

Universidade Federal de Campina Grande
Centro de Ciências e Tecnologia
Programa de Pós-Graduação em Matemática
Curso de Mestrado em Matemática

**Modelos de regressão
log-Birnbaum-Saunders
generalizados para dados com
censura intervalar**

por

Joelson da Cruz Campos [†]

sob orientação da

Prof^{fa}. Dr^a. Michelli Karinne Barros da Silva

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática - CCT - UFCG, como requisito parcial para obtenção do título de Mestre em Matemática.

[†]Este trabalho contou com apoio financeiro da CAPES

Modelos de regressão log-Birnbaum-Saunders generalizados para dados com censura intervalar

por

Joelson da Cruz Campos

Dissertação apresentada ao Corpo Docente do Programa de Pós-Graduação em Matemática - CCT - UFCG, como requisito parcial para obtenção do título de Mestre em Matemática.

Área de Concentração: Estatística

Aprovada por:

Prof^a. Dr^a. Tatiane Ferreira do Nascimento Melo da Silva - UFCG

Prof. Dr. Gilberto da Silva Matos - UFCG

Prof^a. Dr^a. Michelli Karinne Barros da Silva - UFCG

Orientadora

Universidade Federal de Campina Grande
Centro de Ciências e Tecnologia
Programa de Pós-Graduação em Matemática
Curso de Mestrado em Matemática

Dezembro/2011

Resumo

Neste trabalho, propomos o modelo de regressão log-Birnbaum-Saunders generalizado para analisar dados com censura intervalar. As funções escore e a matriz de informação de Fisher observada foram obtidas, bem como foi discutido o processo de estimação dos parâmetros do modelo. Como medida de influência, consideramos o afastamento pela verossimilhança (likelihood displacement) sob vários esquemas de perturbação. Derivamos as matrizes apropriadas para obter a influência local nos parâmetros estimados do modelo e realizamos uma análise residual baseada nos resíduos de Cox-Snell ajustado, Martingale e componente do desvio modificado. Apresentamos também um estudo de simulação de Monte Carlo a fim de investigar o comportamento da distribuição empírica dos resíduos propostos. Finalmente, uma aplicação com dados reais é apresentada.

Palavras-chave: Distribuição Birnbaum-Saunders generalizada; censura intervalar; modelo de regressão; análise de resíduos; influência local.

Abstract

In this work, we propose the model of regression log-Birnbaum-Saunders generalized to analyze data with interval censored. The score functions and the observed Fisher information matrix were obtained, as well as the process for estimating of the parameters was discussed. As measure of influence, we considered the likelihood displacement under several schemes of perturbation. The normal curvatures of local influence were derived and we conducted a residual analysis based on residuals Cox-Snell adjusted, Martingale and modified deviance. A Monte Carlo simulation was carried in order to investigate the behavior of empiric distribution of the proposed residuals. Finally, an application with real data is presented.

Keywords: generalized Birnbaum-Saunders distribution; interval censored; regression model; residuals analysis; local influence.

Agradecimentos

- Primeiramente agradeço a Deus, pois sem Ele jamais teria chegado até aqui.
- A minha Família por todo o incentivo que tive para fazer o mestrado.
- A Tatá por todo seu amor, apoio e compreensão.
- Aos grandes amigos Elvis (pelo apoio nos momentos difíceis no começo da graduação e pelos finais de semana de estudo), Ceiza e Erikarla por toda amizade, confiança e apoio.
- Aos amigos de lar (Flamingo), Ailton (por todo o tempo que estudamos juntos na graduação), Rodrigo, Romildo (pelos estudos de Álgebra linear) e Tonhaunm.
- Aos professores que aceitaram participar da banca para avaliação desse trabalho e pelas suas sugestões para aperfeiçoar o mesmo.
- Aos professores da UFCG que contribuíram para minha formação de maneira direta ou indireta.
- A professora Michelli pela orientação “maternal”, nunca conheci um(a) professor(a) tão preocupado com seus alunos, sem dúvida alguma um exemplo a ser seguido. Não há como descrever em palavras toda gratidão que tenho por todo o conhecimento compartilhado, pela amizade, disponibilidade, paciência e orientação desse trabalho, mas deixo o meu mais sincero muito obrigado!
- Aos professores da UFRN que me deram o conhecimento necessário para que eu pudesse vir à UFCG. Em especial agradeço aos professores Ronaldo (por todo aprendizado que obtive com ele e pelo tempo que trabalhamos juntos), Viviane (por ter incentivado a mim e a Tatá a fazer o mestrado, abrindo nossos olhos nos momentos de dúvida), aos professores Jaques e Gabriela que foram como pais acadêmicos no decorrer da minha graduação e aos professores/amigos Rubens e Márcia.

- Aos professores/amigos Andréa Gusmão, Fátima Feitosa e Fernando Fernandes (In memoriam) com os quais aprendi muito.
- Aos amigos que fiz no decorrer da minha longa caminhada acadêmica em especial a turma de matemática da UFRN de 2006 e a galera do PET de matemática da UFRN.
- As grandes amigas de mestrado e colegas de estudo Fabiana (pessoa mais calma desse mundo) e Aline (aperriada que nem Eu) pelos momentos de estudo e lazer, sem dúvidas vocês foram fundamentais nessa trajetória. Espero levar essa grande amizade que temos por toda a vida.
- A Bárbara (minha irmã por transitividade) pelo help!.
- A CAPES pelo apoio financeiro.

Dedicatória

A Tatá (meu bem) e a minha
família.

Sumário

Lista de Figuras	x
Lista de Tabelas	xiii
Introdução	1
1 Preliminares	4
1.1 Dados censurados	4
1.2 Funções importantes na análise de sobrevivência	6
1.2.1 Função de sobrevivência	6
1.2.2 Função de taxa de falha	8
1.2.3 Função de taxa de falha acumulada	10
1.3 Modelos paramétricos	13
1.4 O método da máxima verossimilhança	14
2 A Distribuição Birnbaum-Saunders e a Distribuição Birnbaum-Saunders Generalizada	17
2.1 Origem e propriedades da distribuição Birnbaum-Saunders	17
2.2 Relação entre as distribuições Birnbaum-Saunders e Senh-normal	24
2.3 A distribuição Birnbaum-Saunders generalizada	26
2.3.1 Definição e algumas propriedades da distribuição BSG	27
2.3.2 A distribuição log-Birnbaum-Saunders generalizada	31
3 O Modelo de regressão log-Birnbaum-Saunders generalizado	32

3.1	A construção dos modelos de regressão log-BS e log-BSG para dados com censura intervalar	32
3.2	Validação do modelo	35
3.2.1	Análise de resíduos	36
3.2.2	Estudo de simulação	39
3.2.3	Influência local	54
4	Aplicação	66
4.1	Descrição dos dados	66
4.2	Modelo log-BS	68
4.3	Modelo log-BS t de student	72
5	Conclusões	78
A	Cálculo das Segundas derivadas	79
A.1	Cálculo de $\frac{\partial^2 l(\theta)}{\partial \alpha^2}$	80
A.2	Cálculo de $\frac{\partial^2 l(\theta)}{\partial \beta_k \partial \alpha}$	81
A.3	Cálculo de $\frac{\partial^2 l(\theta)}{\partial \beta_k \partial \beta_j}$	82
A.4	A matriz hessiana	83
	Bibliografia	84

Lista de Figuras

1.1	Gráfico da função de sobrevivência da distribuição Weibull para alguns valores dos parâmetros (γ, α)	7
1.2	Gráfico da função de taxa de falha instantânea da distribuição Weibull para alguns valores dos parâmetros (γ, α)	11
2.1	Funções de risco da distribuição Birnbaum-Saunders para os valores de α indicados e $\beta = 1$	21
2.2	Gráfico da função densidade da distribuição Birnbaum-Saunders para os valores de $\alpha = 0, 1; 0, 5; 1, 0; 1, 5; 2, 0$ e β fixo igual 1.	21
2.3	Gráfico da função densidade da distribuição Birnbaum-Saunders para os valores de $\beta = 0, 1; 0, 5; 1, 0; 1, 5; 2, 0$ e α fixo igual 1.	22
2.4	Gráfico da densidade normal padrão bivariada e suas curvas de contorno vista sobre o plano x_1 e x_2	27
2.5	Densidade a distribuição BS t de student para $\nu = 3, 5, 7, 30$	29
2.6	Função de risco para a distribuição BS t de student para $\nu = 3, 5, 7, 30$	29
3.1	Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado (r_{C_i}), resíduo Martingale (r_{M_1}) e resíduo componente do desvio modificado (r_{D_i}) com porcentagem de censura à direita de 0% e $\alpha = 1$	41
3.2	Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado (r_{C_i}), resíduo Martingale (r_{M_1}) e resíduo componente do desvio modificado (r_{D_i}) com porcentagem de censura à direita de 10% e $\alpha = 1$	42
3.3	Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado (r_{C_i}), resíduo Martingale (r_{M_1}) e resíduo componente do desvio modificado (r_{D_i}) com porcentagem de censura à direita de 30% e $\alpha = 1$	43

3.4	Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado (r_{C_i}), resíduo Martingale (r_{M_1}) e resíduo componente do desvio modificado (r_{D_i}) com porcentagem de censura à direita de 0% e $\alpha = 2$	44
3.5	Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado (r_{C_i}), resíduo Martingale (r_{M_1}) e resíduo componente do desvio modificado (r_{D_i}) com porcentagem de censura à direita de 10% e $\alpha = 2$	45
3.6	Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado (r_{C_i}), resíduo Martingale (r_{M_1}) e resíduo componente do desvio modificado (r_{D_i}) com porcentagem de censura à direita de 30% e $\alpha = 2$	46
3.7	Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado, Martingale e componente de desvio modificado com $\nu = 3, 4$ e 5 nas três primeiras linhas, respectivamente e 0% de censura à direita sendo $\alpha = 1$	48
3.8	Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado, Martingale e componente de desvio modificado com $\nu = 7, 30$ e 100 nas três primeiras linhas, respectivamente e 0% de censura à direita sendo $\alpha = 1$	49
3.9	Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado, Martingale e componente de desvio modificado com $\nu = 3, 4$ e 5 nas três primeiras linhas, respectivamente e 10% de censura à direita sendo $\alpha = 1$	50
3.10	Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado, Martingale e componente de desvio modificado com $\nu = 7, 30$ e 100 nas três primeiras linhas, respectivamente e 10% de censura à direita sendo $\alpha = 1$	51
3.11	Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado, Martingale e componente de desvio modificado com $\nu = 3, 4$ e 5 nas três primeiras linhas, respectivamente e 30% de censura à direita sendo $\alpha = 1$	52
3.12	Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado, Martingale e componente de desvio modificado com $\nu = 7, 30$ e 100 nas três primeiras linhas, respectivamente e 30% de censura à direita sendo $\alpha = 1$	53
4.1	Sobrevivência estimada para o Modelo (4.1) via o algoritmo de Turnbull	68

4.2	Função de sobrevivência empírica e ajustada para o modelo BS.	69
4.3	Gráfico do resíduo componente de desvio versus índices das observações para o modelo log-BS.	70
4.4	Gráfico de envelope dos resíduos componentes do desvio modificado para o modelo de regressão log-BS.	71
4.5	Gráficos sob o esquema de perturbação do limite superior: (a) $C_i(\theta)$ versus índices; (b) l_{max} versus índices; (c) $C_i(\alpha)$ versus índices; (d) $C_i(\beta)$ versus índices	72
4.6	Função de sobrevivência empírica e ajustada para o modelo BSG.	74
4.7	Gráfico dos resíduos componentes do desvio modificado contra os índices das observações para o modelo de regressão log-BSG.	75
4.8	Gráfico de envelope dos resíduos componente do desvio modificado para o modelo de regressão log-BSG.	75
4.9	Gráficos sob o esquema de perturbação do limite superior: (a) $C_i(\theta)$ versus índices; (b) l_{max} versus índices; (c) $C_i(\alpha)$ versus índices; (d) $C_i(\beta)$ versus índices.	76

Lista de Tabelas

4.1	Tempos (em meses) até a retração de uma das mamas para pacientes com câncer de mama de acordo com dois tratamentos	67
4.2	Estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-BS.	69
4.3	Estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-BS para os dados de câncer de mama menos a observação 3.	73
4.4	Estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-BS para os dados de câncer de mama menos as observações 1, 2, 3 e 48.	73
4.5	Estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-BS-t com 3 graus de liberdade para os dados de câncer de mama.	73
4.6	Estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-BS-t com 3 graus de liberdade para os dados de câncer de mama menos a observação 3. Nesta tabela I.C denota intervalo de confiança.	77
4.7	Estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-BS-t com 3 graus de liberdade para os dados de câncer de mama menos as observações 1, 2, 3 e 48. Nesta tabela I.C denota intervalo de confiança.	77

Introdução

O modelo normal é o modelo de referência na Estatística Teórica e Aplicada. No entanto, sabemos que muitos fenômenos nem sempre se adequam a um modelo normal devido à falta de simetria. Nos dados de sobrevivência, por exemplo, a variável tempo de vida possui suporte nos reais positivos, apresentando uma distribuição assimétrica positiva o que não seria razoável modelá-la por uma distribuição simétrica.

