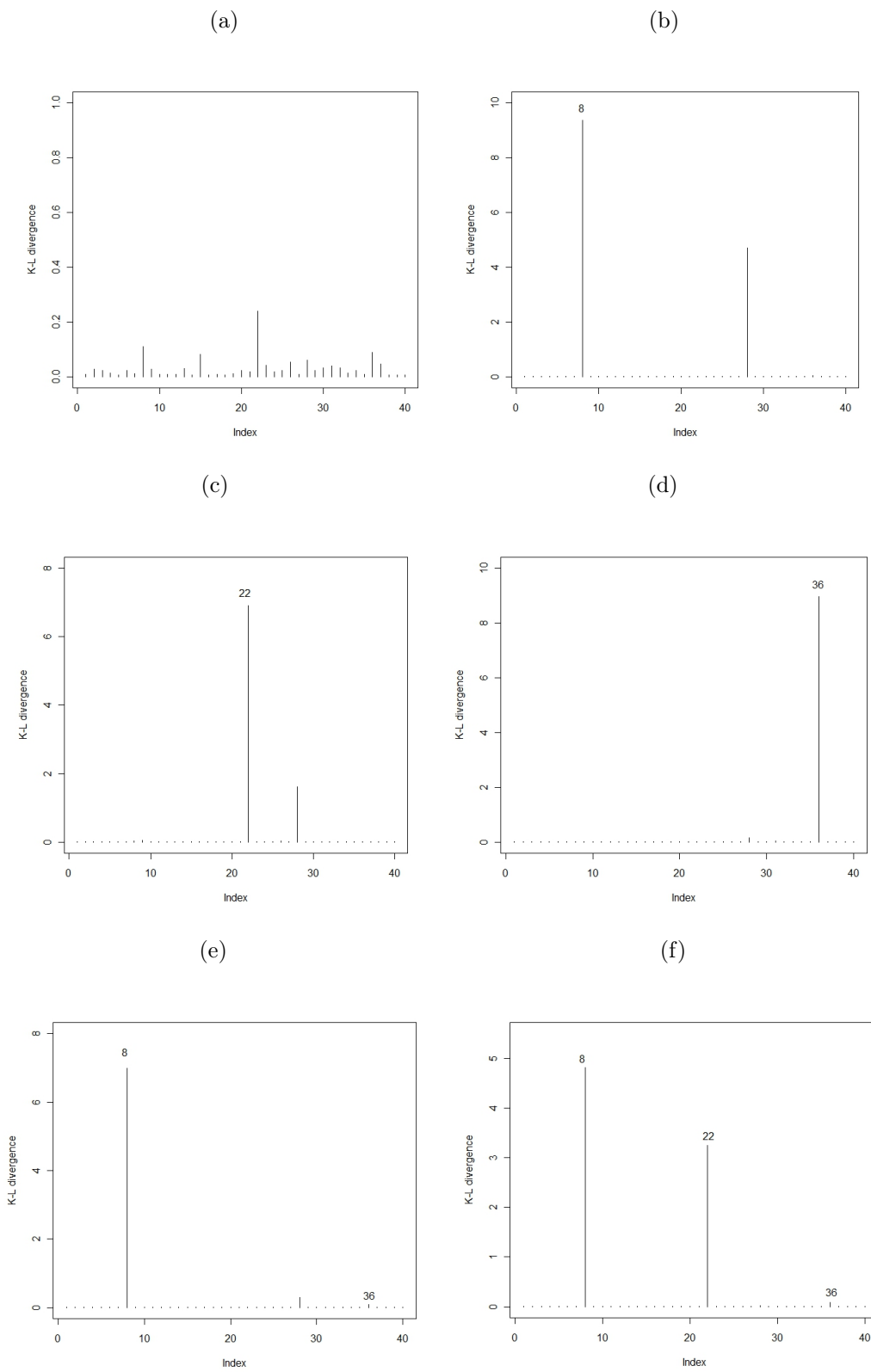


Figura 3.1: Medida de divergência de Kullback-Leibler para os diferentes casos simulados para o modelo log BS.

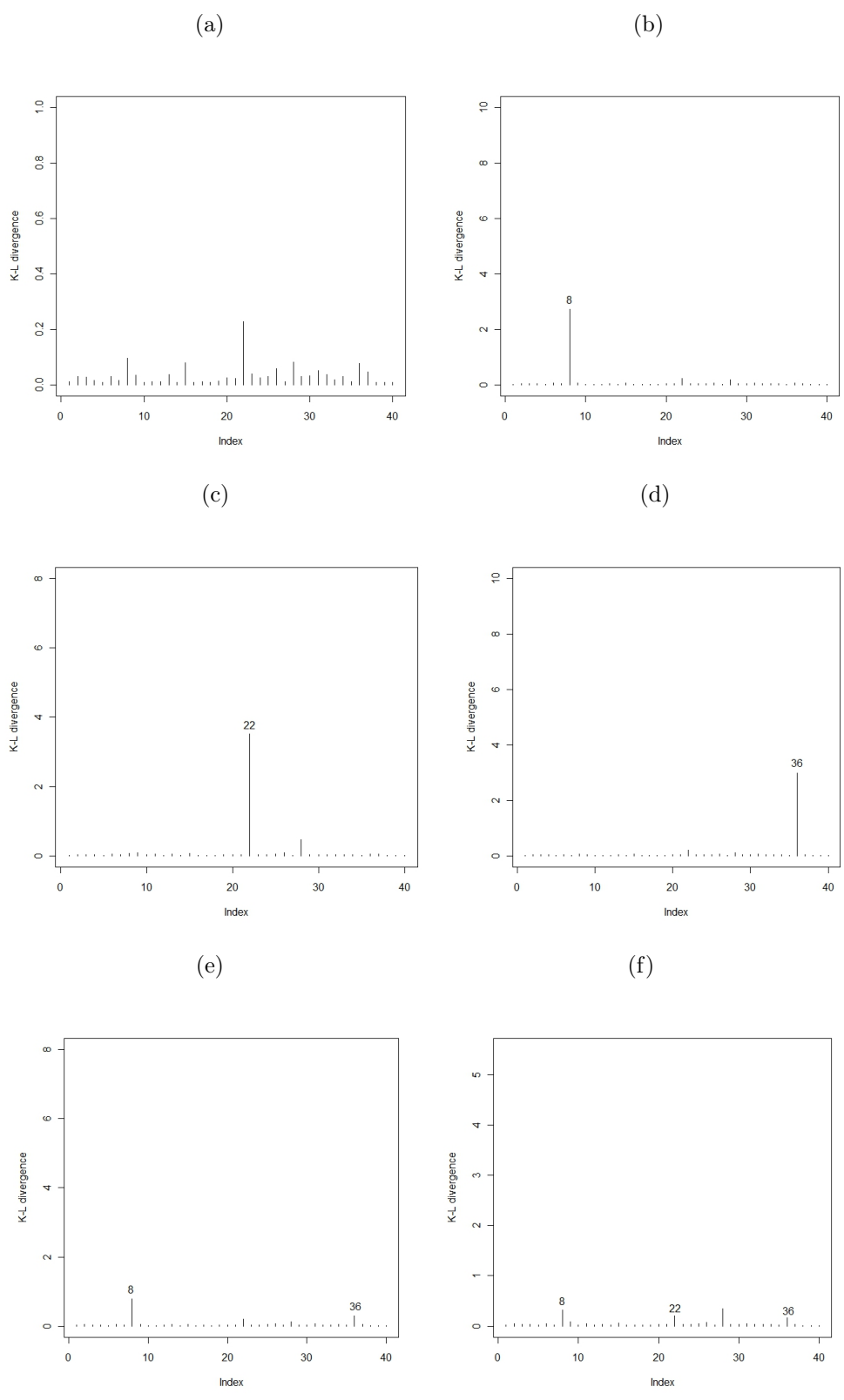


Figura 3.2: Medida de divergência de Kullback-Leibler para os diferentes casos simulados para o modelo log BS-t.

Aplicação a dados de teste de vida

Vamos agora analisar os dados de teste de vida acelerado apresentado por Rieck & Nedelman (1991), em que N_i denota o número de ciclos, até a falha de uma amostra de metal que é submetido a alongamento cíclico e compressão, W_i denota o trabalho por ciclo (MJ/m^3), os autores consideraram o seguinte modelo linear.

$$\log N_i = \beta_1 + \beta_2 \cdot \log W_i + \epsilon_i, \quad i = 1, 2, \dots, 46,$$

sendo ϵ_i segue uma distribuição $\log \text{BS}(\alpha, 0)$.

Consideramos o problema sob a perspectiva Bayesiana porém agora $\epsilon_i \sim \log \text{BSSMN}(\alpha, 0; H)$. Inicialmente foi ajustado usando o modelo $\log \text{BS}$, e a informação *a priori* adotada para $\lambda \sim G(1.2, 0.0001)$, para $\beta_1 \sim N(0, 20)$, $\beta_2 \sim N(0, 20)$. As amostras obtidas para a distribuição *a posteriori* dos parâmetros foram semelhantes à simulação.

Para o modelo $\log \text{BS-t}$ consideramos *a priori* para $\nu \sim \text{Ex}(0.1)\mathbb{I}_{(2, \infty)}(\nu)$. Para o modelo $\log \text{BSSL}$ consideramos *a priori* para $\nu \sim \text{Ex}(0.1)$. Para o modelo $\log \text{BSNC}$ consideramos *a priori* para $\nu \sim \text{beta}(2, 2)$, e $v \sim \text{beta}(2, 2)$. Para aplicar os resultados obtidos, foram geradas observações influentes artificiais perturbando 2 observações (observações 5 e 15) com 2 desvios padrão. Os dados são apresentados na Tabela 3.5.

Os resultados obtidos sob a perspectiva Bayesiana para os dados originais usando a distribuição $\log \text{BS}$ foram $\beta_1=12,1$, $\beta_2 = -1,6$ e $\lambda = 6,0$.

Na Tabela 3.6, apresentamos o resumo *a posteriori* dos parâmetros para os dados de fadiga biaxial para os modelos $\log \text{BS}$, $\log \text{BS-t}$, $\log \text{BSSL}$ e $\log \text{BSNC}$. Na seleção de modelos observamos que os modelos $\log \text{BSSMN}$ se ajustam melhor aos dados que o modelo $\log \text{BS}$. Entre os modelos $\log \text{BSSMN}$ podemos destacar que o modelo $\log \text{BS-t}$ se ajusta melhor aos dados na presença de observações influentes; também apresentamos os gráficos das medidas de divergência de Kullback Leibler para os modelos $\log \text{BSSMN}$, em que essas medidas são reduzidas quando usamos os modelos $\log \text{BSSMN}$. Na Figura 3.5(a) mostra o diagrama de dispersão dos dados, (b) Diagrama de dispersão transformada com os ajustes do modelo de regressão estimada dos modelos $\log \text{BSSMN}$.