De acordo com Colosimo e Giolo (2006), em análise de sobrevivência, a variável resposta é, geralmente, o tempo até a ocorrência de um evento de interesse. Este tempo é denominado tempo de falha, podendo ser o tempo até a morte do paciente, bem como até a cura ou recidiva de uma doença. No entanto, em engenharia, o tempo não é medido necessariamente desta forma. Por exemplo, o envelhecimento de itens (componentes, sistemas, subsistemas, peças, estruturas, órgãos, unidades, etc) nem sempre é medido pelo tempo cronológico. Em algumas ocasiões, o tempo de vida é medido por meio de outras variáveis como a quantidade de quilômetros percorrido, a força aplicada a um material até sua ruptura, o nível de degradação, número de ciclos até que um material seja desgastado (fadiga), etc. Além disso, a terminologia variável tempo de vida (*lifetime variate*) é amplamente utilizada para se referir a qualquer variável aleatória contínua positiva (por exemplo, a quantidade de água da chuva), ver Leiva et al. (2009). O modelo probabilístico associado a uma variável tempo de vida é habitualmente chamado de distribuição de vida. Para mais detalhes sobre distribuições de vida, ver Marshall e Olkin (2007).

Os dados de sobrevivência são caracterizados pela presença de censura que é a observação parcial da resposta e na presença desta, a variável resposta é constituída de dois componentes: tempos de falha e tempos censurados.

O interesse na análise de dados de sobrevivência é estimar a função de sobre-

vivência, que na presença de censura não se torna viável aplicar as técnicas convencionais de análise de dados, pois não podemos descartar as observações censuradas, uma vez que fornecem informações ao nosso estudo e, se não consideradas, levam a conclusões viciadas. Existem duas abordagens estatísticas para estimar a função de sobrevivência, uma delas considera modelos paramétricos e outra modelos não paramétricos. Muitas vezes o uso das duas abordagens é indicado para garantir resultados confiáveis.

Uma importante distribuição de vida (assimétrica) originada de um problema de fadiga de materiais foi desenvolvida por Birnbaum e Saunders (1969a). Esta distribuição é chamada de distribuição Birnbaum-Saunders (BS) e relaciona o tempo até a ocorrência de alguma falha no material com algum dano cumulativo assumido gaussiano. A distribuição BS vem recebendo ultimamente um grande destaque na literatura, tendo sido aplicada na área de engenharia, área médica, de ciências atuariais e de meio ambiente, em que algum tipo de estresse cumulativo pode levar a ocorrência de um evento, ver Leiva, et al. (2007, 2009) e Podlaski (2008).

Díaz-García e Leiva (2005, 2006) apresentam uma generalização da distribuição BS, obtendo-a a partir de distribuições de contornos elípticos, a qual denominaram distribuição Birnbaum-Saunders generalizada (BSG). Segundo os autores, esta generalização baseia-se, além do aspecto elípticos, na busca de distribuições de vida que cresçam mais rapidamente e que possuam caudas com menor ou maior curtose do que a distribuição BS, entre outras propriedades tais como: distribuições de vida que não possuem momentos e que sejam bimodais.

Em Barros et al. (2008) foi discutido o uso da distribuição Birnbaum-Saunders generalizada para analisar dados da área médica, como por exemplo, dados provenientes de estudos de doenças cardíacas ou diferentes tipos de câncer, em que um tipo de dano acumulado provocado por vários fatores de risco é detectado. Então, essa degradação leva a um processo de fadiga, cuja propagação do tempo de sobrevivência pode ser adequadamente modelada pela distribuição BS ou BSG (em particular, BS- t -Student). Em estudos dessa natureza, muitas vezes, o tempo até a ocorrência do evento de interesse não é observado, mas é conhecido o intervalo de tempo onde ocorreu o evento de interesse. Neste caso, estamos na presença de censura intervalar. Neste trabalho, propomos o modelo de regressão log-Birnbaum-Saunders generalizado para analisar dados com censura intervalar. No Capítulo 1 fazemos uma breve revisão dos principais

conceitos de Análise de Sobrevivência. No Capítulo 2 apresentamos a origem da distribuição BS bem como suas principais propriedades, posteriormente apresentamos a distribuição senh-normal e relacionamos a mesma com a distribuição BS. Em seguida, apresentamos a distribuição BS generalizada enunciando suas principais propriedades e concluimos o capítulo apresentando a distribuição senh-esférica e a relação entre a mesma e a distribuição BS generalizada, relação essa que será de grande valia para construção do modelo de regressão log-Birnbaum-Saunders generalizado. No Capítulo 3 construímos o modelo de regressão log-Birnbaum-Saunders generalizado para dados com censura intervalar. Em seguida, apresentamos os resíduos de Cox-Snell ajustado, Martingale e Componente do desvio modificado e fazemos um estudo de simulação no intuito de analisar a distribuição empírica dos mesmos. Por fim, fazemos uma análise de diagnóstico baseado em influência local. No Capítulo 4 aplicamos a teoria estudada a um banco de dados reais. No Capítulo 5 apresentamos as principais conclusões.

Capítulo 1

Preliminares

No decorrer desse capítulo faremos uma breve revisão dos principais conceitos de Análise de Sobrevivência os quais serão de fundamental importância para os estudos que realizaremos nesse trabalho.

1.1 Dados censurados

A variável resposta em estudos de análise de sobrevivência, geralmente, corresponde ao tempo de ocorrência de um evento de interesse. Segundo Colosimo e Giolo (2006) denominamos este tempo de **tempo de falha**, que pode corresponder, por exemplo, ao tempo até a morte de um paciente, ou o tempo até a cura de um paciente dependendo do interesse da situação em estudo.

Os dados de sobrevivência são caracterizados pela presença de **censura** que é a observação parcial da resposta. Essa observação parcial deve-se ao fato de que, por algum motivo, o acompanhamento do paciente foi interrompido, seja porque ele mudou de cidade, ou o estudo foi concluído para análise dos dados, ou o paciente morreu de um motivo diferente do estudado.

Alguns mecanismos de censura são diferenciados em estudos clínicos. Censura do tipo I ocorrem nos estudos que ao serem concluídos por meio de um tempo pré-estabelecido, registram alguns indivíduos que ainda não apresentaram a situação de interesse. Censura do tipo II resultam de estudos que são finalizados após a ocorrência do evento de interesse em número de pré-estabelecido de pacientes. O terceiro mecanismo de censura trata-se da censura aleatória que surge, por exemplo, na área médica

quando um paciente sai do estudo sem que tenha ocorrido a falha por motivos que diferem do evento de interesse.

É importante ressaltar que na engenharia, ocorre um estudo similar ao da análise de sobrevivência, pois comumente há necessidade de se estudar o tempo de vida de um determinado material no intuito de, por exemplo, saber a sua durabilidade de um objeto para que se possa determinar o prazo a ser dado como garantia do produto. Tal estudo quando feito na engenharia recebe o nome de confiabilidade. Para maiores detalhes ver Mann et al. (1974), Meeker e Escobar (1998) e Saunders (2007).

No decorrer de um estudo podem ocorrer ainda censura à direita, censura à esquerda, ou ambas simultaneamente em que temos à censura intervalar. Dizemos que ocorre censura à direita quando o tempo de ocorrência do evento de interesse está à direita do tempo registrado. A censura à esquerda ocorre quando o tempo registrado é maior que o tempo de falha. E a censura intervalar ocorre em estudos com acompanhamentos periódicos em que é conhecido apenas que o evento de interesse ocorreu em um certo intervalo de tempo. Devido o tempo de falha T não ser conhecido exatamente, mas sim pertencer a um intervalo, ou seja, $T \in (L, U)$, tais dados são denominados por dados de sobrevivência intervalar, ou mais usualmente, por dados de censura intervalar. Temos à seguir um exemplo que ilustra o exposto nesse parágrafo.

Considere, por exemplo, um grupo de 20 pacientes submetidos a presença de um certo vírus e suponha que os mesmos irão contraí-lo, porém os vírus ficarão incubados nos pacientes e se manifestarão em um determinado momento desconhecido. Com o intuito de prever o tempo de incubação do vírus os pacientes são testados a cada semana no decorrer de 6 semanas. Denotando por T_i o tempo de incubação do vírus no i -ésimo indivíduo, temos que $T_i \in (U_i, V_i]$, e considerando como unidade a semana, temos que $U_i, V_i \in \{0, 1, 2, 3, 4, 5, 6\}$. Temos, por exemplo, que se $T_2 \in (2, 3]$ significa que a doença se manifestou no segundo paciente no decorrer da terceira semana e que não se sabe ao certo o dia, ou seja, temos uma censura intervalar. Notemos que neste conjunto de dados podem ocorrer os seguintes casos particulares: $T_i \in (0, 1]$, ou seja, antes da primeira análise a doença se manifestou, dizemos nesse caso que o indivíduo sofreu censura à esquerda; $T_i \geq 6$ ou $T_i \in [6, \infty)$, nesse caso dizemos que o indivíduo sofreu censura à direita; o indivíduo contraiu a doença no dia da avaliação, onde nesse caso dizemos que o indivíduo teve um tempo de falha exato.

1.2 Funções importantes na análise de sobrevivência

Em estudos de sobrevivência, em geral, consideramos a variável aleatória T que representa o tempo de falha como uma variável aleatória não-negativa e contínua. Tal variável aleatória é especificada pela função de sobrevivência ou pela função de taxa de falha. No que segue apresentaremos estas funções e algumas outras funções úteis no estudo de sobrevivência.

1.2.1 Função de sobrevivência

De acordo com Kalbfleisch e Prentice (2002) temos a definição a seguir.

Definição 1.1. *A probabilidade de uma observação não falhar, ou seja, sobreviver até o tempo t , denominada de função de sobrevivência é dada por*

$$S(t) = P(T > t) = 1 - F(t),$$

em que $F(t)$ representa a função de distribuição acumulada (fda) da variável aleatória T .

Esta é uma das principais funções probabilísticas usada para descrever estudos de sobrevivência.

Exemplo 1.1. *Considere T uma variável aleatória Weibull com parâmetros γ e α cuja função de densidade é dada por*

$$f(t) = \frac{\gamma}{\alpha^\gamma} t^{\gamma-1} \exp \left\{ - \left(\frac{t}{\alpha} \right)^\gamma \right\}, \quad t \geq 0$$

em que γ e α são parâmetros positivos e representam, respectivamente, os parâmetros de forma e escala. A função de sobrevivência de T é dada por

$$S(t) = \exp \left\{ - \left(\frac{t}{\alpha} \right)^\gamma \right\}.$$

Observando a Figura (1.1) vemos que a função de sobrevivência, $S(t)$, da distribuição Weibull é não-crescente, contínua e satisfaz $S(0) = 1$ e $\lim_{t \rightarrow \infty} S(t) = 0$. Na Propriedade (1.1), veremos que isto não é válido apenas para a distribuição Weibull e sim para qualquer distribuição contínua.

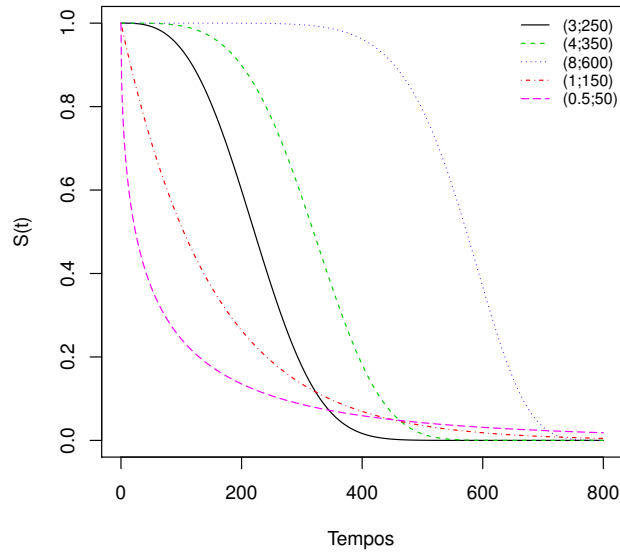


Figura 1.1: Gráfico da função de sobrevivência da distribuição Weibull para alguns valores dos parâmetros (γ, α) .

Propriedade 1.1. *Se T é uma variável aleatória absolutamente contínua e não-negativa, então sua função de sobrevivência S possui as seguintes propriedades:*

- (i) S é monótona não-crescente.
- (ii) $S(0) = 1$ e $\lim_{t \rightarrow \infty} S(t) = 0$.
- (iii) S é contínua.

Demonstração. Para demonstrarmos essas propriedades faremos uso das propriedades da função de distribuição acumulada (fda) $F : \mathbb{R} \rightarrow [0, 1]$, definida por $F(t) = P(T \leq t)$, de uma variável aleatória contínua X não negativa, encontradas em James (2006), a saber:

- (a) F é não-decrescente.
- (b) $F(0) = P(T \leq 0) = 0$ e $\lim_{t \rightarrow \infty} F(t) = 1$.
- (c) F é contínua se, e somente se, X é uma variável aleatória contínua.

(i) Sejam $t_1, t_2 \in \mathbb{R}$, tais que $t_1 \leq t_2$. Como F é não-decrescente pelo item (a) temos que $F(t_1) \leq F(t_2)$, ou seja, $-F(t_1) \geq -F(t_2) \Rightarrow 1 - F(t_1) \geq 1 - F(t_2)$, donde concluímos que $S(t_1) \geq S(t_2)$.

(ii) Do item (b) obtemos que

$$S(0) = 1 - F(0) = 1$$

e

$$\lim_{t \rightarrow \infty} S(t) = \lim_{t \rightarrow \infty} [1 - F(t)] = 1 - \lim_{t \rightarrow \infty} F(t) = 1 - 1 = 0.$$

(iii) Temos que a função constante 1 é contínua e do item (c) temos que $F(t)$ é contínua. Como a diferença de funções contínuas é uma função contínua segue que $S(t) = 1 - F(t)$ é uma função contínua.

■

Propriedade 1.2. *Se T é uma variável aleatória absolutamente contínua e não-negativa, então*

$$f(t) = -\frac{d}{dt}S(t),$$

nos pontos em que S é derivável.

Demonstração. Sendo T uma variável aleatória contínua, segue que

$$S(t) = 1 - \int_{-\infty}^t f(x)dx.$$

Daí, nos pontos em que S é derivável segue do Teorema Fundamental do Cálculo que

$$\frac{d}{dt}S(t) = -f(t) \Rightarrow f(t) = -\frac{d}{dt}S(t).$$

■

1.2.2 Função de taxa de falha

Sendo T uma variável aleatória contínua não negativa a probabilidade de ocorrer falha em um intervalo de tempo $[t_1, t_2)$, ou seja, $P(t_1 \leq T < t_2)$ pode ser expressa por

$$S(t_1) - S(t_2).$$

A função taxa de falha, $\lambda(t_1, t_2)$, no intervalo $[t_1, t_2)$ é definida por Colosimo e Giolo (2006) como sendo a probabilidade de que a falha ocorra neste intervalo, dado que não ocorreu antes de t_1 , dividida pelo comprimento do intervalo, ou seja,

$$\begin{aligned}
\lambda(t_1, t_2) &= \frac{P(t_1 \leq T < t_2 | T \geq t_1)}{t_2 - t_1} \\
&= \frac{P(t_1 \leq T < t_2)}{P(T \geq t_1)(t_2 - t_1)} \\
&= \frac{P(T \geq t_1) - P(T \geq t_2)}{P(T \geq t_1)(t_2 - t_1)} \\
&= \frac{S(t_1) - S(t_2)}{S(t_1)(t_2 - t_1)}.
\end{aligned} \tag{1.1}$$

De modo geral, escrevendo o intervalo da forma $[t, t + \Delta t)$, a Equação (1.1) anterior assume a seguinte forma:

$$\lambda(t) = \frac{S(t) - S(t + \Delta t)}{\Delta t S(t)}, \tag{1.2}$$

o que nos leva então a definição à seguir.

Definição 1.2. *A função de taxa de falha instantânea, também conhecida como função de risco, de uma variável aleatória não negativa, T , é dada por*

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}.$$

A propriedade à seguir relaciona a função de taxa de falha instantânea com a função de sobrevivência.

Propriedade 1.3. *Se T é uma variável aleatória contínua, não negativa, com função taxa de falha instantânea $h(t)$ e função de sobrevivência $S(t)$, então*

$$h(t) = -\frac{d}{dt} (\log(S(t))).$$

Demonstração. Por definição temos que a função taxa de falha instantânea é dada por

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t},$$

utilizando (1.2) obtemos,

$$\begin{aligned}
 h(t) &= \lim_{\Delta t \rightarrow 0} \frac{S(t) - S(t + \Delta t)}{\Delta t S(t)} \\
 &= -\frac{1}{S(t)} \lim_{\Delta t \rightarrow 0} \frac{S(t + \Delta t) - S(t)}{\Delta t} \\
 &= -\frac{1}{S(t)} \frac{d}{dt} (S(t)) \\
 &= -\frac{d}{dt} (\log(S(t))).
 \end{aligned}$$

■

Exemplo 1.2. Usaremos agora a propriedade (1.3) para obter a função de taxa de falha instantânea da distribuição Weibull. Para isso, lembremos que no Exemplo (1.1) obtivemos que $S(t) = \exp \left\{ - \left(\frac{t}{\alpha} \right)^\gamma \right\}$. Assim,

$$\begin{aligned}
 h(t) &= -\frac{d}{dt} \log \left\{ \exp \left[- \left(\frac{t}{\alpha} \right)^\gamma \right] \right\} \\
 &= -\frac{d}{dt} \left[- \left(\frac{t}{\alpha} \right)^\gamma \right] \\
 &= \frac{\gamma}{\alpha^\gamma} t^{\gamma-1}.
 \end{aligned}$$

A Figura (1.2) ilustra graficamente a função taxa de falha instantânea da distribuição Weibull para diferentes valores dos parâmetros (γ, α) .

No caso da distribuição Weibull, notamos que $h(t)$ é crescente para valores de $\gamma > 1$, decrescente para valores de $\gamma < 1$ e constante para $\gamma = 1$.

1.2.3 Função de taxa de falha acumulada

Definição 1.3. Seja T uma variável aleatória contínua não negativa, então a função de taxa de falha acumulada de T é dada por

$$\Lambda(t) = \int_0^t h(u) du.$$

Conforme consta em Colosimo e Giolo (2006), como o próprio nome diz, essa função fornece a taxa de falha acumulada do indivíduo. Esta função não tem uma interpretação direta, porém é bastante útil na avaliação da função de maior interesse que é a de taxa de falha instantânea, $h(t)$. Uma vez que na estimação não paramétrica $\Lambda(t)$ possui um estimador com propriedades ótimas enquanto a função de taxa de falha instantânea é difícil de ser estimada.

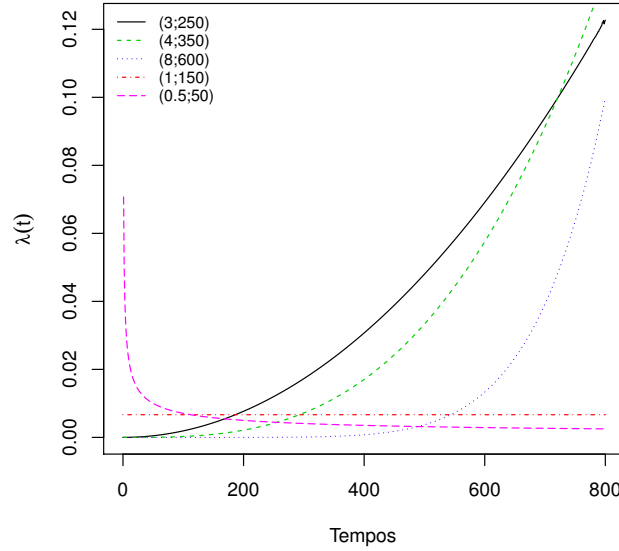


Figura 1.2: Gráfico da função de taxa de falha instantânea da distribuição Weibull para alguns valores dos parâmetros (γ, α) .

Propriedade 1.4. *Seja T uma variável aleatória contínua não negativa, então sua função de taxa de falha acumulada e sua função de sobrevivência estão relacionadas por meio da seguinte equação*

$$\Lambda(t) = -\log(S(t)).$$

Demonstração. Por definição, $\Lambda(t) = \int_0^t h(t)dt$, usando a Propriedade (1.3) temos que, $\Lambda(t) = \int_0^t -\frac{d}{dt}(\log S(t)) dt$. Assim, pelo Teorema Fundamental do Cálculo e do fato de que $S(0) = 1$ concluímos que $\Lambda(t) = -\log(S(t))$. ■

Exemplo 1.3. *Usando o Exemplo (1.1) e a Propriedade (1.4) temos que a função taxa de falha acumulada para a distribuição Weibull é*

$$\Lambda(t) = -\log(S(t)) = -\log \left\{ \exp \left[-\left(\frac{t}{\alpha} \right)^\gamma \right] \right\} = \left(\frac{t}{\alpha} \right)^\gamma.$$

O teorema que apresentamos agora é de grande valia no cálculo de esperança de v.a's contínuas não negativas, o mesmo encontra-se demonstrado em James (2006).

Teorema 1. *Se T é uma variável aleatória que assume apenas valores não-negativos, então*

$$\mathbb{E}(T) = \int_0^{\infty} P(T > t)dt = \int_0^{\infty} S(t)dt.$$

De acordo com Kalbfleish e Prentice (2002) outras duas funções, são de interesse na análise de sobrevivência, o tempo médio de vida e a vida média residual, esta última mede o tempo médio de vida restante de um indivíduo/observação com idade t . A seguir temos as definições das mesmas.

Definição 1.4. *O tempo médio de vida para uma observação é dado por*

$$t_m = \mathbb{E}(T).$$

Segue do Teorema 1 que,

$$t_m = \int_0^{\infty} S(t)dt.$$

Definição 1.5. *A vida média residual de uma observação condicionada a um tempo de vida t é dado por*

$$Vmr(t) = \mathbb{E}(T - t|T \geq t).$$

Notemos que a $Vmr(t)$ pode ser simplificada da seguinte forma,

$$\begin{aligned} \mathbb{E}(T - t|T \geq t) &= \frac{\int_t^{\infty} (u - t)f(u)du}{S(t)} \\ &\stackrel{*}{=} \frac{\int_t^{\infty} S(u)du}{S(t)}. \end{aligned}$$

Em (*) usamos integração por partes, a Propriedade (1.2) e o fato de que $\mathbb{E}(T) < \infty$ implica que $\lim_{t \rightarrow \infty} tF(t) = 0$. Observemos que o tempo médio de vida e a vida média residual estão relacionadas por meio da seguinte equação.

$$Vmr(0) = \mathbb{E}(T) = t_m.$$

Temos ainda as seguintes relações:

$$S(t) = \frac{Vmr(0)}{Vmr(t)} \exp \left\{ - \int_0^t \frac{du}{Vmr(u)} \right\}$$

e

$$h(t) = \left(\frac{dVmr(t)}{dt} + 1 \right) / Vmr(t).$$

1.3 Modelos paramétricos

Comumente há um interesse em se estimar a função de sobrevivência e isso pode ser feito por meio de métodos paramétricos ou não paramétricos. Considerando apenas dados não censurados, ou seja, que tenhamos o tempo de falha exato um estimador natural para a função de sobrevivência avaliada em t é

$$\hat{S}(t) = \frac{\text{número de indivíduos que não falharam até o tempo } t}{\text{número de indivíduos no estudo}}.$$

Como na prática, o conjunto de dados amostrais geralmente apresenta censura é necessário uso de técnicas para que se possa acomodar a informação contida nestas observações. Entre essas técnicas temos o estimador de Nelson-Aalen, proposto por Nelson (1972) e estudado por Aalen (1978), e o estimador de Kaplan-Meier, proposto por Kaplan e Meier (1958), que é o mais conhecido e utilizado não só em estudos de sobrevivência como também de confiabilidade. O estimador de Kaplan-Meier também é conhecido como estimador limite-produto. Para dados com censura intervalar temos o estimador limite-produto modificado proposto por Turnbull (1976), para maiores detalhes sobre o mesmo veja Colosimo e Giolo (2006). Neste trabalho nos concentraremos na estimação paramétrica. Para proceder com tal estimação, comumente nos estudos de análise de sobrevivência, faz-se uso de algumas distribuições de probabilidade as quais historicamente apresentaram-se adequadas para descrever o tempo de falha em determinadas situações, são elas:

A distribuição exponencial conhecida por ser a única que possui taxa de falha constante. Cox e Snell (1981) utilizaram o modelo exponencial para descrever o tempo de vida de pacientes adultos com Leucemia.