Tabela 3.5: Dados fadiga biaxial (Bronw & Miller (1974)).

W	N	W	N	W	N	W	N
11,5	3280	24	804	40,1	750	60,3	283
13	5046	24,6	1093	40,1	316	60,5	212
14,3	1563	25,2	1125	43	456	62,1	327
15,6	4707	25,5	884	44,1	552	62,8	373
16	977	26,3	1300	46,5	355	66,5	125
17,3	2834	27,9	852	47,3	242	67	187
19,3	2266	28,3	580	48,7	190	67,1	135
21,1	2208	28,4	1066	52,9	127	67,9	245
21,5	1040	28,6	1114	56,6	185	68,8	137
22,6	700	30,9	386	59,9	255	75,4	200
22,6	1583	31,9	745	60,2	195	100,5	190
24	482	34,5	736				

Tabela 3.6: Resumo *a posteriori* dos parâmetros para os dados de fadiga biaxial.

	Log-BS		Log BS-t		log BSSL		log BSNC	
	média	DP	média	DP	média	DP	média	Dp
β_1	12,9	0,5	12,6	0,5	12,6	0,5	12,7	0,5
β_2	-1,8	0,1	-1,7	0,1	-1,7	0,1	-1,8	0,1
λ	3,3	0,7	6,5	2,1	10,9	4,8	7,9	4,6
ν			6,1	4,3	1,7	1,0	0,4	0,2
v							0,3	0,2
	B	DIC	B	DIC	B	DIC	B	DIC
	-41,59	79,15	-37,33	72,52	-37,47	72,53	-38,94	74,11

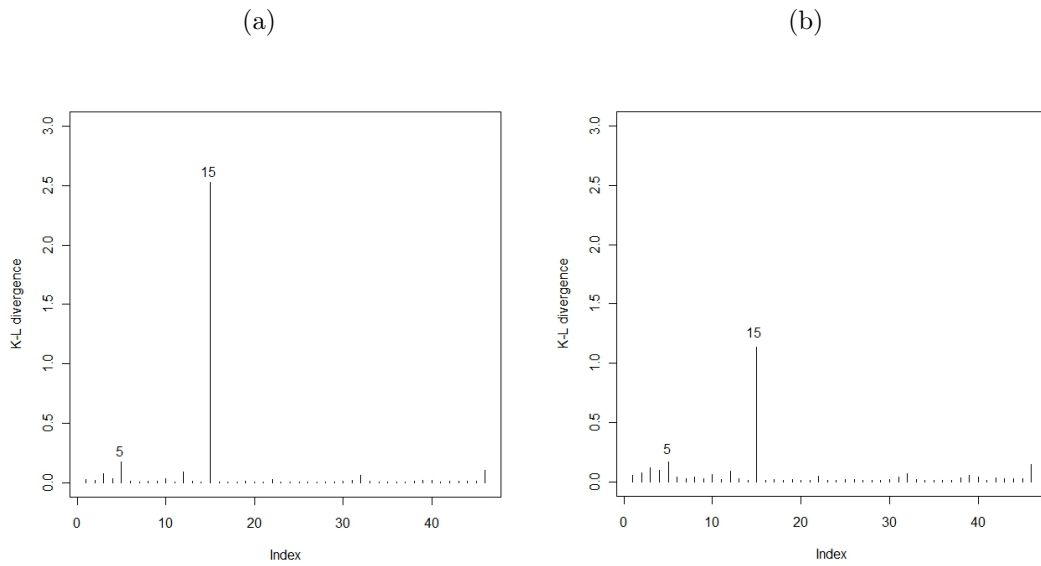


Figura 3.3: Medida de divergência K-L para os dados de fadiga, (a) log BS, (b) log BS-t.

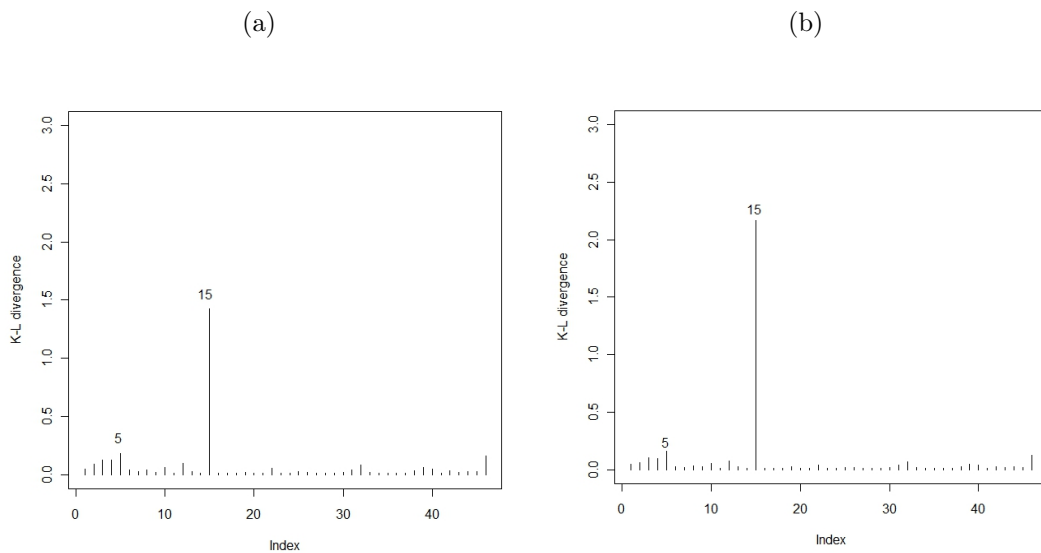
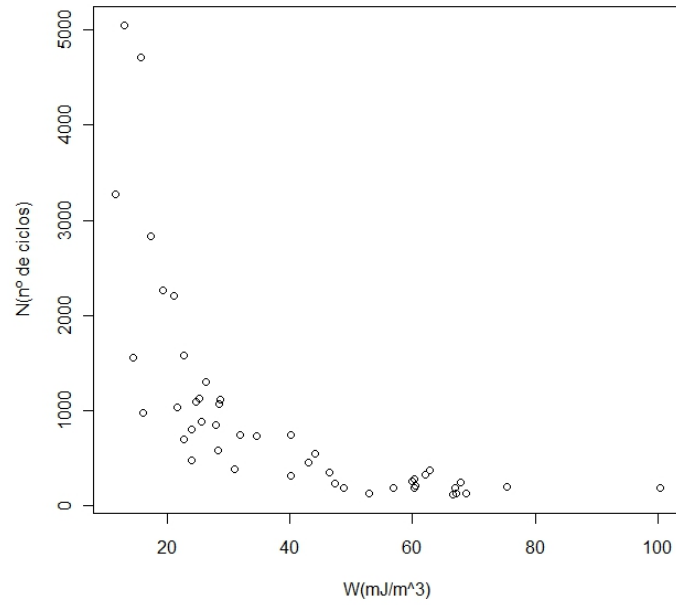


Figura 3.4: Medida de divergência K-L para os dados de fadiga, (a) log BSSL, (b) log BSNC.

(a)



(b)

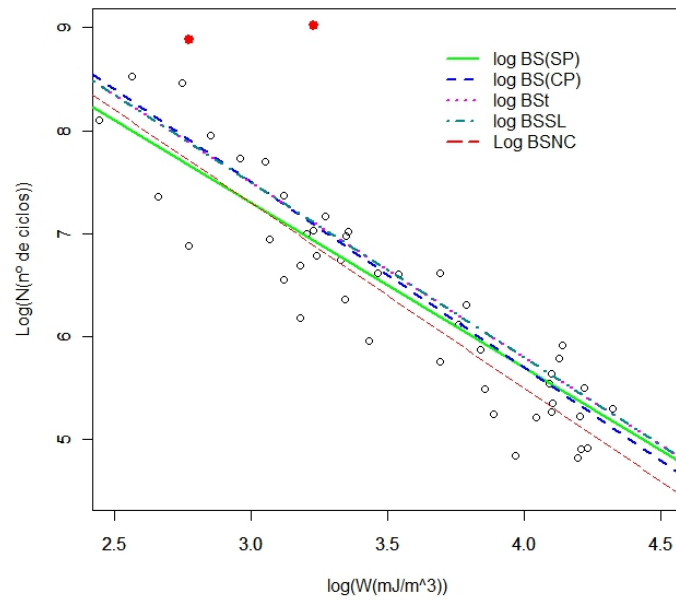


Figura 3.5: (a) Diagrama de dispersão dos dados, (b) Diagrama de dispersão transformada com os ajustes do modelo de regressão estimada dos modelos \log BSSMN.

3.3 Modelos não lineares

As hipóteses de linearidade entre duas ou mais variáveis é forte e muitas vezes não são cumpridas, pois existem problemas nos quais a relação entre as variáveis resposta e a função dos parâmetros não é linear. Nestas circunstâncias existem outros tipos de funções que poderiam ser usadas para ajustar os dados, mediante funções não lineares. Este tipo de funções têm muitas aplicações em química, física, biologia, etc. Um modelo não linear é um modelo no qual ao menos um dos parâmetros aparece em forma não linear. Considere

$$y = \mu(\boldsymbol{\beta}, x) + \epsilon, \quad (3.18)$$

y sendo a variável resposta, e $\boldsymbol{\beta}$ o vetor de parâmetros.

- Por exemplo, em química, uma amostra de gás para manter uma temperatura constante, considera-se o volume ν , a pressão p satisfazendo a seguinte relação $p\nu^\gamma = c$, considerando $y = p$, $x = \nu^{-1}$

$$y = cx^\gamma + \epsilon = \mu(x; c, \gamma) + \epsilon,$$

em que γ é uma constante para cada gás a qual deve ser estimada, c é uma constante a ser estimada. Este modelo particular é um modelo não linear, pois é não linear em um dos parâmetros γ .