A distribuição Weibull que é largamente utilizada em aplicações práticas devido ao fato dela apresentar uma grande variedade de formas, todas com uma propriedade básica: a sua função taxa de falha é monótona conforme vimos no Exemplo (1.2).

A distribuição Log-normal que assim como a distribuição de Weibull é muito utilizada para caracterizar tempo de vida de indivíduos. Em estudos de confiabilidade a distribuição Log-normal pode ser usada, por exemplo, para modelar a fadiga de metais, semicondutores, diodos e isolamento elétrica.

A distribuição Log-logística que em situações práticas, se apresenta como uma alternativa à Weibull e a log-normal.

A distribuição Gama e a Gama Generalizada que inclui como caso particular a distribuição Weibull utilizada por Brown e Flood (1947) para descrever o tempo de vida de copos de vidro circulando em uma cafeteria e, também, por Birnbaum e Saunders (1958) para descrever o tempo de vida de materiais eletrônicos e por isso tornou-se amplamente utilizada em estudos de confiabilidade.

A distribuição Birnbaum-Saunders que conforme veremos teve como motivação um problema da área de confiabilidade. Na área médica sua utilização é mais recente e ocorreu através de Leiva et al. (2007) e Barros et al. (2008) onde a mesma modelou o tempo de vida de pacientes com mieloma múltiplo e câncer de pulmão.

Existem outras distribuições de probabilidade apropriadas para modelar tempo de falhas de materiais e tempo de vida em situações clínicas. Dentre elas estão as distribuições log-gama, Rayleigh, normal inversa, Gompertz e Birnbaum Saunders generalizada. Para mais detalhes dessas distribuições ver Marshall e Olkin (2007). No decorrer desse trabalho iremos dar ênfase as distribuições Birnbaum-Saunders e Birnbaum-Saunders generalizada.

1.4 O método da máxima verossimilhança

Uma vez escolhido o modelo probabilístico a ser usado, comumente precisamos estimar os parâmetros do mesmo, para tal finalidade existe algumas técnicas, dentre as mais conhecidas temos o método dos momentos, o método dos mínimos quadrados e o método da máxima verossimilhança. Apresentaremos o método da máxima verossimilhança e uma adaptação do mesmo para quando estamos na presença de dados com censura.

Lembremos que escolhido o modelo probabilístico a ser usado e tomada uma amostra, o método da máxima verossimilhança busca o parâmetro com maior probabilidade de ter gerado tal amostra entre todos os possíveis parâmetros para a distribuição.

Afim de se obter o estimador de máxima verossimilhança conforme mencionado acima deve-se tomar uma amostra aleatória (simples) de observações t_1, \dots, t_n de uma certa população de interesse em que todas as observações são não-censuradas. Suponhamos também que a população é caracterizada pela sua função de densidade $f(t)$ a qual depende de um parâmetro θ . A função de verossimilhança para um parâmetro

genérico θ desta população é expressa por:

$$\mathbb{L}(\theta) = \prod_{i=1}^n f(t_i; \theta).$$

Temos assim, que \mathbb{L} é uma função de θ e encontrar o estimador de máxima verossimilhança corresponde a encontrar o valor $\hat{\theta}$ no qual a função \mathbb{L} atinge seu máximo para a amostra fixada. Vale salientar também que na expressão acima θ pode representar um único parâmetro ou um conjunto de parâmetros.

Nota-se que de acordo com o procedimento anterior não está claro como usar o método da máxima verossimilhança na presença de dados censurados. De acordo com Klein e Moeschberger (2003), na presença de censura cada indivíduo contribui para a verossimilhança com uma informação específica. Um indivíduo que, por exemplo, apresente um tempo exato de falha, contribui para a função de verossimilhança com a probabilidade de ocorrência do evento de interesse neste tempo. Esta contribuição é dada pela fdp de T neste respectivo tempo. Por outro lado, a contribuição de cada indivíduo censurado à direita é dada pela função de sobrevivência de T avaliada no último tempo de visita. Similarmente, a contribuição de um indivíduo censurado à esquerda é dada pela fda de T avaliada no tempo da primeira visita. Finalmente, a contribuição de um indivíduo, que apresente um tempo de falha em um certo intervalo, é dada pela probabilidade de que o tempo de ocorrência do evento pertença a este intervalo. Em resumo, a forma geral da função de verossimilhança na presença de censura é

$$\mathbb{L}(\theta) = \left[\prod_{i \in L} (1 - S(t_i)) \right] \left[\prod_{i \in I} (S(t_{i-1}) - S(t_i)) \right] \left[\prod_{i \in C} S(t_i) \right] \left[\prod_{i \in D} f_T(t_i) \right] \quad (1.3)$$

onde,

- L é o conjunto dos índices para o qual a i -ésima observação sofreu censura a esquerda.
- I é o conjunto dos índices para o qual a i -ésima observação sofreu censura intervalar.
- C é o conjunto dos índices para o qual a i -ésima observação sofreu censura a direita.

- D é o conjunto dos índices para o qual a i -ésima observação falhou.

Neste trabalho estudaremos um conjunto de dados com censura à direita e censura intervalar.

Capítulo 2

A Distribuição Birnbaum-Saunders e a Distribuição Birnbaum-Saunders Generalizada

Nesse capítulo apresentaremos a origem da distribuição Birnbaum-Saunders segundo Birnbaum e Saunders (1969a) e as principais propriedades da mesma. Em seguida, apresentaremos a distribuição Senh-normal e relacionaremos a mesma com a distribuição Birnbaum-Saunders. Por fim, definiremos as distribuições de contorno elíptico e usaremos essas distribuições para definir a distribuição Birnbaum-Saunders generalizada conforme Díaz-García e Leiva (2005, 2006) exibindo também suas principais propriedades.

2.1 Origem e propriedades da distribuição Birnbaum-Saunders

Quando expomos um material a situações de estresse e tensão é natural que o mesmo venha a sofrer algum dano estrutural, a este dano damos o nome de **fadiga**. É comum na área de confiabilidade o interesse em modelar os processos de fadiga para descrever a variação aleatória dos tempos de falha associados aos materiais expostos à fadiga sob diferentes padrões e forças cíclicas. Com intuito de predizer a performance dos materiais sob diferentes condições podemos caracterizar os mesmos por parâmetros

de distribuições de tempo de vida. Os modelos probabilísticos mais utilizados para descrever a vida por fadiga são os modelos gama, lognormal e weibull, que apresentam um bom ajuste na região central da distribuição. Comumente estamos interessados nos percentis mais baixos ou mais altos da distribuição, regiões essas que geralmente possuem poucos dados, fazendo com que o ajuste dos modelos mencionados não seja bom. Birnbaum e Saunders (1969a) com o conhecimento sobre um tipo particular de fadiga criaram uma nova distribuição denominada de distribuição Birnbaum-Saunders a qual apresentou-se como uma excelente alternativa quando o interesse é analisar o ajuste de um modelo nos percentis mais baixos ou mais altos da distribuição. No que se segue, apresentaremos o processo utilizado por Birnbaum e Saunders (1969a) para o desenvolvimento dessa nova distribuição.

Para obter essa nova família de distribuições, Birnbaum e Saunders (1969a) consideraram um material sujeito a um padrão cíclico de tensão e força. Devido a exposição a situações de tensão e força os autores assumiram que a falha por fadiga ocorria devido ao início, crescimento e extensão final de uma rachadura dominante. Em outras palavras, a falha ocorria quando essa rachadura dominante excedia um comprimento w . Além disso, também assumiram que a sequência de tensão imposta ao material era a mesma de ciclo para ciclo. Assim, Birnbaum e Saunders (1969a) fizeram algumas suposições:

1. A extensão incremental da rachadura X_i resultante da aplicação da i -ésima oscilação de carga é uma variável aleatória com uma distribuição que só depende da rachadura atual causada pela tensão neste ciclo.

Segue dessa suposição que as extensões das rachaduras em diferentes ciclos são independentes.

2. A extensão da rachadura durante o $(j + 1)$ -ésimo ciclo é

$$Y_{j+1} = X_{jm+1} + \cdots + X_{jm+m}, \text{ para } j = 0, 1, \cdots,$$

em que X_{jm+i} é a extensão da rachadura (possivelmente microscópica) após a i -ésima oscilação de carga do $(j + 1)$ -ésimo ciclo.

3. A extensão total da rachadura, Y_j , devido ao j -ésimo ciclo é uma variável aleatória que segue uma distribuição com média μ e variância σ^2 , $\forall j = 1, 2, \cdots$.

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

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

com fda dada por

$$H_n(w) = P(W_n \leq w),$$

para $n = 1, 2, \dots$.

Seja C o número de ciclos requeridos até que seja observada a falha. A função de distribuição da variável aleatória C é

$$P(C \leq n) = P\left(\sum_{j=1}^n Y_j > w\right) = P(W_n > w) = 1 - H_n(w).$$

Agora, notemos que

$$\begin{aligned} P(C \leq n) &= P\left(\sum_{j=1}^n \frac{Y_j - \mu}{\sigma\sqrt{n}} > \frac{w - n\mu}{\sigma\sqrt{n}}\right) \\ &= 1 - P\left(\sum_{j=1}^n \frac{Y_j - \mu}{\sigma\sqrt{n}} \leq \frac{w - n\mu}{\sigma\sqrt{n}}\right). \end{aligned}$$

Como os Y_j 's são independentes supondo que os mesmos possuam a mesma distribuição segue do Teorema Central do Limite que

$$\begin{aligned} P(C \leq n) &\cong 1 - \Phi\left(\frac{w - n\mu}{\sigma\sqrt{n}}\right) \\ &= 1 - \Phi\left(-\frac{n\mu - w}{\sigma\sqrt{n}}\right) \\ &= \Phi\left(\frac{\mu\sqrt{n}}{\sigma} - \frac{w}{\sigma\sqrt{n}}\right), \end{aligned} \tag{2.1}$$

em que $\Phi(\cdot)$ representa a fda da distribuição normal padrão.

Usando a equação (2.1) Birnbaum e Saunders (1969a) definiram uma distribuição contínua de vida. Segundo os autores, substituindo n por uma variável aleatória real não negativa t , então a variável aleatória T é a extensão contínua da variável aleatória discreta C . Assim, T representa o tempo total até que ocorra a falha. Logo, tomando

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

a fda de T pode ser reescrita como

$$F_T(t) = P(T \leq t) = \Phi \left[\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right], \quad t > 0, \quad (2.2)$$

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

Segue, também, que uma variável aleatória $T \sim BS(\alpha, \beta)$ pode ser obtida a partir de uma variável aleatória Z com distribuição normal padrão mediante a seguinte relação

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

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

Segue diretamente de (2.2) e do fato da distribuição normal padrão ser simétrica em torno do zero que se $T \sim BS(\alpha, \beta)$, então a função de sobrevivência T , $S(t)$, é dada por

$$S_T(t) = 1 - F_T(t) = 1 - \Phi \left[\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right]. \quad (2.4)$$

Derivando-se (2.2) com relação a t obtém-se facilmente que se $T \sim BS(\alpha, \beta)$, então sua fdp é dada por

$$f_T(t) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2\alpha^2} \left(\frac{t}{\beta} + \frac{\beta}{t} - 2 \right) \right] \frac{t^{-3/2}(t + \beta)}{2\alpha\sqrt{\beta}}, \quad (2.5)$$

em que $t, \alpha, \beta > 0$.

Se $T \sim BS(\alpha, \beta)$ a função de taxa de falha instantânea de T , $h(t)$, é dada por

$$h_T(t) = \frac{f_T(t)}{S_T(t)},$$

em que $f_T(t)$ e $S_T(t)$ são dados em (2.5) e (2.4), respectivamente.

De acordo com a Figura (2.1) notamos que a função de taxa de falha instantânea da distribuição BS tem o seguinte comportamento: a mesma assume o valor zero em $t = 0$, cresce até um valor máximo e depois decresce até uma constante positiva.

A Figura (2.2) ilustra porque α é chamado de parâmetro de forma, uma vez que fixado o β e variando o α a forma do gráfico da densidade é alterada. Na Figura (2.3) temos a ilustração do porque β é chamado de parâmetro de escala uma vez que fixado α e variando β notamos que a forma do gráfico da distribuição é a mesma e apenas

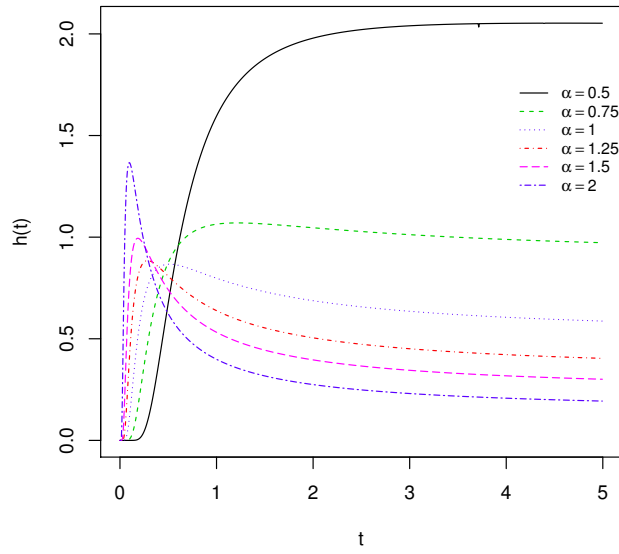


Figura 2.1: Funções de risco da distribuição Birnbaum-Saunders para os valores de α indicados e $\beta = 1$.

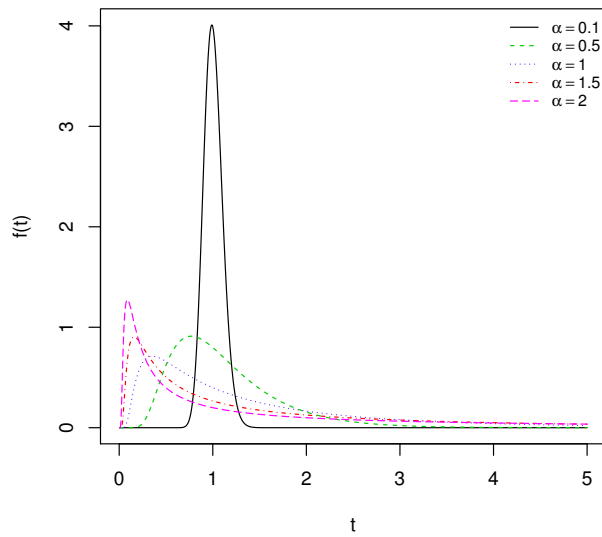


Figura 2.2: Gráfico da função densidade da distribuição Birnbaum-Saunders para os valores de $\alpha = 0, 1; 0, 5; 1, 0; 1, 5; 2, 0$ e β fixo igual 1.

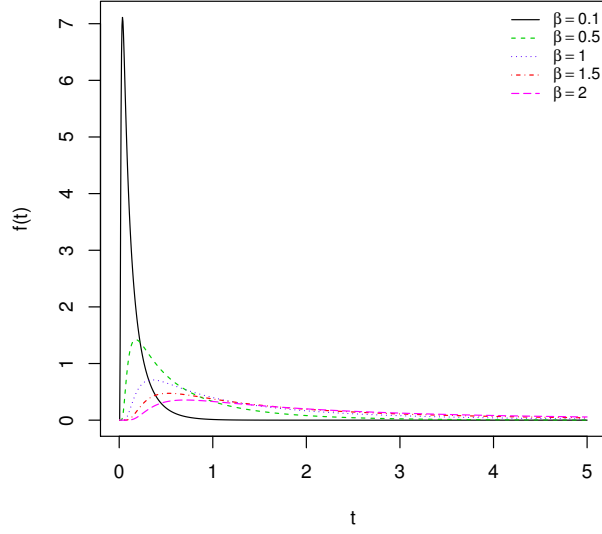


Figura 2.3: Gráfico da função densidade da distribuição Birnbaum-Saunders para os valores de $\beta = 0, 1; 0, 5; 1, 0; 1, 5; 2, 0$ e α fixo igual 1.

a escala é alterada. O parâmetro β corresponde também a mediana da distribuição conforme veremos posteriormente.

Se T é uma variável aleatória com distribuição Birnbaum Saunders, então T goza das seguintes propriedades:

Propriedade 2.1. Se $T \sim BS(\alpha, \beta)$ e $c > 0$, então $cT \sim BS(\alpha, c\beta)$.

Demonstração. Se $T \sim BS(\alpha, \beta)$ temos que $P(T \leq t) = \Phi \left\{ \frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right\}$.

Assim,

$$\begin{aligned} P(cT \leq t) &= P\left(T \leq \frac{t}{c}\right) \\ &= \Phi \left\{ \frac{1}{\alpha} \left(\sqrt{\frac{t}{c\beta}} - \sqrt{\frac{c\beta}{t}} \right) \right\}. \end{aligned}$$

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

Propriedade 2.2. Se $T \sim BS(\alpha, \beta)$, então $1/T \sim BS(\alpha, 1/\beta)$.

Demonstração. Seja $Y = 1/T$ temos que

$$\begin{aligned}
 F_Y(y) &= P(Y \leq y) \\
 &= P(1/T \leq y) \\
 &= P\left(T \geq \frac{1}{y}\right) \\
 &= 1 - F_T\left(\frac{1}{y}\right) \\
 &= \Phi\left[-\frac{1}{\alpha}\left(\sqrt{\frac{1}{y\beta}} - \sqrt{y\beta}\right)\right] \\
 &= \Phi\left[\frac{1}{\alpha}\left(\sqrt{y\beta} - \sqrt{\frac{1}{y\beta}}\right)\right].
 \end{aligned}$$

Segue que $Y \sim BS(\alpha, 1/\beta)$. ■

Propriedade 2.3. Se $T \sim BS(\alpha, \beta)$, então $\text{Med}(T) = \beta$.

Demonstração. Imediata, basta notar que $F_T(\beta) = \Phi(0) = \frac{1}{2}$. ■

Propriedade 2.4. Se $T \sim BS(\alpha, \beta)$, então existem a média e a variância de T e as mesmas são dadas, respectivamente, por

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

Demonstração. Veja Birnbaum-Saunders (1969b). ■

Para dados não censurados, Birnbaum-Saunders (1969a) propuseram a estimação dos parâmetros da distribuição Birnbaum-Saunders pelo método da máxima verossimilhança enquanto Ng, Kundu e Balakrishnan (2003) propõe o método dos momentos modificado. No caso de dados com censura intervalar construiremos a função logaritmo da verossimilhança de acordo com Klein e Moeschberger (2003) e faremos a estimação dos parâmetros fazendo uso do método da máxima verossimilhança em que obtemos um estimador com boas propriedades.

2.2 Relação entre as distribuições Birnbaum-Saunders e Senh-normal

No decorrer dessa seção apresentamos a distribuição senh-normal e suas principais características. Além disso, relacionamos a mesma com a distribuição Birnbaum-Saunders, relação essa que será de extrema importância para construção do modelo de regressão log-BS.

Definição 2.1. Dizemos que uma variável aleatória Y , segue a distribuição senh-normal com parâmetros α, γ e σ , denotado simplesmente por $Y \sim SN(\alpha, \gamma, \sigma)$, se sua fda é dada por

$$F_Y(y) = \Phi\left(\frac{2}{\alpha} \sinh\left(\frac{y - \gamma}{\sigma}\right)\right), \quad (2.6)$$

em que $\alpha > 0$, $\gamma \in \mathbb{R}$ e $\sigma > 0$ são respectivamente chamados de parâmetros de forma, de posição e de escala, enquanto $\Phi(\cdot)$ corresponde a fda da distribuição normal padrão.

Se $Y \sim SN(\alpha, \gamma, \sigma)$ segue de (2.6) e do fato da distribuição normal padrão ser simétrica em torno da origem que

$$S_Y(y) = \Phi\left(-\frac{2}{\alpha} \sinh\left(\frac{y - \gamma}{\sigma}\right)\right).$$

Derivando (2.6) com relação a y obtemos que a função de densidade de uma variável aleatória $Y \sim SN(\alpha, \gamma, \sigma)$ é dada por

$$f_Y(y) = \phi\left(\frac{2}{\alpha} \sinh\left(\frac{y - \gamma}{\sigma}\right)\right) \frac{2 \cosh[(y - \gamma)/\sigma]}{\alpha \sigma}, \quad y \in \mathbb{R}.$$

em que $\phi(\cdot)$ é a fdp da normal padrão.

Esta distribuição foi desenvolvida por Rieck e Nedelman (1991) mediante a transformação

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

onde $Z \sim N(0, 1)$. A mesma têm como principais características: simetria, graus de curtose maiores e menores que o modelo normal, admite unimodalidade e bimodalidade e possui como caso particular a distribuição log-Birnbaum-Saunders (log-BS).

A propriedade à seguir relaciona as distribuições Birnbaum-Saunders e senh-normal.

Propriedade 2.5. Uma variável aleatória $T \sim BS(\alpha, \beta)$ se, e somente se,

$$Y = \log(T) \sim SN(\alpha, \log(\beta), 2).$$

Demonstração. Suponhamos primeiramente que $T \sim BS(\alpha, \beta)$. Assim,

$$\begin{aligned} P(Y \leq y) &= P(\log(T) \leq y) \\ &= P(T \leq e^y) \\ &= \Phi \left\{ \frac{1}{\alpha} \left(\sqrt{\frac{e^y}{\beta}} - \sqrt{\frac{\beta}{e^y}} \right) \right\} \\ &= \Phi \left\{ \frac{1}{\alpha} \left(e^{\frac{y}{2}} \beta^{-\frac{1}{2}} - e^{-\frac{y}{2}} \beta^{\frac{1}{2}} \right) \right\} \\ &= \Phi \left\{ \frac{1}{\alpha} \left[\exp \left(\frac{y - \log(\beta)}{2} \right) - \exp \left(\frac{\log(\beta) - y}{2} \right) \right] \right\} \\ &= \Phi \left\{ \frac{1}{\alpha} \left[\exp \left(\frac{y - \log(\beta)}{2} \right) - \exp \left(-\frac{y - \log(\beta)}{2} \right) \right] \right\} \\ &= \Phi \left\{ \frac{2}{\alpha} \sinh \left(\frac{y - \log(\beta)}{2} \right) \right\}, \end{aligned}$$

ou seja, $Y \sim SN(\alpha, \log(\beta), 2)$.

Suponhamos agora que $Y = \log(T) \sim SN(\alpha, \log(\beta), 2)$. Assim,

$$\begin{aligned} P(T \leq t) &= P(\log(T) \leq \log(t)) \\ &= P(Y \leq \log(t)) \\ &= \Phi \left\{ \frac{2}{\alpha} \sinh \left(\frac{\log(t) - \log(\beta)}{2} \right) \right\} \\ &= \Phi \left\{ \frac{1}{\alpha} \left[\exp \left(\frac{\log(t) - \log(\beta)}{2} \right) - \exp \left(-\frac{\log(t) - \log(\beta)}{2} \right) \right] \right\} \\ &= \Phi \left\{ \frac{1}{\alpha} \left[\exp \left(\frac{\log(t) - \log(\beta)}{2} \right) - \exp \left(\frac{\log(\beta) - \log(t)}{2} \right) \right] \right\} \\ &= \Phi \left\{ \frac{1}{\alpha} \left(t^{\frac{1}{2}} \beta^{-\frac{1}{2}} - t^{-\frac{1}{2}} \beta^{\frac{1}{2}} \right) \right\} \\ &= \Phi \left\{ \frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right\}, \end{aligned}$$

ou seja, $T \sim BS(\alpha, \beta)$. ■

2.3 A distribuição Birnbaum-Saunders generalizada

A distribuição Birnbaum-Saunders Generalizada (BSG) proposta por Díaz-García e Leiva (2005) flexibiliza a distribuição do estresse cumulativo, que ao invés de ser assumida gaussiana é agora assumida simétrica. Isso gera uma nova classe de distribuições assimétricas com domínio positivo e caudas mais leves e/ou mais pesadas do que as caudas da distribuição BS clássica. A fim de apresentarmos a distribuição BSG precisamos primeiramente conhecer uma classe especial de distribuições denominada de distribuições de contornos elípticos, ou simplesmente, distribuições elípticas.