- Outro exemplo é um modelo com muita aplicação na indústria elétrica e química, é dada por

$$y = \alpha + \delta x + \beta \rho^x + \epsilon \quad 0 < \rho < 1.$$

Recentemente Lemonte & Cordeiro(2009) propuseram uma classe de modelos de regressão não linear,

$$Y_i = \mu(\boldsymbol{\beta}, x_i) + \epsilon_i \quad i = 1, 2, \dots, n,$$

sendo Y_i é a variável resposta (o logaritmo do tempo de vida da i -ésima observação). Os autores consideraram uma estrutura não linear para a média dos parâmetros $\mu_i = \mu(\boldsymbol{\beta}, x_i)$, em que $\mu(\cdot; \cdot)$ é uma função injetiva, contínua e duas vezes diferenciável com respeito ao vetor de parâmetros $\boldsymbol{\beta}$, \boldsymbol{x}_i é um vetor de valores de variáveis explicativas associado com a i -ésima resposta observável y_i , e ϵ_i são variáveis aleatórias independentes e identicamente distribuídas log BS $(\alpha, 0, 2)$. Além disso, assumiram

que a matriz das derivadas $X = X(\beta) = \partial\mu/\partial\beta^T$ tem posto p , dada uma amostra aleatória o log-verossimilhança $l(\beta, \alpha)$ pode ser escrita como,

$$l(\beta, \alpha) = \sum_{i=1}^n \log(\xi_{i1}) - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2,$$

em que

$$\xi_{i1} = \xi_{i1}(\theta) = \frac{2}{\alpha} \cosh\left(\frac{y_i - \mu_i}{2}\right), \quad \xi_{i2}(\theta) = \sinh\left(\frac{y_i - \mu_i}{2}\right), \quad i = 1, \dots, n.$$

Um algoritmo para calcular o estimador de máxima verossimilhança $\hat{\beta}$ e $\hat{\alpha}$ de β e α é dado por Lemonte & Cordeiro (2009),

$$\left(X^{(m)T} X^{(m)}\right) \beta^{(m+1)} = X^{(m)T} \zeta^{(m)}, \quad \alpha^{(m+1)} = \frac{1}{2} \alpha^{(m)} \left(1 + \bar{\xi}_2^{(m)}\right), \quad m = 0, 1, \dots,$$

em que

$$\zeta^{(m)} = X^{(m)} \beta^{(m)} + \frac{4}{\psi(\alpha^m)} s^m, \quad s = s(\theta) = (s_1, \dots, s_n)^T \quad \text{com}$$

$$s_i = (\xi_{i1} \xi_{i2} - \xi_{i2} / \xi_{i1}) / 2, \quad \bar{\xi}_2^{(m)} = 1/n \sum_{i=1}^n \xi_{i2}^{2(m)}, \quad \psi(\alpha) = 2 + \frac{4}{\alpha^2} - \frac{\sqrt{2\pi}}{\alpha} \psi_0(\alpha) \text{ e}$$

$$\psi_0(\alpha) = \left\{ 1 - \operatorname{erf}\left(\frac{\sqrt{2}}{\alpha}\right) \right\} \exp\left(\frac{2}{\alpha^2}\right),$$

sendo que erf denota a função $\operatorname{erf}(x) = (2/\sqrt{\pi}) \int_0^x \exp(-t^2) dt$.

Existe um pacote em R chamado “library bs”, no qual se pode estimar os parâmetros sob a perspectiva clássica. Neste estudo abordamos o problema de estimação sob a perspectiva Bayesiana, em que os erros têm distribuição log BSSMN.

3.3.1 Modelo não linear log-Birnbaum-Saunders com mistura de escala normal log BSSMN

Consideremos o seguinte modelo

$$Y_i = \mu(\beta, x_i) + \epsilon_i \quad i = 1, 2, \dots, n, \quad (3.19)$$

sendo y_i o log do tempo de vida observado para o i -ésimo indivíduo, $\boldsymbol{\beta}=(\beta_1, \dots, \beta_p)^T$ um vetor de parâmetros desconhecidos para ser estimados, $\mu(\cdot)$ uma função não linear com respeito aos parâmetros, \mathbf{x} contém os valores das variáveis exploratórias, e $\epsilon_i \sim \log \text{BSSMN}(\alpha, 0; H)$.

A função de verossimilhança para os parâmetros do modelo (3.19), para um conjunto de n observações é dada por

$$L(\alpha, \boldsymbol{\beta}; H|\mathbf{x}, \mathbf{y}) \propto \prod_{i=1}^n \frac{1}{\alpha} \cosh\left(\frac{y_i - \mu(\boldsymbol{\beta}, x_i)}{2}\right) \phi_{SMN}\left(\frac{2}{\alpha} \sinh\left(\frac{y_i - \mu(\boldsymbol{\beta}, x_i)}{2}\right); H\right).$$

Como no caso linear, introduzindo a variável de mistura, temos que a função verossimilhança ampliada para os parâmetros pode ser expressa por

$$\begin{aligned} L(\alpha, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{u}|\mathbf{y}, \mathbf{x}) &\propto \alpha^{-n} \prod_{i=1}^n u_i^{\frac{1}{2}} \cosh\left(\frac{y_i - \mu(\boldsymbol{\beta}, x_i)}{2}\right) \\ &\times \exp\left(-\frac{2}{\alpha^2} \sum_{i=1}^n u_i \sinh^2\left(\frac{y_i - \mu(\boldsymbol{\beta}, x_i)}{2}\right)\right) \times h(u_i, \boldsymbol{\nu}), \end{aligned} \quad (3.20)$$

em que $h(u_i, \boldsymbol{\nu})$ é a função de distribuição de probabilidade da variável de mistura U_i ; $\mu(\boldsymbol{\beta}, x_i)$ é uma função não linear. Como no caso do modelo linear consideramos a reparametrização $\lambda = 1/\alpha^2$, assim a verossimilhança é dada por

$$\begin{aligned} L(\lambda, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{u}|\mathbf{y}, \mathbf{x}) &\propto \lambda^{\frac{n}{2}} \prod_{i=1}^n u_i^{\frac{1}{2}} \cosh\left(\frac{y_i - \mu(\boldsymbol{\beta}, x_i)}{2}\right) \\ &\times \exp\left(-2\lambda \sum_{i=1}^n u_i \sinh^2\left(\frac{y_i - \mu(\boldsymbol{\beta}, x_i)}{2}\right)\right) \times h(u_i, \boldsymbol{\nu}). \end{aligned} \quad (3.21)$$

Como no caso linear consideramos uma distribuição *a priori* conjunta própria para os parâmetros do modelo (3.19). As *priori* a serem consideradas dependem da função não linear que vai ser trabalhada, também consideramos os parâmetros independentes com a distribuição *a priori* dada por (3.6) e *a posteriori* conjunta é dada por

$$\begin{aligned} \pi(\lambda, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{u}|\mathbf{x}, \mathbf{y}) &\propto \lambda^{\frac{n}{2}} \prod_{i=1}^n \left[u_i^{\frac{1}{2}} \cosh\left(\frac{y_i - \mu(\boldsymbol{\beta}, x_i)}{2}\right) \right] \\ &\times \exp\left\{-2\lambda \sum_{i=1}^n u_i \sinh^2\left(\frac{y_i - \mu(\boldsymbol{\beta}, x_i)}{2}\right)\right\} \times h(u_i, \boldsymbol{\nu}) \\ &\times \pi(\lambda, \boldsymbol{\beta}, \boldsymbol{\nu}). \end{aligned} \quad (3.22)$$

Finalmente para aplicar MCMC precisamos das condicionais *a posteriori*, as quais são semelhantes ao caso linear, dadas na Seção 3.1.

3.4 Aplicação

Dados artificiais

Para examinar o desempenho das medidas de diagnóstico dos modelos propostos, consideramos um conjunto de dados simulados com 40 observações para o problema (3.19), em que $\mu(\boldsymbol{\beta}, x_i) = \beta_1 + \beta_2 \exp(\beta_3/x_i)$, $i=1, \dots, 40$, $\epsilon_i \sim \text{log BS}(0.4, 0)$. Geramos $x_i \sim G(1, 0.01)$ fixamos $\boldsymbol{\beta} = [10, -5, -18]$. Para mostrar os resultados dos modelos propostos, perturbamos as observações 1, 6, 11, 27, 30, 35, 36 e algumas combinações delas; as perturbações foram com relação à variável resposta, $y_i^* = y_i + 3S_y$.

Para analisarmos os dados gerados sob a perspectiva Bayesiana, inicialmente consideramos o modelo log BS, *a priori* para o parâmetro $\lambda \sim G(1.2, 0.0001)$, além disso *a priori* para $\beta_1 \sim N(0, 100)$, e $\beta_2 \sim N(-10, 100)$, $\beta_3 \sim N(-20, 100)$. Geramos uma cadeia com 100000 iterações cada, as 5000 primeiras foram descartadas para minimizar a possível influência do ponto inicial da cadeia. As restantes foram selecionadas de 5 em 5 para evitar a correlação nas séries, resultando em uma amostra de 19000 valores. Todo o estudo foi implementado no sistema *Openbugs* conjuntamente com o software estatístico R. A convergência das cadeias foi verificada por métodos gráficos e o método de Geweke.

Para analisar a robustez dos modelos propostos (log BSSMN), foi ajustado o modelo log BS-t, consideramos *a priori* para o parâmetro $\nu \sim \text{Ex}(0.1)\mathbb{I}_{(2, \infty)}(\nu)$. As amostras obtidas para distribuição *a posteriori* dos parâmetros foram semelhantes ao caso log BS.

Na Tabela 3.7 apresentamos os resultados para o modelo log BS, com diferentes dados perturbados. Na primeira linha mostramos os resultados sem nenhuma perturbação, em que podemos observar que os valores estimados são próximos aos verdadeiros valores, quando os dados são perturbados temos que as estimativas são distantes dos verdadeiros valores para cada um dos casos, especificamente para o parâmetro β_3 .

Na Tabela 3.8 apresentamos os resultados com algumas observações perturbadas para o modelo log BS-t. Na primeira linha dispomos os dados sem perturbação, e observamos que as estimativas deste modelo são parecidas com o modelo log BS; além disso observamos que os graus de liberdade estimadas neste caso são altos, o

que nos leva a interpretar que o modelo log BS-t se aproxima de log BS.

Quando temos observações perturbadas as estimativas do modelo log BS-t são próximas aos valores reais. Em todos os casos quando se perturba uma observação, duas, três e cinco observações, além disso, o valor dos graus de liberdade estimados não são tão elevados como quando foi estimado sem perturbação.

Observamos ainda na Tabela 3.9 os valores do DIC e B. Para cada um dos casos, notamos uma diferença notável nos valores, e segundo ambos critérios de seleção, o modelo log BS-t se ajusta melhor aos dados, em cada um dos casos apresentados.

Na Tabela 3.10 apresentamos os valores da medida de divergência de Kullback-Leibler e suas respectivas calibrações, notamos que ao utilizar o modelo log BS-t os valores das estimativas de K-L em alguns casos reduzem consideravelmente, o que pode ser visualizado com os gráficos da medida K-L para os casos perturbados.

Tabela 3.7: Sumário *a posteriori* do modelo de regressão não linear log BS para dados simulados.