Definição 2.2. Dizemos que uma variável aleatória X segue uma distribuição de contorno elíptico se sua função característica tem a forma

$$\psi(t) = \exp(it\mu)\varphi(t^2\sigma^2), \quad (2.7)$$

com $\varphi : \mathbb{R}^+ \rightarrow \mathbb{R}$, ou se a fdp de X é dada por

$$f_X(x) = f\left[\frac{(x-\mu)^2}{\sigma^2}\right] = f(u); \quad x \in \mathbb{R}, \quad u > 0, \quad (2.8)$$

em que $f(u)$ é a fdp de X .

Notação: $X \sim EC(\mu, \sigma^2; \varphi)$ ou $X \sim EC(\mu, \sigma^2; f)$ de acordo com (2.7) ou (2.8), respectivamente.

Essa classe de distribuições possui esse nome devido ao fato do contorno de suas funções de densidades multivariadas possuírem a forma elíptica. A Figura (2.4) ilustra o comentário feito.

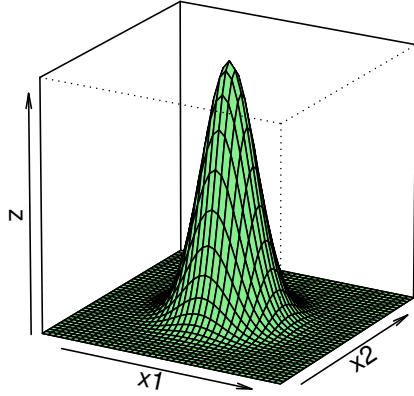
São exemplos de distribuições nessa classe de modelos as seguintes distribuições: Normal, Normal contaminada; t -Student; Exponencial Potência; Cauchy e Logística.

No decorrer das últimas décadas diversos autores tem trabalhado com distribuições de contornos elípticos, entre eles podemos citar, Fang e Anderson (1990), Gupta e Varga (1993), Galea et al. (2000), Barros (2007) e Pulgar (2009).

Na seção à seguir apresentaremos a distribuição BSG proposta por Díaz-García e Leiva (2005, 2006).

Distribuição Normal Bivariada para

$$\mu_1 = 0, \mu_2 = 0, \sigma_{11} = 1, \sigma_{22} = 1, \sigma_{12} = 0$$



Curvas de contorno

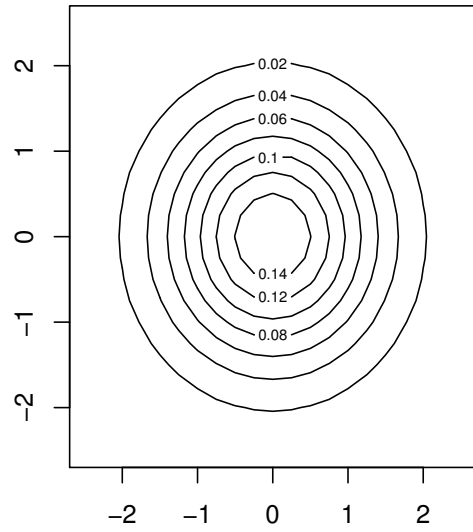


Figura 2.4: Gráfico da densidade normal padrão bivariada e suas curvas de contorno vista sobre o plano x_1 e x_2 .

2.3.1 Definição e algumas propriedades da distribuição BSG

Díaz-García e Leiva (2005, 2006) desenvolveram um novo modelo de sobrevivência com base em distribuições elípticas. O modelo proposto pelos mesmos consiste numa generalização da distribuição BS e foi obtido a partir de (2.3) relaxando a suposição de que $Z \sim N(0,1)$. Em outras palavras, eles consideraram

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

em que $U \sim EC(0,1; f)$. Assim, a nova variável aleatória T segue uma distribuição denominada de Birnbaum-Saunders generalizada, denotada por $T \sim BSG(\alpha, \beta; f)$.

Os parâmetros α e β , assim como na distribuição Birnbaum-Saunders, correspondem respectivamente, aos parâmetros de forma e escala e f é como dada em (2.8). De acordo com Barros et al. (2008) essa nova classe de distribuições assim obtida é mais flexível por conter distribuições que possuem caudas mais ou menos pesadas que a distribuição BS e distribuições bimodais, além de incluir como caso particular a

distribuição BS.

Definição 2.3. Dizemos que $T \sim BSG(\alpha, \beta; f)$ se sua função de densidade é dada por

$$f_T(t) = \frac{t^{-3/2}(t + \beta)}{2\alpha\beta^{1/2}} f\left(\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right), \quad t > 0, \quad (2.9)$$

em que f é a densidade de uma variável aleatória que pertence a classe de distribuições de contornos elípticos.

Fazendo

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

obtemos que

$$\frac{d}{dt} a_t(\alpha, \beta) = \frac{t^{-3/2}(t + \beta)}{2\alpha\beta^{1/2}}.$$

Assim, a equação (2.9) pode ser reescrita da seguinte forma:

$$f_T(t) = f(a_t(\alpha, \beta)) \frac{d}{dt} a_t(\alpha, \beta),$$

onde f é como dada em (2.8).

As funções de sobrevivência e de taxa de falha instantânea são dadas respectivamente, por

$$S_T = (t; \alpha, \beta) = 1 - F(a_t(\alpha, \beta))$$

e

$$h_T(t; \alpha, \beta) = \frac{f_T(t; \alpha, \beta)}{S_t(t; \alpha, \beta)} = \left(\frac{f(a_t(\alpha, \beta))}{1 - F(a_t(\alpha, \beta))}\right) \frac{d}{dt} a_t(\alpha, \beta).$$

Na Figura (2.5) temos o gráfico da função densidade de uma variável aleatória com distribuição BS- t de student para diferentes valores de ν . Através da mesma, notamos que quanto maior for ν mais próximo da distribuição BS fica a distribuição BS- t de student. Enquanto isso, na Figura (2.6) temos o gráfico da função de taxa de falha instantânea (ou função de risco) da distribuição BS- t de student para alguns valores de ν , através da qual notamos que essa função é não monótona, uma vez que a mesma se inicia no zero, atinge seu máximo e posteriormente decresce. Além disso, notamos também que quanto maior for ν maior a taxa de falha instantânea.

A distribuição Birnbaum-Saunders generalizada goza das seguintes propriedades que são apresentadas em Barros (2007).

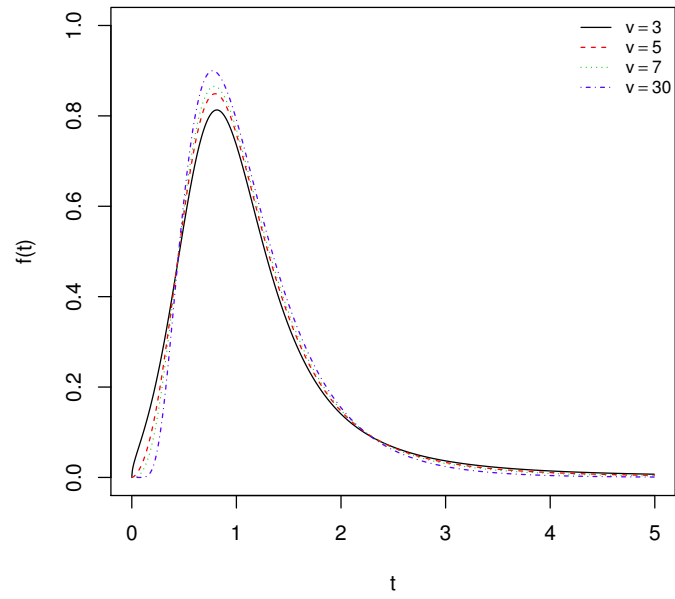


Figura 2.5: Densidade a distribuição BS t de student para $\nu = 3, 5, 7, 30$.

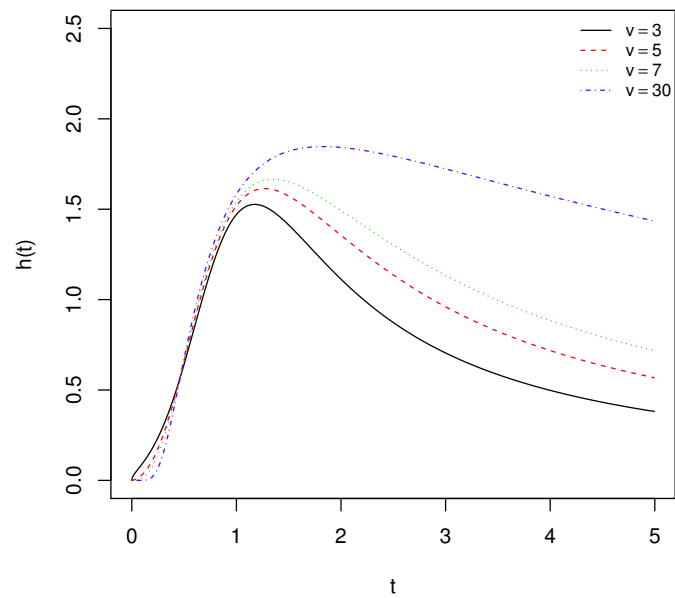


Figura 2.6: Função de risco para a distribuição BS t de student para $\nu = 3, 5, 7, 30$.

Propriedade 2.6. Se $T \sim BSG(\alpha, \beta; f)$, então

i) $cT \sim BSG(\alpha, c\beta; f) \forall c > 0$ (proporcionalidade)

ii) $T^{-1} \sim BSG(\alpha, \beta^{-1}; f)$ (reciprocidade)

iii) Se $T \sim BSG(\alpha, \beta; f)$ então $\beta T^{-1} \sim BSG(\alpha, 1; f)$.

Propriedade 2.7. Uma variável aleatória $T \sim BSG(\alpha, \beta; f)$ possui momentos se, e somente se, a função de distribuição de f possui momentos. Nesse caso,

$$\mathbb{E} \left(\left[\frac{T}{\beta} \right]^r \right) = \sum_{k=0}^r \binom{2r}{2k} \sum_{s=0}^k \binom{k}{s} \frac{(2(r+s-k))!}{(r+s-k)!} \frac{\varphi^{(r+s-k)}(0)}{i^{2(r+s-k)}} \left(\frac{\alpha}{2} \right)^{2(r+s-k)}.$$

Note que os momentos de T ficam em função da função de distribuição que gera a BSG.

Como consequência das propriedades de reciprocidade e proporcionalidade temos que as variáveis aleatórias βT^{-1} e $\beta^{-1}T$ têm a mesma distribuição e, portanto, obtemos que os momentos negativos de T são dados por

$$\mathbb{E}(T^{-n}) = \frac{\mathbb{E}(T^n)}{\beta^{2n}}. \quad (2.10)$$

A propriedade à seguir nos oferece uma maneira alternativa de encontrar os momentos da distribuição BSG.

Propriedade 2.8. Se $T \sim BSG(\alpha, \beta; f)$ e todos os momentos de T existem, então o n -ésimo momento de T é

$$\mathbb{E}(T^n) = \frac{\alpha^{2n}\beta^n}{2} \mathbb{E}(Z^{2n}) - \frac{1}{2} \left\{ \sum_{k=1}^n \binom{2n}{k} \frac{\mathbb{E}(T^{n-k})}{\beta^{-k}(-1)^k} + \beta^{2n} \sum_{k=n+1}^{2n-1} \binom{2n}{k} \frac{\mathbb{E}(T^{k-n})}{\beta^k(-1)^k} \right\} \quad (2.11)$$

para $n = 1, 2, \dots$, em que $Z \sim EC(0, 1; f)$.

Usando (2.10) e (2.11) recursivamente obtemos que os quatro primeiros momentos da distribuição BSG, são dados respectivamente por

$$(a) \quad \mathbb{E}(T) = \frac{\beta}{2}(2 + V_1\alpha^2),$$

$$(b) \quad \mathbb{E}(T^2) = \frac{\beta^2}{2}(2 + 4V_1\alpha^2 + V_2\alpha^4),$$

$$(c) \quad \mathbb{E}(T^3) = \frac{\beta^3}{2}(2 + 9V_1\alpha^2 + 6V_2\alpha^4 + V_3\alpha^6) \text{ e}$$

$$(d) \mathbb{E}(T^4) = \frac{\beta^4}{2}(2 + 16V_1\alpha^2 + 20V_2\alpha^4 + 8V_3\alpha^6 + V_4\alpha^8)$$

em que $V_n = \mathbb{E}(U^{2n})$ e $U \sim \text{EC}(0, 1; f)$.

Propriedade 2.9. *Seja $T \sim \text{BSG}(\alpha, \beta; f)$ tal que $\mathbb{E}(T^n)$ existe. Então, a variância e os coeficientes de variação, assimetria e curtose de T são dados, respectivamente, por*

$$\begin{aligned} \text{Var}(T) &= \frac{\beta^2}{4}(2\alpha^4V_2 + 4\alpha^2V_1 - \alpha^4V_1^2), \\ \gamma(T) &= \frac{(2V_2\alpha^4 + 4V_1\alpha^2 - V_1^2\alpha^4)^{1/2}}{2 + V_1\alpha^2}, \\ \alpha_3(T) &= \frac{2(6V_2\alpha + 2V_3\alpha^3 - 6V_1^2\alpha - 3V_1V_2\alpha^3 + V_1^3\alpha^3)}{(2V_2\alpha^2 + 4V_1 - V_1^2\alpha^2)^{3/2}} \\ &e \\ \alpha_4(T) &= \frac{16V_2 + (32V_3 - 48V_1V_2 + 24V_1^3)\alpha^2 + (8V_4 - 16V_1V_3 + 12V_1^2V_2 - 3V_1^4)\alpha^4}{(2V_2\alpha^2 + 4V_1 - V_1^2\alpha^2)^2}. \end{aligned}$$

2.3.2 A distribuição log-Birnbaum-Saunders generalizada

Díaz-García e Dominguez Molina (2006) apresentaram a distribuição log-Birnbaum-Saunders generalizada também conhecida como seno hiperbólica esférica, ou simplesmente senh-esférica, denotada por $SS(\alpha, \gamma, \sigma; f)$, cuja densidade é dada por

$$f_Y(y) = \frac{1}{\sigma} \left[\frac{2}{\alpha} \cosh \left(\frac{y - \gamma}{\sigma} \right) \right] f \left[\left(\frac{2}{\alpha} \sinh \left(\frac{y - \gamma}{\sigma} \right) \right)^2 \right],$$

em que $y \in \mathbb{R}$, $\alpha > 0$, $\gamma \in \mathbb{R}$, $\sigma > 0$, $\gamma = \log(\beta)$ e f é como dado em (2.8). Em outras palavras, se $T \sim \text{BSG}(\alpha, \beta; f)$, então

$$Y = \log(T) \sim SS(\alpha, \gamma, 2; f). \quad (2.12)$$

Notemos que quando f é a fdp da distribuição normal, ou seja, $T \sim \text{BS}(\alpha, \beta)$ temos que

$$Y = \log(T) \sim SN(\alpha, \log(\beta), 2).$$

As funções de distribuição acumulada e de sobrevivência para $Y \sim SS(\alpha, \gamma, 2; f)$ são dadas respectivamente por,

$$F_Y(y) = F \left[\frac{2}{\alpha} \sinh \left(\frac{y - \gamma}{\sigma} \right) \right], \quad (2.13)$$

e

$$S_Y(y) = F \left[-\frac{2}{\alpha} \sinh \left(\frac{y - \gamma}{\sigma} \right) \right], \quad (2.14)$$

em que F é a fda da distribuição elíptica que gera a distribuição BSG.

Capítulo 3

O Modelo de regressão log-Birnbaum-Saunders generalizado

Neste capítulo apresentamos o modelo de regressão log-BSG, seguindo a ideia da construção do modelo de regressão log-BS feito por Rieck e Nedelman (1991). Em seguida, de acordo com Hashimoto et al. (2010) obteremos o modelo de regressão log-BSG para dados censurados. Discutimos também a estimação dos parâmetros do modelo e apresentamos os resíduos de Cox-Snell ajustado, Martingale e Componente do desvio modificado. Apresentamos também um estudo de simulação de Monte Carlo para os modelos log-BS e log-BS- t de student com intuito de investigar o comportamento da distribuição empírica dos resíduos apresentados. Concluindo o capítulo, como medida de influência, apresentamos o afastamento pela verossimilhança sob diversos esquemas de perturbação.

3.1 A construção dos modelos de regressão log-BS e log-BSG para dados com censura intervalar

Para construirmos o modelo de regressão log-Birnbaum-Saunders generalizado para dados com censura intervalar seguiremos inicialmente a mesma ideia utilizada por Rieck e Nedelman (1991) para a construção do modelo log-BS relaxando ape-

nas a hipótese que $T \sim BS(\alpha_i, \eta_i)$. Consideremos então uma sequência de variáveis aleatórias independentes T_1, T_2, \dots, T_n onde $T_i \sim BS(\alpha_i, \eta_i; f)$ para $i = 1, \dots, n$, em que, f corresponde a fdp que gera a BSG. Além disso, suponhamos que a distribuição de cada variável aleatória T_i dependa de um vetor de p variáveis explicativas $\mathbf{x}_i^\top = (1, x_{i1}, x_{i2}, \dots, x_{ip})$, ou seja, estamos assumindo que

- (i) $\eta_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta})$ para $i = 1, \dots, n$, em que $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^\top$ é um vetor de parâmetros desconhecidos.
- (ii) O parâmetro de forma α_i independe do vetor de covariáveis \mathbf{x}_i^\top , ou seja, $\alpha_i = \alpha$ para $i = 1, \dots, n$.

Assim, considerando que $\zeta_i \sim BSG(\alpha, 1; f)$ com base na Propriedade (2.6) item (i) temos que $T_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta}) \zeta_i \sim BSG(\alpha, \exp(\mathbf{x}_i^\top \boldsymbol{\beta}); f)$, para $i = 1, \dots, n$. Considerando $y_i = \log(T_i)$, obtemos

$$y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \epsilon_i, \quad i = 1, \dots, n. \quad (3.1)$$

cuja componente aleatória ϵ_i é o $\log(\zeta_i)$ que por (2.12) tem distribuição $SS(\alpha, 0, 2; f)$.

Logo, estamos na presença de um modelo de regressão log-linear para a distribuição BSG. Porém, na presença de censura intervalar, os dados observados constituem de um intervalo $(\log(u_i), \log(v_i))$, para cada indivíduo, no qual esses intervalos são conhecidos e inclui y_i com probabilidade um, ou seja, $P(\log(u_i) \leq y_i \leq \log(v_i)) = 1$ e se $\log(v_i) = \infty$, então o logaritmo do tempo é censurado à direita para y_i .

O nosso objetivo nesse modelo será estimar $[\alpha, \boldsymbol{\beta}^\top]^\top$, para isso vamos assumir que os tempos de sobrevivência são censurados à direita e em um intervalo de tempo.

Considerando dados na presença de censura intervalar e de censura à direita, segue de (1.3), (2.13) e (2.14) que a função de verossimilhança para o modelo log-BSG será dada por

$$L(\theta) = \prod_{i \in I} (F(\xi_{i2v}) - F(\xi_{i2u})) \prod_{i \in C} F(-\xi_{i2u}) \quad (3.2)$$

em que,

$$\xi_{i2u} = \frac{2}{\alpha} \sinh\left(\frac{\log(u_i) - \mu_i}{2}\right), \quad \xi_{i2v} = \frac{2}{\alpha} \sinh\left(\frac{\log(v_i) - \mu_i}{2}\right) \quad \text{e} \quad \mu_i = \mathbf{x}_i^\top \boldsymbol{\beta}.$$

Daí, temos que o logaritmo da função de verossimilhança é dada por

$$\begin{aligned} l(\boldsymbol{\theta}) &= \sum_{i \in I} \log(F(\xi_{i2v}) - F(\xi_{i2u})) + \sum_{i \in C} \log(F(-\xi_{i2u})) \\ &= \sum_{i \in I} l_1(\alpha, \xi_{i2u}, \xi_{i2v}) + \sum_{i \in C} l_2(\alpha, \xi_{i2u}) \end{aligned} \quad (3.3)$$

com,

$$l_1(\alpha, \xi_{i2u}, \xi_{i2v}) = \log(F(\xi_{i2v}) - F(\xi_{i2u})) \quad (3.4)$$

e

$$l_2(\alpha, \xi_{i2u}) = \log(F(-\xi_{i2u})). \quad (3.5)$$

As funções escores para α e β_j ($j = 1, \dots, p$), são dadas, respectivamente, por

$$U_\alpha(\boldsymbol{\theta}) = \frac{1}{\alpha} \sum_{i \in I} \frac{f(\xi_{i2u})\xi_{i2u} - f(\xi_{i2v})\xi_{i2v}}{F(\xi_{i2v}) - F(\xi_{i2u})} + \frac{1}{\alpha} \sum_{i \in C} h(\xi_{i2u})\xi_{i2u}$$

e

$$U_{\beta_j}(\boldsymbol{\theta}) = \frac{1}{2} \sum_{i \in I} x_{ij} \left(\frac{\xi_{i1u}f(\xi_{i2u}) - \xi_{i1v}f(\xi_{i2v})}{F(\xi_{i2v}) - F(\xi_{i2u})} \right) + \frac{1}{2} \sum_{i \in C} x_{ij}\xi_{i1u}h(\xi_{i2u})$$

em que, $f(a) = \frac{\partial F(a)}{\partial a}$ com F correspondendo a fda da distribuição que gerou a BSG, $\xi_{i1u} = \frac{2}{\alpha} \cosh\left(\frac{\log(u_i) - \mu_i}{2}\right)$ e $\xi_{i1v} = \frac{2}{\alpha} \cosh\left(\frac{\log(v_i) - \mu_i}{2}\right)$.

Os estimadores de máxima verossimilhança dos coeficientes de regressão e do parâmetro de forma são soluções das equações $\mathbf{U}_{\beta_j}(\boldsymbol{\theta}) = 0$ ($j = 1 \dots p$) e $U_\alpha(\boldsymbol{\theta}) = 0$, em que $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top)^\top$, $\mathbf{U}_{\beta_j}(\boldsymbol{\theta}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \beta_j}$ e $U_\alpha(\boldsymbol{\theta}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \alpha}$. Ocorre que neste caso as equações não possuem soluções analíticas, fazendo com que seja necessário a utilização de métodos iterativos para a obtenção das raízes. As mesmas podem ser obtidas numericamente maximizando o logaritmo da função de verossimilhança usando um algoritmo de otimização não linear, por exemplo, Newton-Raphson ou um algoritmo quase newton (BFGS, por exemplo). Para obtenção dos estimadores de máxima verossimilhança para o modelo apresentado nesse trabalho fizemos uso do método BFGS e implementamos um algoritmo no software `0x Doornik` (2011).

A inferência assintótica para o vetor de parâmetros de $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top)^\top$ pode ser feita com base na normalidade assintótica do estimador de máxima verossimilhança $\hat{\boldsymbol{\theta}}$, dada por

$$\hat{\boldsymbol{\theta}} \sim N_{(p+1)}(\boldsymbol{\theta}, \boldsymbol{\Sigma}_\theta),$$

em que Σ_{θ} é a matriz de variâncias e covariâncias de $\hat{\theta}$, que pode ser aproximada pelo inverso da matriz de informação de Fisher $-\ddot{\mathbf{L}}_{\theta\theta}^{-1}$, avaliada em $\hat{\theta}$, obtida a partir de

$$\ddot{\mathbf{L}}_{\theta\theta} = \begin{bmatrix} \ddot{\mathbf{L}}_{\alpha\alpha} & \ddot{\mathbf{L}}_{\alpha\beta} \\ \ddot{\mathbf{L}}_{\beta\alpha} & \ddot{\mathbf{L}}_{\beta\beta} \end{bmatrix} = \begin{bmatrix} \text{tr}(\mathbf{M}) & \mathbf{K}^{\top} \mathbf{X} \\ \mathbf{X}^{\top} \mathbf{K} & \mathbf{X}^{\top} \mathbf{D} \mathbf{X} \end{bmatrix}, \quad (3.6)$$

em que, $\mathbf{M} = \text{diag}(m_1(\boldsymbol{\theta}), m_2(\boldsymbol{\theta}), \dots, m_n(\boldsymbol{\theta}))$, com

$$m_i(\boldsymbol{\theta}) = \begin{cases} -\frac{1}{\alpha^2} \left(\frac{f(\xi_{i2u})\xi_{i2u} - f(\xi_{i2v})\xi_{i2v}}{F(\xi_{i2v}) - F(\xi_{i2u})} \right) + \frac{1}{\alpha} \left[\frac{\partial}{\partial \alpha} \left(\frac{f(\xi_{i2u})\xi_{i2u} - f(\xi_{i2v})\xi_{i2v}}{F(\xi_{i2v}) - F(\xi_{i2u})} \right) \right], & \text{se } i \in I \\ \frac{\xi_{i2u}^2}{\alpha^2} \left(\frac{f'(\xi_{i2u})F(-\xi_{i2u}) + f^2(\xi_{i2u})}{F^2(-\xi_{i2u})} \right), & \text{se } i \in C \end{cases},$$