Casos	β_1		β_2		β_3		λ	
	média	DP	média	DP	média	DP	média	DP
Nenhum	10,1	0,3	-5,1	0,3	-17,9	2,5	6,2	1,4
1	9,9	0,5	-5,4	0,6	-37,9	8,9	0,8	0,2
6	9,6	0,5	-4,9	0,6	-38,2	10,5	0,8	0,2
30	11,6	2,2	-5,6	2,1	-8,8	4,4	1,0	0,2
1;6	10,2	1,1	-5,8	1,7	-41,7	14,3	0,5	0,2
35;36	11,4	3,8	-4,8	3,8	-8,8	7,4	0,5	0,1
11;25	10,5	1,6	-3,8	1,6	-17,0	10,8	0,3	0,1
1;6;11	10,5	1,2	-3,9	1,2	-23,8	12,0	0,3	0,1
1;11;27;30;36	9,7	3,8	-2,5	3,9	-12,8	11,5	0,3	0,1

Tabela 3.8: Sumário *a posteriori* do modelo de regressão não linear log BS-t para dados simulados.

Casos perturbados	β_1		β_2		β_3		λ		ν	
	média	DP	média	DP	média	DP	média	DP	média	DP
Nenhum	10,1	0,3	-5,2	0,3	-18,3	2,8	6,9	1,7	17,1	11,3
1	10,1	0,3	-5,3	0,3	-18,6	2,6	8,5	2,7	2,8	0,8
6	10,1	0,3	-5,2	0,3	-18,7	3,2	7,6	2,5	3,0	0,8
30	10,2	0,4	-5,2	0,3	-17,4	2,9	8,6	2,6	3,0	0,9
1;6	10,1	0,3	-5,2	0,3	-19,5	3,2	7,4	2,3	2,4	0,4
11;25	10,2	0,3	-5,3	0,3	-18,3	2,4	7,6	2,8	2,4	0,3
35;36	10,3	0,4	-5,2	0,3	-16,7	3,1	7,4	2,4	2,4	0,4
1;6;11	10,1	0,4	-5,3	0,3	-19,1	3,6	7,2	2,7	2,2	0,2
1;11;27;30;36	10,3	0,5	-5,1	0,5	-16,1	3,3	5,7	2,1	2,2	0,2

Tabela 3.9: Critério de comparação de modelos para os dados simulados, do modelo não linear.

Casos perturbados	Log-BS		LogBS-t	
	B	DIC	B	DIC
SP	-23,22	46,19	-24,38	48,82
1	-71,43	121,91	-37,31	70,48
6	-69,84	121,80	-37,34	70,43
30	-63,56	110,72	-36,57	68,38
1;6	-72,93	138,64	-43,11	84,18
35;36	-70,29	133,29	-43,40	84,25
1;6;11	-79,78	152,75	-49,70	98,21
1;11;27;30;36	-77,96	151,85	-61,08	121,21

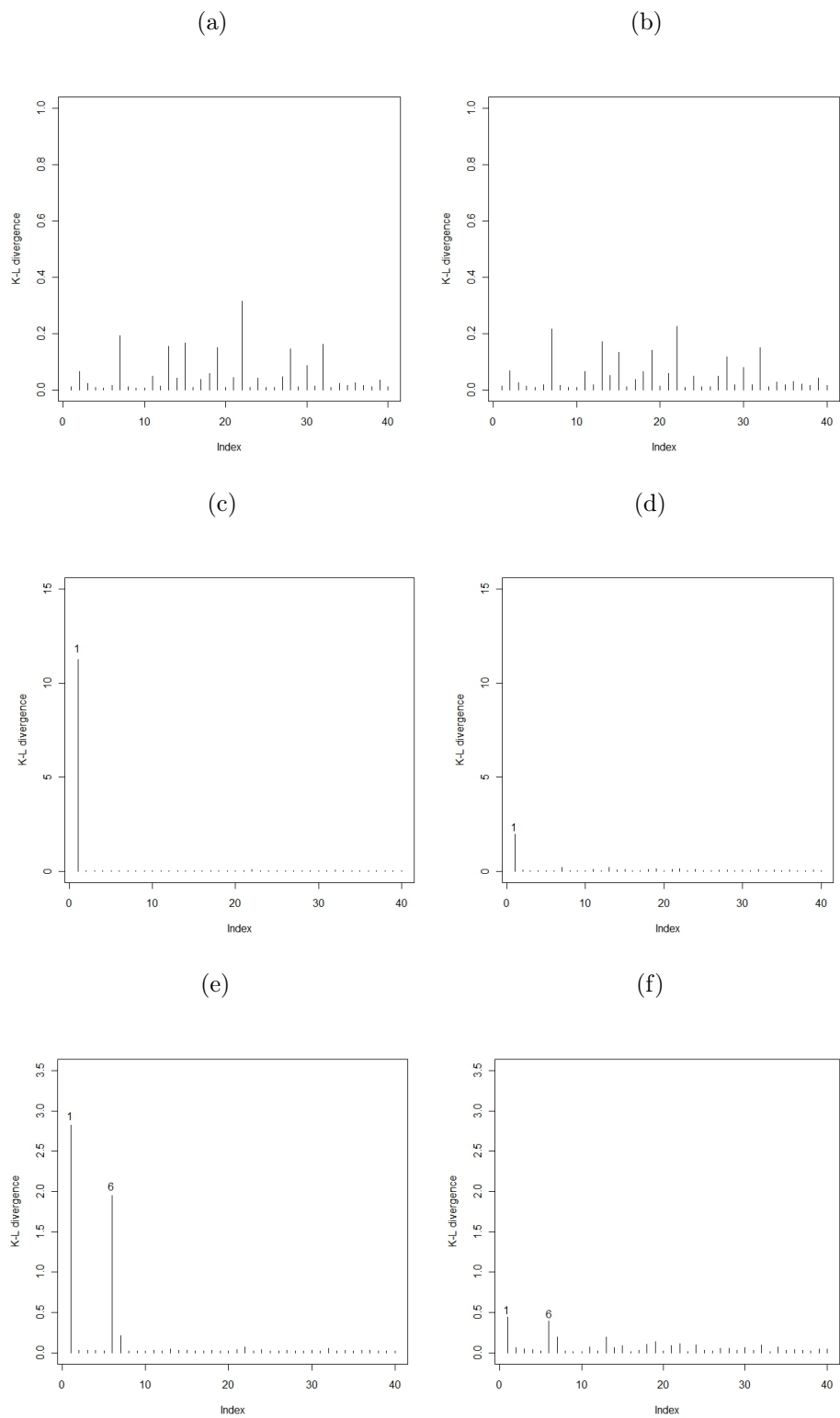


Figura 3.6: Medida de divergência de Kullback-Leibler para os dados simulados para o modelo não linear, (a) log BS (SP), (b) log BS-t (SP); (c) e (e) log BS; (d) e (f) log BS-t

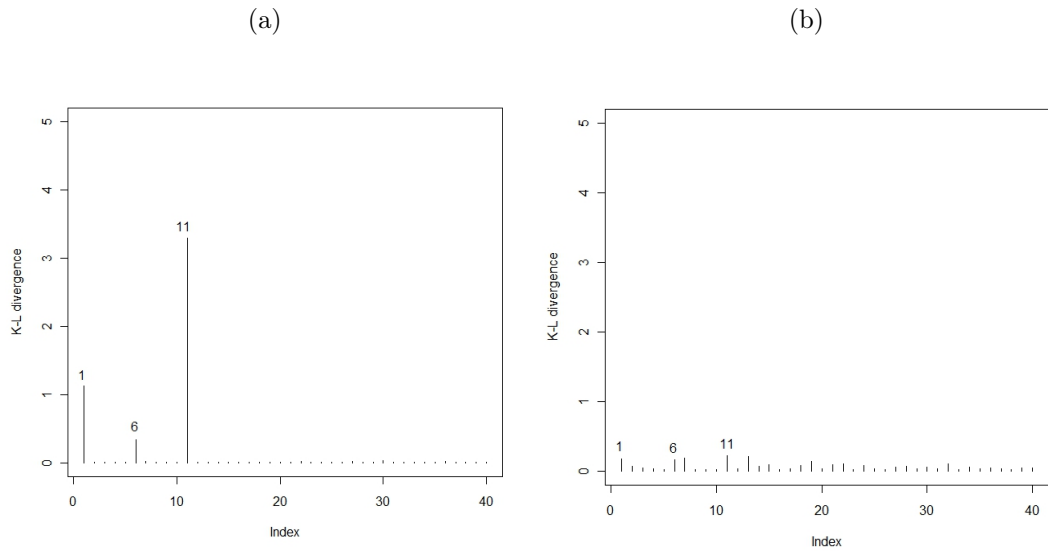


Figura 3.7: Medida de divergência de Kullback-Leibler para os dados simulados com as observações 1, 6, e 11 perturbadas, para o modelo não linear, (a) log BS, (b) log BS-t.

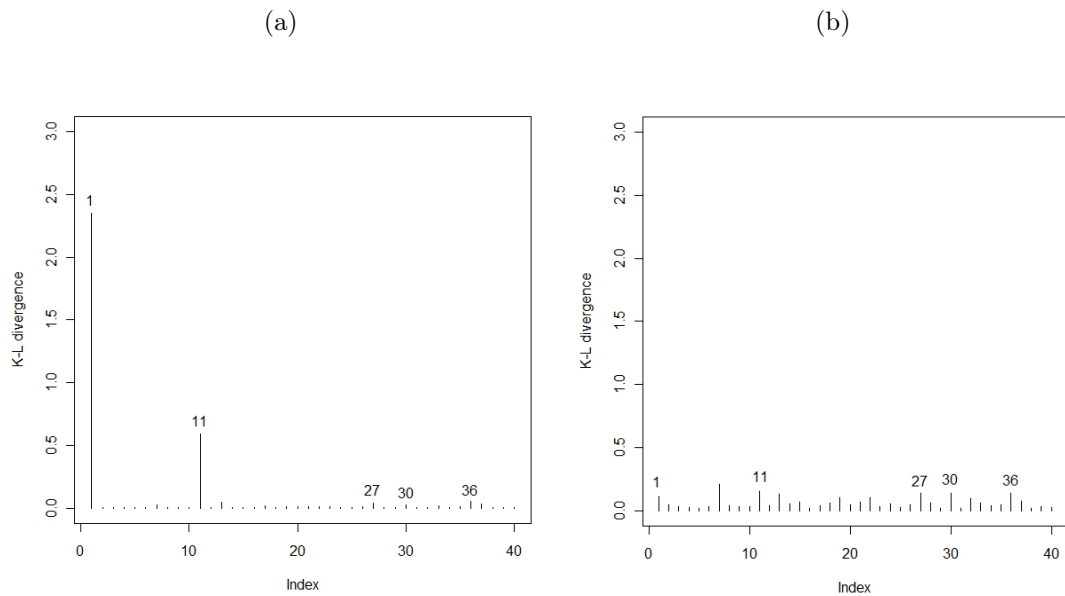


Figura 3.8: Medida de divergência de Kullback-Leibler para os dados simulados com as observações 1; 11; 27; 30 e 36 perturbadas, para o modelo não linear, (a) Log BS, (b) log BS-t.

Tabela 3.10: Medida de divergência de Kullback-Leibler e suas calibrações para dados simulados do modelo de regressão não linear.

Pert.	Id. casos	Log BS		Log BS-t	
		$K(P, P_i)$	Cal	$K(P, P_i)$	Cal
SP	1	0,01	0,57	0,01	0,58
SP	6	0,01	0,58	0,02	0,60
SP	30	0,08	0,69	0,07	0,68
	1	11,26	1,0	1,98	1,0
	6	10,10	1,0	1,91	1,0
	30	8,32	1,0	2,23	1,0
1;6	1	2,82	1,0	0,44	0,88
	6	1,95	1,0	0,39	0,87
11;25	11	3,06	1,0	0,54	0,90
	25	2,08	1,0	0,35	0,85
35;36	35	2,52	1,0	0,62	0,92
	36	1,47	0,98	0,59	0,91
1;6;11	1	1,12	0,97	0,17	0,76
	6	0,33	0,84	0,16	0,76
	11	3,22	1,0	0,22	0,79
1;11;27	1	2,35	1,0	0,11	0,72
30;36	11	0,59	0,91	0,15	0,75
	27	0,03	0,63	0,13	0,74
	30	0,02	0,60	0,13	0,74
	36	0,06	0,66	0,14	0,75

Aplicação não linear

Para o conjunto de dados apresentados por Rieck & Nedelman (1991) sobre a vida de uma peça metálica (Tabela 3.5), consideramos o modelo de regressão não linear apresentado por Lemonte & Cordeiro(2009) dada por

$$y_i = \beta_1 + \beta_2 \exp(\beta_3/W_i) + \epsilon_i, \quad i = 1, \dots, 46, \quad (3.23)$$

onde $y_i = \log(N_i)$, ϵ_i segue uma distribuição log BS.

Consideramos o problema sob a perspectiva Bayesiana onde $\epsilon_i \sim \log \text{BSSMN}(\alpha, 0; H)$ neste caso a função não linear é dada por $\mu(\beta, x_i) = \beta_1 + \beta_2 \exp(\beta_3/x_i)$. Inicialmente foi ajustado o modelo log BS, consideramos uma *priori* para $\lambda \sim G(a, b)$ neste caso $a = 1, 2$ e $b = 0, 0001$, *a priori* para $\beta_1 \sim N(0, 100)$, e $\beta_2 \sim N(-10, 100)$, $\beta_3 \sim N(-20, 100)$, foi perturbado uma observação(observação 5) com 3 DP. As amostras geradas foram semelhantes à simulação.

A fim de demonstrar a robustez do modelo log BSSMN, consideramos o modelo log BS-t, neste caso *a priori* para ν foi semelhante à simulação, as amostras obtidas para distribuição *a posteriori* dos parâmetros também foram semelhantes ao caso log BS. Os resultados obtidos são apresentados na Tabela. 3.11.

Lemonte & Cordeiro (2009) obtiveram o EMV dos parâmetros para o modelo (3.23) os quais foram $\beta_1 = 8.98$, $\beta_2 = -5.18$, $\beta_3 = -22.51$. Usando inferência Bayesiana, os estimadores dos parâmetros sem nenhuma perturbação para o modelo log BS foram $\beta_1 = 9.4$, $\beta_2 = -5.6$, $\beta_3 = 21.0$.

Pode-se observar na Tabela. 3.11, que as estimativas dos parâmetros apresentam uma diferença, sobre todo no valor do parâmetro β_3 além disso os valores de B e DIC indicam que o modelo log BS-t, se ajusta melhor aos dados. Na tabela 3.12 apresentamos a medida de Kullback Leibler, onde observamos que a mesma diminui quando trabalhamos com o modelo log BS-t, além disso apresentamos o gráfico das medidas K-L, e na Figura 3.10 apresentamos o gráfico dos dados transformados com os ajustes do modelo log BS e log BS-t.

Tabela 3.11: Resumo *a posteriori* dos parâmetros para os dados de fadiga biaxial para o modelo de regressão não linear.

	Log-BS		Log BS-t	
	média	DP	média	DP
β_1	10,7	1,5	10,0	1,5
β_2	-6,9	1,3	-6,2	1,3
β_3	-17,2	5,8	-20,2	6,9
λ	3,8	0,8	7,4	2,0
ν			6,2	4,5
	B	DIC	B	DIC
	-41,42	65,48	-34,51	55,59

Tabela 3.12: Medida de de divergência K-L, para o modelo não linear

Caso	log BS		log BS-t	
	K-L	cal	K-L	cal
5	5,32	1,00	1,97	0,98

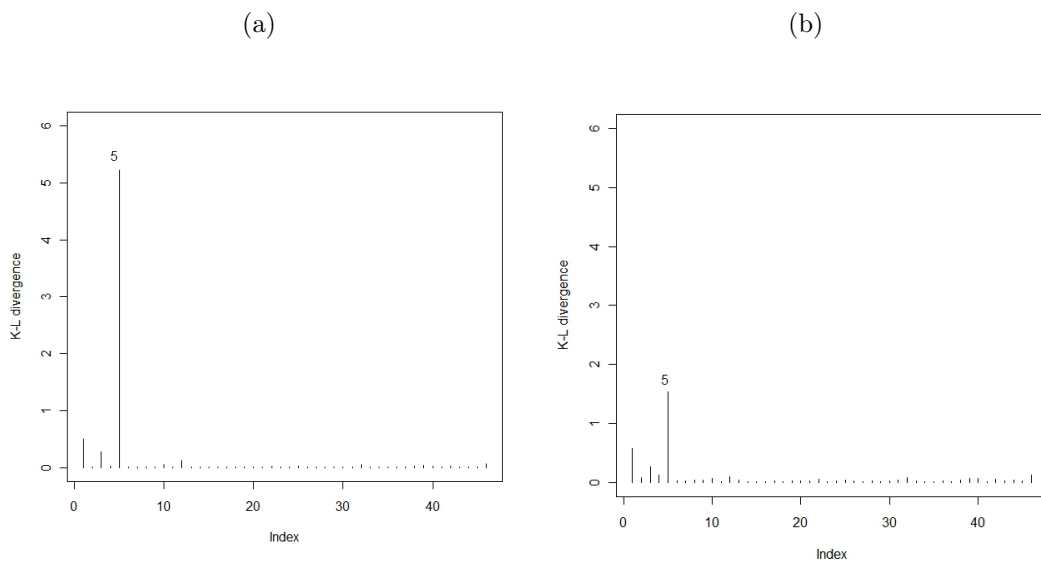


Figura 3.9: (a) Medida de divergência K-L para os dados da aplicação usando log BS, (b) log BS-t; para o modelo de regressão não linear.

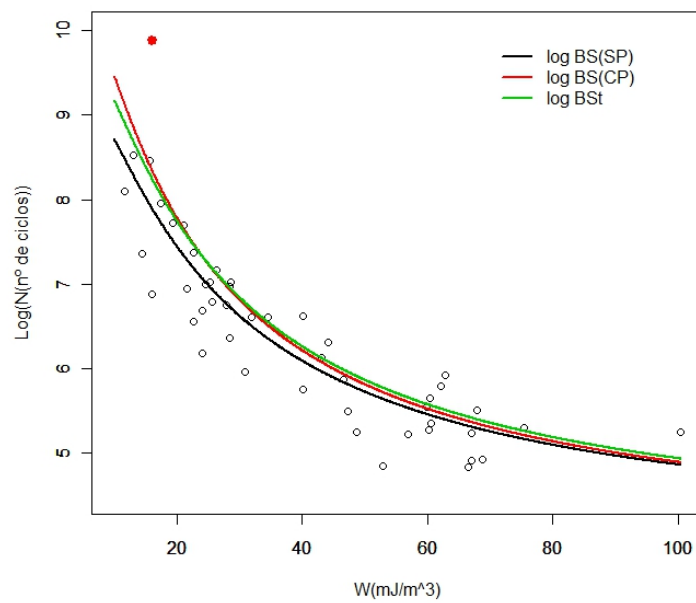


Figura 3.10: Ajuste dos dados de fadiga biaxial.

3.5 Conclusão e observações

Neste estudo apresentamos o modelo de regressão linear e não linear com erros log Birnbaum-Saunders com mistura de escala normal (log BSSMN), como uma alternativa aos modelos de regressão tradicionais. Na simulação e aplicação observamos que os modelos propostos, apresentam estimativas mais robustas na presença de observações influentes.

Capítulo 4

Modelo de regressão geométrico Birnbbaum-Saunders

Em muitos problemas médicos, tais como doenças cardíacas crônicas e diferentes tipos de cancer, um dano cumulativo causado por vários fatores de risco é apresentado. Esta degradação leva a um processo de fadiga, cuja propagação do tempo de vida pode ser modelado adequadamente pela distribuição Birnbbaum-Saunders (BS). Recentemente Leiva *et al.* (2008b), aplicou a versão clássica da distribuição (BS)(modelo gerado de uma normal) para modelar o tempo de sobrevivência de pacientes com múltiplo mieloma.

Neste trabalho propomos uma nova família de distribuição baseado no modelo BS em um cenário de causas competitivas, assumindo que o número de causas é modelado pela distribuição Geométrica e os tempos de vida associados a cada uma dessas causas é modelado pela distribuição Birnbbaum-Saunders.

4.1 Formulação do problema

Para um indivíduo na população, seja N o número de causas não observados do evento de interesse para este indivíduo. Assumimos que N segue uma distribuição geométrica com parâmetro θ e a distribuição de probabilidade é dada por

$$P(N = n) = (1 - \theta)\theta^{n-1}, \quad n = 1, \dots \quad (4.1)$$

O tempo para o j -ésimo caso produzir o evento de interesse é denotado por Z_j , $j=1, \dots, N$. Assumimos que a condicional de Z_j sobre N são iid. com distribuição BS,

além disso, assumimos que Z_1, Z_2, \dots são independentes de N .

O tempo para o evento de interesse ser observado, pode ser definido pela variável aleatória

$$T = \min(Z_1, \dots, Z_N). \quad (4.2)$$

Proposição 4.1. *Se a variável aleatória T é definido como em (4.2), considerando as equações (1.1) e (4.1), T tem distribuição Geométrica Birnbaum-Saunders (GBS) com f.d.p. dada por,*

$$f_T(t) = \frac{(1 - \theta)t^{-\frac{3}{2}}(t + \beta) \exp\left[-\frac{1}{2\alpha^2}\left(\frac{t}{\beta} - \frac{\beta}{t} - 2\right)\right]}{2\alpha\sqrt{\beta}\sqrt{2\pi}\left(1 - \theta\Phi\left[-\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right]\right)^2}, \quad t > 0, \quad (4.3)$$

Prova: A distribuição acumulada da variável aleatória T dada em (4.2) dado $N = m$ é dada por

$$\begin{aligned} F_T(t|N = m) &= Pr[\min(Z_1, Z_2, \dots, Z_m) \leq t] \\ &= 1 - Pr[\min(Z_1, Z_2, \dots, Z_m) > t] \\ &= 1 - Pr^m(Z_1 > t). \\ &= 1 - \left(1 - \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right]\right)^m. \end{aligned}$$

Por tanto a distribuição marginal de T é dada por,

$$\begin{aligned} F_T(t) &= 1 - \sum_{m=1}^{\infty} \left(1 - \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right]\right)^m (1 - \theta)\theta^{m-1} \\ &= \frac{\Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right]}{1 - \theta\Phi\left[-\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right]}. \end{aligned} \quad (4.4)$$

Assim $f_T(t) = dF_T(t)/dt$ o que prova a Proposição.

□

O parâmetro β controla a escala, o parâmetro α e θ controla a forma, quando o parâmetro θ se aproxima a zero, a distribuição GBS vai para a distribuição BS. Na

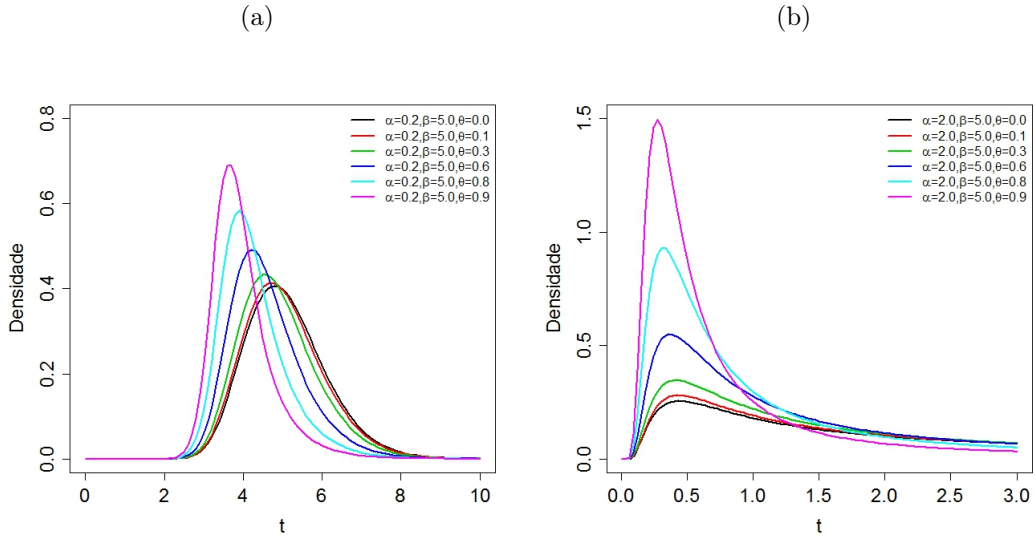


Figura 4.1: Densidade da distribuição GBS com $\theta=0; 0,1; 0,3; 0,6; 0,8; 0,9$; $\beta=5$; $\alpha=0,2$ (gráfico (a)), $\alpha=2$ (gráfico (b)).

Figura 4.1 mostramos a densidade da distribuição GBS para alguns valores de θ , os gráficos mostram que a distribuição GBS é flexível e que o valor θ tem um efeito substancial de assimetria e curtose.

A função de sobrevivência do modelo GBS é dada por,

$$S(t) = \frac{(1 - \theta)\Phi \left[-\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right]}{1 - \theta\Phi \left[-\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right]} \quad t > 0. \quad (4.5)$$

Por (1.1) e (4.5) é fácil verificar que a função de risco é dada por

$$r_T(t) = \frac{r_{BS}(t)}{1 - \theta\Phi \left[-\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right]}, \quad t > 0, \quad (4.6)$$

em que $r_{BS}(t)$ é a função de risco da distribuição BS, observamos pela equação (4.6) que $r_T(t)/r_{BS}(t)$ é crescente em t , com $0 < \theta < 1$, além disso $r_{BS}(t) \leq r_T(t)$ e $\lim_{t \rightarrow \infty} r_T(t) = \frac{1}{2\alpha^2\beta}$ e temos que o comportamento do limite da função de risco $r_T(t)$, é o mesmo que o comportamento da função de risco da distribuição BS. Na Figura 4.2 se apresenta algumas formas da função de risco da distribuição GBS.

O r -ésimo momento da distribuição GBS é dada por

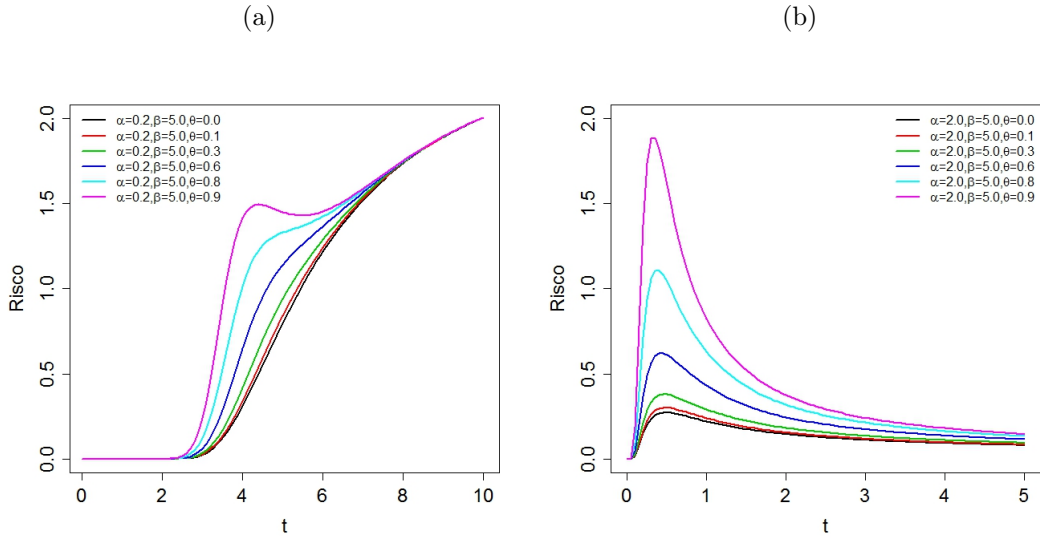


Figura 4.2: Função de risco da distribuição GBS. Os parâmetros são fixados com $\theta=0; 0,1; 0,3; 0,6; 0,8; 0,9$ e $\beta=5$, $\alpha=0,2$ (gráfica (a)); $\alpha=2$ (gráfica (b))

$$\mu' = E(T^r) = (1 - \theta)A_r(\alpha, \beta, \theta), \quad (4.7)$$

onde,

$$A_r(\alpha, \beta, \theta) = \int_0^{\infty} r t^{r-1} \frac{\Phi[-u_t(\alpha, \beta)]}{1 - \theta \Phi[-u_t(\alpha, \beta)]} dt,$$

onde,

$$u_t(\alpha, \beta) = \frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right).$$

Notamos na Equação (4.7) para $\theta = 0$, os momentos correspondem com os momentos da distribuição BS. A medida de assimetria e curtose pode ser calculada pela equação (4.7) numericamente com o software R (usando a função "integrate") ou algum outro método alternativo.

Finalmente na Figura 4.3 se apresenta algumas formas da curtose e assimetria, observamos que são crescentes com respeito a α e θ .

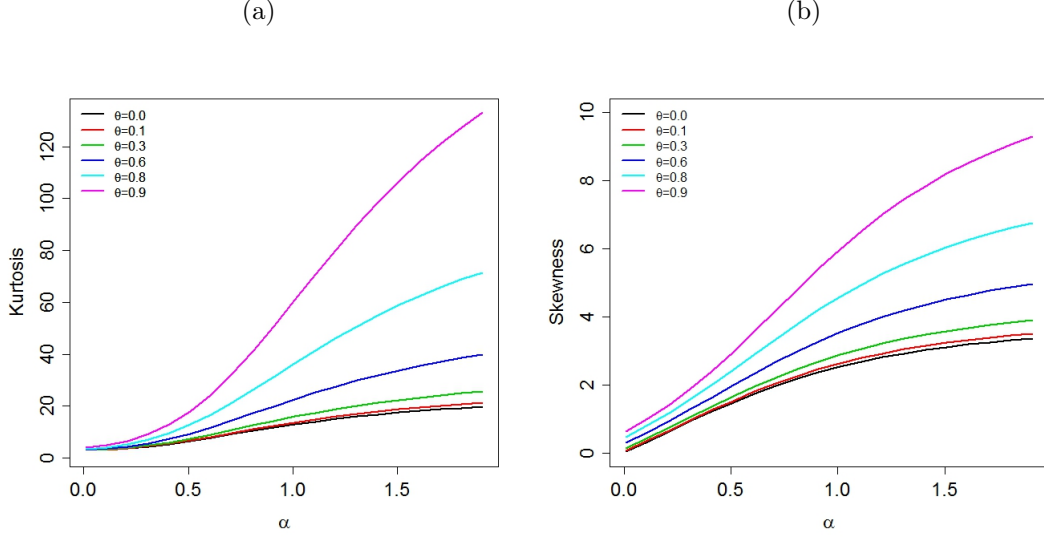


Figura 4.3: (a) Gráfico da curtose com $\mu = \log(5)$ fixo e $\theta=0; 0,1; 0,3; 0,6; 0,8; 0,9$; (b) Skewness com parâmetro $\mu=\log 5$ fixo.

4.2 Modelo de regressão log GBS

Seja T uma variável aleatória com função de distribuição acumulada dada por (4.4). Uma variável aleatória $Y = \log T$ tem uma distribuição log GBS (LGBS), considerando $\eta = \log(\beta)$, a função de distribuição acumulada de Y , é dada por,

$$F_Y(y) = \frac{\Phi \left[\frac{2}{\alpha} \sinh \left(\frac{y-\eta}{2} \right) \right]}{1 - \theta \Phi \left[-\frac{2}{\alpha} \sinh \left(\frac{y-\eta}{2} \right) \right]}, \quad y \in \mathfrak{R} \quad (4.8)$$

e

$$f_Y(y) = \frac{\alpha^{-1}(1 - \theta) \cosh \left(\frac{y-\eta}{2} \right) \phi \left[\frac{2}{\alpha} \sinh \left(\frac{y-\eta}{2} \right) \right]}{\left(1 - \theta \Phi \left[-\frac{2}{\alpha} \sinh \left(\frac{y-\eta}{2} \right) \right] \right)^2}, \quad (4.9)$$

com $-\infty < y < \infty$, $0 < \theta < 1$, $\alpha > 0$ e $-\infty < \eta < \infty$ e $\phi(\cdot)$ é a função de distribuição de probabilidade normal padrão, na Figura 4.4(a) apresentamos o gráfico de densidade da distribuição LGBS, observamos que a medida que o parâmetro α aumenta, a curtose diminui. Na Figura 4.4(b) podemos observar que o parâmetro θ está relacionado com a forma da distribuição e a medida que ela diminui ela tende para a distribuição log BS (LBS).

Considere o seguinte modelo de regressão

$$Y_i = \mu(x_i, \beta) + b_i \quad i = 1, 2, \dots, n, \quad (4.10)$$

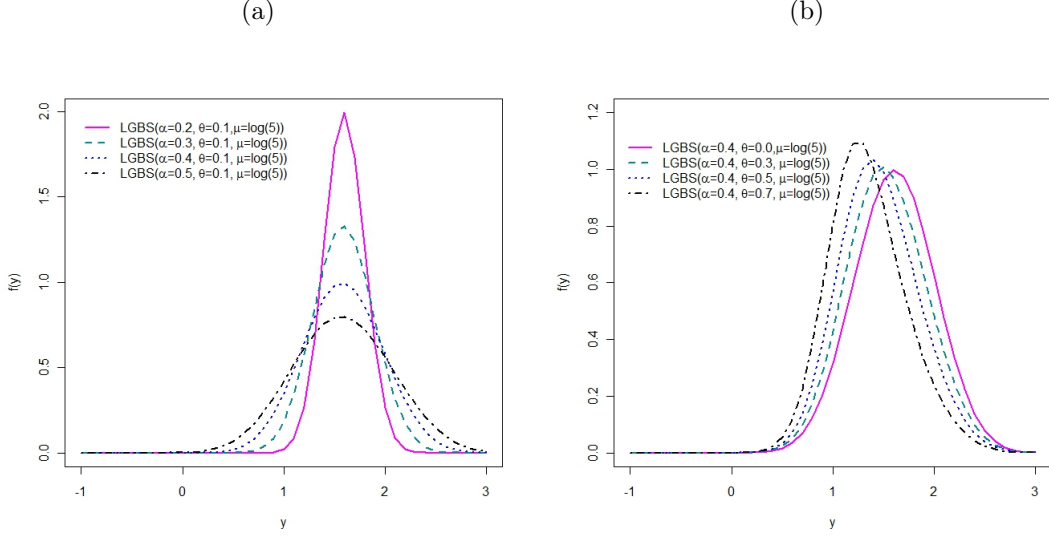


Figura 4.4: Densidade da distribuição LGBS.

em que y_i é o log-tempo de sobrevivência, $\beta=(\beta_1, \dots, \beta_p)^T$ é um vetor de parâmetros desconhecidos que serão estimados, x_i contém os valores da variável exploratória, $\mu(\cdot)$ é uma função linear ou não linear com respeito aos parâmetros, b_i são os erros com f.d.p. dadas por.

$$f_B(b_i) = \frac{\alpha^{-1}(1-\theta) \cosh\left(\frac{b_i}{2}\right) \phi\left[\frac{2}{\alpha} \sinh\left(\frac{b_i}{2}\right)\right]}{\left(1 - \theta \Phi\left[-\frac{2}{\alpha} \sinh\left(\frac{b_i}{2}\right)\right]\right)^2}, \quad u \in \mathfrak{R}. \quad (4.11)$$

Os momentos de B podem ser calculados a partir de,

$$E[B^r] = \int_{-\infty}^{\infty} b^r \frac{\alpha^{-1}(1-\theta) \cosh\left(\frac{b}{2}\right) \phi\left[\frac{2}{\alpha} \sinh\left(\frac{b}{2}\right)\right]}{\left(1 - \theta \Phi\left[-\frac{2}{\alpha} \sinh\left(\frac{b}{2}\right)\right]\right)^2} db. \quad (4.12)$$

O cálculo da equação (4.12), não é analiticamente tratável, portanto métodos numéricos ou alguns outros métodos alternativos poderiam ser usados. Na Tabela 4.1 apresentamos alguns valores para o valor esperado dos erros para diferentes valores de α e θ , onde observamos que quando $\theta=0$, $E(b) = 0$ para qualquer valor de α .

Tabela 4.1: Valores Esperados dos erros do modelo de regressão LGBS.

$\theta \backslash \alpha$	0,1	0,3	0,5	0,7	0,9	1
0	0	0	0	0	0	0
0,2	0	-0,03	-0,06	-0,08	-0,11	0,12
0,4	0	-0,09	-0,14	0,19	-0,24	-0,27
0,6	0	0,15	-0,25	-0,34	-0,43	-0,47
0,8	0	-0,26	-0,43	-0,59	-0,74	-0,81

4.3 Inferência Bayesiana para o modelo LGBS

Considere um conjunto de n observações $(x_1, y_1), \dots, (x_n, y_n)$ com $x_i = (x_{i,1}, \dots, x_{i,p})^\top$, e $b_i \sim \log \text{GBS}(\alpha, \theta, 0)$ onde a distribuição de probabilidade associada a b_i é dada por (4.11), a função de verossimilhança para α, β, θ é dada por,

$$L(\theta, \alpha, \beta) = \alpha^{-n}(1 - \theta)^n \prod_{i=1}^n \frac{\cosh\left(\frac{y_i - \mu(x_i, \beta)}{2}\right) \phi\left[\frac{2}{\alpha} \sinh\left(\frac{y_i - \mu(x_i, \beta)}{2}\right)\right]}{\left(1 - \theta \Phi\left[-\frac{2}{\alpha} \sinh\left(\frac{y_i - \mu(x_i, \beta)}{2}\right)\right]\right)^2}. \quad (4.13)$$

Distribuições a priori e a posteriori

Com respeito as *priori*, atribuímos uma *priori* própria, portanto a *posteriori* esta bem definida e é própria, consideramos que os parâmetros são independentes, com densidade dada por,

$$\pi(\alpha, \beta, \theta) = \pi(\theta)\pi(\alpha) \prod_{j=1}^p \pi(\beta_j), \quad (4.14)$$

em que as *priori* para os β_j , vão depender do espaço paramétrico onde ela esta definida em qualquer dos casos(função linear ou não linear). A *priori* para o parâmetro $\theta \sim \text{beta}(a, b)$ com $a < b$ e a *priori* para o parâmetro $\alpha \sim G(u, v)$ com $u \ll v$, assim a distribuição *a posteriori* conjunta para os parâmetros α, β, θ é dada por

$$\begin{aligned} \pi(\theta, \alpha, \beta | \mathbf{x}, \mathbf{y}) &= \alpha^{-n}(1 - \theta)^n \prod_{i=1}^n \frac{\cosh\left(\frac{y_i - \mu(x_i, \beta)}{2}\right) \exp\left[-\frac{2}{\alpha^2} \sinh^2\left(\frac{y_i - \mu(x_i, \beta)}{2}\right)\right]}{\left(1 - \theta \Phi\left[-\frac{2}{\alpha} \sinh\left(\frac{y_i - \mu(x_i, \beta)}{2}\right)\right]\right)^2} \\ &\times \pi(\alpha, \beta, \theta). \end{aligned} \quad (4.15)$$

A partir desta equação as marginais não podem ser obtidas de forma analítica, portanto para aplicar métodos MCMC, precisamos das condicionais *a posteriori*. Assim a condicional *a posteriori* para os parâmetros são dadas por,

$$\pi(\theta|\mathbf{x}, \mathbf{y}, \alpha, \boldsymbol{\beta}) \propto (1 - \theta)^n \prod_{i=1}^n \left(1 - \theta \Phi \left[-\frac{2}{\alpha} \sinh \left(\frac{y_i - \mu(x_i, \boldsymbol{\beta})}{2} \right) \right] \right)^{-2} \pi(\theta), \quad (4.16)$$

$$\pi(\boldsymbol{\beta}|\mathbf{x}, \mathbf{y}, \alpha, \theta) \propto \prod_{i=1}^n \frac{\cosh \left(\frac{y_i - \mu(x_i, \boldsymbol{\beta})}{2} \right) \exp \left[-\frac{2}{\alpha^2} \sinh^2 \left(\frac{y_i - \mu(x_i, \boldsymbol{\beta})}{2} \right) \right]}{\left(1 - \theta \Phi \left[-\frac{2}{\alpha} \sinh \left(\frac{y_i - \mu(x_i, \boldsymbol{\beta})}{2} \right) \right] \right)^2} \prod_{i=1}^n \pi(\beta_j), \quad (4.17)$$

$$\pi(\alpha|\mathbf{x}, \mathbf{y}, \boldsymbol{\beta}, \theta) \propto \prod_{i=1}^n \frac{\exp \left[-\frac{2}{\alpha^2} \sinh^2 \left(\frac{y_i - \mu(x_i, \boldsymbol{\beta})}{2} \right) \right]}{\left(1 - \theta \Phi \left[-\frac{2}{\alpha} \sinh \left(\frac{y_i - \mu(x_i, \boldsymbol{\beta})}{2} \right) \right] \right)^2} \times \pi(\alpha). \quad (4.18)$$

Como nenhuma das expressões têm forma fechada, usaremos o algoritmo de metrópolis Hastings.

4.4 Aplicação linear

Neste caso consideramos $\mu(\cdot)$ uma função linear, para a aplicação foram considerados os dados de teste de vida acelerado apresentado por Rieck & Nedelman (1991) (Tabela 3.5) e o modelo de regressão dada por

$$\log N_i = \beta_1 + \beta_2 \cdot \log W_i + b_i, \quad i = 1, 2, \dots, 46,$$

em que a f.d.p. para b_i é dada por (4.11). Consideramos o problema desde uma perspectiva Bayesiana. Inicialmente foi ajustado o modelo LBS, as distribuições *a priori* usadas foram $\lambda \sim (1.2, 0.0001)$, $\beta_1 \sim N(0, 12)$, $\beta_2 \sim N(0, 12)$, foi perturbada uma observação (observação 8) com 3 DP.

Foi gerada uma cadeia com 100000 iterações cada, as 5000 primeiras foram descartadas para reduzir a possível influência do ponto inicial da cadeia. As restantes foram selecionadas de 5 em 5 para evitar a correlação nas séries, resultando numa amostra com 19000 observações.

Para o modelo LGBS *a priori* para $\theta \sim \text{beta}(2, 6)$ e para os parâmetros α e β , consideramos as mesmas *priori* que o modelo LBS. As amostras obtidas para distribuição *a posteriori* dos parâmetros foram semelhantes ao caso LBS.

Tabela 4.2: Resumo *a posteriori* LGBS modelo linear.

	<i>LBS</i> (SP)		<i>LGBS</i> (SP)		<i>LBS</i> (CP)		<i>LGBS</i> (CP)	
	média	DP	média	DP	média	DP	média	DP
β_1	12,1	0,4	12,1	0,4	13,4	0,8	13,1	0,6
β_2	-1,6	0,1	-1,6	0,1	-1,9	0,2	-1,6	0,2
α	0,4	0,04	0,4	0,05	0,9	0,1	1,0	0,2
θ			0,2	0,1			0,8	0,1
B	-26,43		-26,78		-69,83		-50,93	
DIC	52,73		50,67		116,74		85,45	

Tabela 4.3: Medida de de divergência K-L.

Caso	<i>LBS</i>		<i>LGBS</i>	
	K-L	cal	K-L	cal
8	12,41	1,00	7,3	1,00

Na Tabela 4.2 podemos observar que as estimativas do modelo LGBS se comportam melhor que o modelo LBS, quando há presença de observações influentes, podemos notar que o ajuste do parâmetro β_2 tem uma diferença com respeito ao ajuste do modelo LGBS; além disso, temos que os valores B e DIC indicam que o modelo LGBS ajusta melhor ao comportamento dos dados. Com respeito as medidas de divergência de Kullback Leibler, o modelo LGBS reduz essas medidas estimadas mediante o modelo LBS (Tabela 4.3).

Um ponto importante no ajuste deste modelo, é que a média dos erros do modelo (4.10) não é zero, portanto o valor estimado de β_1 , pode ser modificado, encontrando $E(b)$ da equação (4.12), neste caso o valor esperado $E(b) = -0,8$, portanto o valor de $\beta_1 = 12,3$.

Aplicação não linear

Para o conjunto de dados apresentados por Rieck & Nedelman (1991) sobre a vida de uma peça metálica (Tabela 3.5), consideramos o modelo (3.23), com f.d.p. para os erros (b_i) dada por (4.11). Consideramos as *priori* próprias e independentes, e o cálculo das condicionais a posteriori são dadas nas equações (4.16), (4.18) e (4.17); perturbamos uma observação (observação 8) com 3 DP. Inicialmente foi ajustado o modelo log BS, as distribuições *a priori* usadas para $\alpha \sim G(1, 0.0001)$, $\beta_1 \sim N(0, 100)$, $\beta_2 \sim N(0, 100)$, $\beta_3 \sim N(0, 100)$. As amostras obtidas para distribuição *a posteriori* dos parâmetros foram semelhantes ao caso LGBS linear.

Para o modelo LGBS consideramos *a priori* para $\theta \sim \text{beta}(2, 6)$ e a *a priori* para α e β , foram semelhantes ao caso LBS (não linear). Os resultados obtidos são apresentados na seguinte Tabela.

Tabela 4.4: Resumo *a posteriori* LGBS não linear.

	<i>LBS</i> (SP)		<i>LGBS</i> (SP)		<i>LBS</i> (CP)		<i>LGBS</i> (CP)	
	média	DP	média	DP	média	DP	média	DP
β_1	9,4	1,0	9,5	1,1	9,6	1,2	10,7	1,4
β_2	-5,5	0,8	-5,6	0,9	-6,5	1,0	-5,7	1,1
β_3	-21,2	6,9	-20,8	6,7	-28,0	9,4	-22,6	9,5
α	0,41	0,04	0,43	0,05	0,8	0,1	1,0	0,3
θ			0,2	0,1			0,8	0,1
B	-26,28		-26,45		-66,53		-51,09	
DIC	46,76		49,28		113,16		83,87	

Tabela 4.5: Medida de de divergência K-L, não linear.

	<i>LBS</i>		<i>LGBS</i>	
Caso	K-L	cal	K-L	cal
8	9,61	1,00	7,94	1,00

Na Tabela 4.4 podemos observar que as estimativas dos parâmetros do modelo LGBS sem perturbação (SP) e do modelo LBS (SP) são próximas, mas, quando te-

mos observações influentes tem uma considerável diferença, sobretudo na estimativa do parâmetro β_3 . Na comparação de modelos os valores de B e DIC para o modelo LGBS indicam que se ajustam melhor aos dados se comparado com o modelo LBS na presença de observações influentes. Na Tabela 4.5 se apresenta a medida de divergência K-L, onde podemos notar que essa medida diminui quando temos o modelo LGBS.

Na estimação dos parâmetros para o modelo não linear, podemos corrigir a estimativa do parâmetro β_1 encontrando o valor esperado dos erros $E(b)$, neste caso o valor encontrado foi, $E(b) = -0,8$; portanto o valor corrigido para $\beta_1 = 9,9$.

Assim os valores estimados usando o modelo LGBS pode ser usada como uma alternativa ao modelo LBS na presença de observações influentes.

Capítulo 5

Considerações Finais

Neste estudo apresentamos procedimentos inferenciais, desde uma perspectiva Bayesiana para modelos de regressão considerando algumas extensões da distribuição Birnbaum-Saunders (BS). O procedimento Bayesiano é baseado em métodos de simulação via cadeias de Markov (MCMC). Os algoritmos relacionados ao procedimento MCMC são de fácil implementação computacional e os resultados podem ser usados de forma segura. Além disso, consideramos um estudo de diagnóstico baseado na medida de divergência de Kullback-Leibler para detectar os pontos influentes na distribuição *a posteriori* dos parâmetros do modelo de regressão.

A seguir apresentamos algumas propostas de continuidade do trabalho:

1. Desenvolvimento de metodologia Bayesiana para o modelo de regressão na presença de dados censurados considerando as extensões da distribuição BS consideradas neste estudo;
2. Considerar outras medidas de divergência (Peng & Dey, 1995) no estudo de diagnóstico;
3. Desenvolvimento de metodologia Bayesiana e de diagnóstico para modelos de sobrevivência com fração de cura considerando a distribuição BS para os tempos latentes e outras distribuições para o número de causas latentes;

Acreditamos que as técnicas desenvolvidas neste trabalho podem ser aplicadas a outros modelos, como modelos de regressão log-BSSMN com erros de medição nas variáveis explicativas e efeitos aleatórios.

Referências Bibliográficas

- Achcar, J. & Espinosa, M. (1992). Use of bayesian methods in accelerated life tests considering a log-linear model for the birnbaum–saunders distribution. *Biometrics Bulletin*, **10**, 18–22.
- Balakrishnan, N., Leiva, V., Sanhueza, A. & Vilca, F. (2009). Estimation in the Birnbaum–Saunders distribution based on scale-mixture of normals and the EM-algorithm. *SORT (Statistics and Operations Research Transactions)*, **33**(2), 171–192.
- Birnbaum, Z. W. & Saunders, S. C. (1969). A new family of life distributions. *Journal of Applied Probability*, **6**(2), 319–327.
- Cancho, V., Ortega, E. & Paula, G. (2010). On estimation and influence diagnostics for log-Birnbaum-Saunders Student-t regression models: Full Bayesian analysis. *Journal of Statistical Planning and Inference*, **140**(9), 2486–2496.
- Cancho, V. G., Louzada-Neto, F. & Barriga, G. D. C. (2011). The poisson-exponential lifetime distribution. *Comput. Stat. Data Anal.*, **55**, 677–686.
- Chib, S. & Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *American Statistician*, **49**(4), 327–335.
- Cho, H., Ibrahim, J., Sinha, D. & Zhu, H. (2009). Bayesian case influence diagnostics for survival models. *Biometrics*, **65**(1), 116–124.
- Cook, R. (1986). Assessment of local influence. *Journal of the Royal Statistical Society. Series B (Methodological)*, **48**, 133–169.
- Cook, R. & Weisberg, S. (1982). *Residuals and influence in regression*, volume 5. Chapman and Hall, New York.

- Cowles, M. & Carlin, B. (1996). Markov chain monte carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, **91**(434), 883–904.
- Cox, D. R. & Snell, E. J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society B*, **30**(2), 248–275.
- Desmond, A. (1985). Stochastic models of failure in random environments. *Canadian Journal of Statistics*, **13**(3), 171–183.
- Desmond, A. (1986). On the relationship between two fatigue-life models. *Reliability, IEEE Transactions on*, **35**(2), 167–169.
- Diaz-Garcia, J. & Dominguez-Molina, J. (2006). Some generalisations of Birnbaum-Saunders and sinh-normal distributions. **1**(35), 1709–1727.
- Diaz-Garcia, J. & Leiva-Sanchez, V. (2005). A new family of life distributions based on the elliptically contoured distributions. *Journal of Statistical Planning and Inference*, **128**(2), 445–457.
2007
- Farfán Carrasco, J., M. (2007). Modelo de regressão log-weibull modificado e a nova distribuição weibull modificada generalizada. Dissertação de mestrado, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba.
- Gelfand, A. & Dey, D. (1994). Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society. Series B (Methodological)*, **56**(3), 501–514.
- Gilks, W. R. & Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Applied Statistics*, **41**(2), 337–348.
- Gómez, H., Olivares-Pacheco, J. & Bolfarine, H. (2009). An extension of the generalized Birnbaum-Saunders distribution. *Statistics & Probability Letters*, **79**(3), 331–338.
- Kelker, D. (1970). Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhyā: The Indian Journal of Statistics, Series A*, **32**(4), 419–430.

- Lange, K. & Sinsheimer, J. (1993). Normal/independent distributions and their applications in robust regression. *Journal of Computational and Graphical Statistics*, **2**(2), 175–198.
- Lange, K., Little, R. & Taylor, J. (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistical Association*, **84**(408), 881–896.
- Leiva, V., Barros, M., Paula, G. A. & Sanhueza, A. (2008a). Generalized Birnbaum-Saunders distributions applied to air pollutant concentration. *Environmetrics*, **19**(3), 235–250.
- Leiva, V., Riquelme, M., Balakrishnan, N. & Sanhueza, A. (2008b). Lifetime analysis based on the generalized Birnbaum–Saunders distribution. *Computational Statistics and Data Analysis*, **52**(4), 2079–2097.
- Lemonte, A. & Cordeiro, G. (2009). Birnbaum–Saunders nonlinear regression models. *Computational Statistics and Data Analysis*, **53**(12), 4441–4452.
- Lucas, A. (1997). Robustness of the student t based m-estimator. *Communications in Statistics-Theory and Methods*, **26**(5), 1165–1182.
- McCulloch, R. (1989). Local model influence. *Journal of the American Statistical Association*, **84**(406), 473–478.
- Owen, W. & Padgett, W. (1999). Accelerated test models for system strength based on Birnbaum-Saunders distributions. *Lifetime Data Analysis*, **5**(2), 133–147.
- Peng, F. & Dey, D. (1995). Bayesian analysis of outlier problems using divergence measures. *Canadian Journal of Statistics*, **23**(2), 199–213.
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Rieck, J. & Nedelman, J. (1991). A log-linear model for the Birnbaum-Saunders distribution. *Technometrics*, **33**(1), 51–60.
- Rosa, G., Padovani, C. & Gianola, D. (2003). Robust linear mixed models with normal/independent distributions and bayesian mcmc implementation. *Biometrical Journal*, **45**(5), 573–590.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B*, **64**, 583–639.

Vilca-Labra, F. & Leiva, V. (2006). A new fatigue life model based on the family of skew-elliptical distributions. *Communications in Statistics – Theory and Methods*, **35**(2), 229–244.