$\mathbf{K} = (k_1(\boldsymbol{\theta}), k_2(\boldsymbol{\theta}), \dots, k_n(\boldsymbol{\theta}))^{\top}$, com

$$k_i(\boldsymbol{\theta}) = \begin{cases} \frac{g_{iu} - g_{iv}}{F(\xi_{i2v}) - F(\xi_{i2u})} + \frac{(f(\xi_{i2v})\xi_{i1v} - f(\xi_{i2u})\xi_{i1u})(f(\xi_{i2u})\xi_{i2u} - f(\xi_{i2v})\xi_{i2v})}{(F(\xi_{i2v}) - F(\xi_{i2u}))^2}, & \text{se } i \in I \\ \xi_{i1u} \left[\left(\frac{f'(\xi_{i2u})}{F(-\xi_{i2u})} + \left(\frac{f(\xi_{i2u})}{F(-\xi_{i2u})} \right)^2 \right) \xi_{i2u} + h(\xi_{i2u}) \right] & \text{se } i \in C \end{cases}$$

e $\mathbf{D} = \text{diag}(d_1(\boldsymbol{\theta}), d_2(\boldsymbol{\theta}), \dots, d_n(\boldsymbol{\theta}))$, com

$$d_i(\boldsymbol{\theta}) = \begin{cases} \frac{a_{iu} - a_{iv}}{F(\xi_{i2v}) - F(\xi_{i2u})} + \frac{(\xi_{i1u}f(\xi_{i2u}) - \xi_{i1v}f(\xi_{i2v}))^2}{[F(\xi_{i2v}) - F(\xi_{i2u})]^2} & \text{se } i \in I \\ \xi_{i2u}h(\xi_{i2u}) + \xi_{i1u}^2 \left(\frac{f'(\xi_{i2u})}{F(-\xi_{i2u})} + \left(\frac{f(\xi_{i2u})}{F(-\xi_{i2u})} \right)^2 \right) & \text{se } i \in C \end{cases}.$$

Os detalhes desses cálculos estão no Apêndice A.

3.2 Validação do modelo

De acordo com Paula (2010) uma etapa importante na análise de um ajuste de um modelo de regressão é a verificação de possíveis afastamentos das suposições feitas para o modelo, para a componente aleatória e para a parte sistemática do modelo, bem como

a existência de observações discrepantes com alguma interferência desproporcional ou inferencial nos resultados do ajuste. Essa etapa é conhecida como análise de diagnóstico e inicia-se com a análise de resíduos com o intuito de identificar a presença de pontos aberrantes e avaliar a adequação da distribuição da variável resposta para o modelo. Paula (2010) cita Cox e Snell (1968) como uma importante referência nesse quesito uma vez que nesse artigo os mesmos apresentaram uma forma bastante geral de definir resíduos. Atkinson (1985) propõe, por meio de simulações de Monte Carlo, a construção de bandas de confiança para os resíduos as quais denominou de envelope.

Uma técnica bastante conhecida para avaliar o impacto da retirada de uma observação particular nas estimativas do modelo de regressão é a deleção de casos. Segundo Paula (2010) uma das propostas mais inovadoras na área de diagnóstico em regressão foi apresentada por Cook (1986) que propõe avaliar a influência conjunta das observações sob pequenas perturbações no modelo e/ou aos dados, ao invés da avaliação pela retirada individual ou conjunta de pontos. Essa metodologia, foi denominada de influência local. Posteriormente, faremos uso dessas técnicas (deleção de casos e influência local) e daremos maiores detalhes sobre as mesmas.

3.2.1 Análise de resíduos

Nesta seção exibimos os resíduos de Cox-Snell ajustado e martingale apresentados em Farrington (2000) e os resíduos componente do desvio modificado conforme apresentado em Hashimoto et al. (2010). Posteriormente, faremos um estudo de simulação para verificar o comportamento de tais resíduos.

Resíduos de Cox-Snell ajustado

Quando estamos na presença de dados censurados os resíduos não podem ser obtidos ignorando esse fato. De acordo com Farrington (2000) um tipo de resíduo comumente utilizado para dados com censura à direita é o resíduo de Cox-Snell, Cox e Snell (1968). Esses resíduos são baseados nas observações de uma variável aleatória de tempo, T , com função de sobrevivência $S(t)$ tal que a variável aleatória $-\log S(T)$ é distribuída exponencialmente com média um.

De acordo com Farrington (2000) essa lógica pode ser aplicada em casos de cen-

sura intervalar. Dada uma amostra de intervalos censurados $(a_1, b_1], \dots, (a_n, b_n]$ e funções de sobrevivências estimadas \hat{S}_i , os intervalos $(-\log \hat{S}_i(a_i), -\log \hat{S}_i(b_i)]$, $i = 1, \dots, n$ podem ser aproximados pela amostra de censuras intervalares de uma distribuição exponencial com média um.

Uma deficiência na versão do resíduo de Cox-Snell para dados com censura intervalar é que os mesmos não são fáceis de manusear por se tratarem de intervalos. Segundo Farrington (2000), uma maneira de superar esse problema é substituir os resíduos intervalares por valores esperados sob a distribuição de uma exponencial de média um. Sendo U uma variável aleatória com distribuição exponencial de média um, definimos

$$\begin{aligned} r_{C_i}(a_i, b_i) &= E\{U|U \in [-\log(S_i(a_i)), -\log(S_i(b_i))]\} \\ &= \frac{\int_{-\log(S_i(a_i))}^{-\log(S_i(b_i))} ue^{-u} du}{\int_{-\log(S_i(a_i))}^{-\log(S_i(b_i))} e^{-v} dv} \\ &= \frac{S_i(a_i)\{1 - \log S_i(a_i)\} - S_i(b_i)\{1 - \log S_i(b_i)\}}{S_i(a_i) - S_i(b_i)} \end{aligned}$$

lembrando que $S_i(x_i) = e^{-x_i}$, $x_i \geq 0 \quad \forall i$.

Por simplicidade, escrevemos r_{C_i} para $r_{C_i}(a_i, b_i)$, quando não houver ambiguidade quanto a censura intervalar. Com isso, definimos o resíduo de Cox-Snell ajustado para dados com censura intervalar por

$$\hat{r}_{C_i} = \frac{\hat{S}(a_i)\{1 - \log \hat{S}(a_i)\} - \hat{S}(b_i)\{1 - \log \hat{S}(b_i)\}}{\hat{S}(a_i) - \hat{S}(b_i)}.$$

Para intervalos da forma $(\log(u_i), \log(v_i))$ temos

$$\hat{r}_{C_i} = \frac{\hat{S}[\log(u_i)]\{1 - \log \hat{S}[\log(u_i)]\} - \hat{S}[\log(v_i)]\{1 - \log \hat{S}[\log(v_i)]\}}{\hat{S}[\log(u_i)] - \hat{S}[\log(v_i)]}.$$

Tomando por base o modelo de regressão dado em (3.1) os resíduos do Cox-Snell ajustado fica dado por

$$\hat{r}_{C_i} = \begin{cases} \frac{(1 - F(\hat{\xi}_{i2u}))(1 - \log(1 - F(\hat{\xi}_{i2u}))) - (1 - F(\hat{\xi}_{i2v}))(1 - \log(1 - F(\hat{\xi}_{i2v})))}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})}, \\ \text{se } i \in I \\ 1 - \log(1 - F(\hat{\xi}_{i2u})), \\ \text{se } i \in C \end{cases}$$

em que ξ_{i2u} , ξ_{i2v} e $F(\cdot)$ são como dados em (3.2).

Resíduo de martingale

De acordo com Farrington (2000) e Hashimoto et al. (2010) apresentamos uma adaptação para os resíduos de Cox-Snell, a qual foi discutida por Lagakos (1980). Tal adaptação dos resíduos de Cox-Snell resulta em um novo resíduo chamado de resíduos Martingale (Barlow e Prentice, 1988). Esses resíduos são assimétricos e assume valor máximo em 1 e mínimo em $-\infty$, o mesmo é definido por

$$\hat{r}_{M_i} = 1 - \hat{r}_{C_i} = \frac{\hat{S}[\log(u_i)] \log \hat{S}[\log(u_i)] - \hat{S}[\log(v_i)] \log \hat{S}[\log(v_i)]}{\hat{S}[\log(u_i)] - \hat{S}[\log(v_i)]}.$$

Assim, para o modelo de regressão dado em (3.1) temos que

$$\hat{r}_{M_i} = \begin{cases} \frac{(1 - F(\hat{\xi}_{i2u})) \log(1 - F(\hat{\xi}_{i2u})) - (1 - F(\hat{\xi}_{i2v})) \log(1 - F(\hat{\xi}_{i2v}))}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})}, & \text{se } i \in I \\ \log(1 - F(\hat{\xi}_{i2u})), & \text{se } i \in C \end{cases}.$$

Resíduos componente do desvio modificado

O terceiro resíduo que estudaremos trata-se dos resíduos componente do desvio modificado, definido por:

$$r_{D_i} = \text{senal}(r_{M_i})[-2\{r_{M_i} + \delta_i \log(\delta_i - r_{M_i})\}]^{1/2},$$

em que δ_i é a variável indicadora de censura intervalar. Os mesmos são uma tentativa de tornar os resíduos martingale mais simétricos em torno do zero. Para o modelo dado em (3.1) os resíduos componente do desvio modificado são dados por,

$$\hat{r}_{D_i} = \begin{cases} \text{sinal}(\hat{r}_{M_i}) \left[-2 \left\{ \frac{\rho_{i2u} - \rho_{i2v}}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} + \log \left(1 - \frac{\rho_{i2u} - \rho_{i2v}}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} \right) \right\} \right]^{1/2}, \\ \text{se } i \in I \\ \text{sinal}(\hat{r}_{M_i})[-2 \log(1 - F(\hat{\xi}_{i2u}))], \\ \text{se } i \in C \end{cases},$$

em que $\rho_{i2u} = (1 - \Phi(\hat{\xi}_{i2u})) \log(1 - \Phi(\hat{\xi}_{i2u}))$ e $\rho_{i2v} = (1 - \Phi(\hat{\xi}_{i2v})) \log(1 - \Phi(\hat{\xi}_{i2v}))$.

O gráfico dos resíduos contra os tempos são uma ferramenta importante na análise, pois os mesmos nos dão uma forma de verificar a adequação do modelo ajustado e auxiliam a encontrar observações atípicas. Para o cálculo dos resíduos correspondentes aos modelos de regressão log-BS e log-BS- t de student, basta substituir F pelas respectivas fda que geram a BSG, a saber, Φ e Φ_t .

3.2.2 Estudo de simulação

Com o intuito de investigar o comportamento da distribuição empírica dos resíduos Cox-Snell ajustado, Martingale e componente do desvio modificado para os modelos de regressão log-BS e log-BS- t de student com censura intervalar e à direita, fizemos os seguintes estudos de simulação.

Caso 1 - Estudo de simulação para o modelo log-BS

Geramos os logaritmos dos tempos de sobrevivência y_1, y_2, \dots, y_n por meio de simulações de Monte Carlo considerando o modelo de regressão dado em (3.1), em que f corresponde a densidade da normal padrão, ou seja, o modelo de regressão log-BS. Os y_i foram gerados a partir da relação

$$y_i = x_0\beta_0 + x_1\beta_1 + \log \left\{ \left[\alpha U/2 + \sqrt{((\alpha^2 U^2)/4) + 1} \right]^2 \right\},$$

em que $U \sim N(0, 1)$, fixamos $\beta_0 = 4$, $\beta_1 = 2$, α sendo 1 e posteriormente 2, e os x_i com $i = 1, \dots, n$ foram gerados de uma distribuição uniforme(0, τ), onde τ foi ajustado até que fossem obtidas as porcentagens de censura à direita de 0%, 10% e 30% para amostras de tamanhos $n = 50, 100$ e 200 . Os logaritmos dos tempos com censura intervalar $\log(u_i)$ e $\log(v_i)$ foram gerados de forma que $y_i \in (\log(u_i), \log(v_i))$ com

probabilidade um, conforme em Hashimoto et al. (2010), seguindo o mesmo enfoque adotado por Zhao (2004).

Para cada combinação de n , α e porcentagem de censura, geramos 1000 amostras e para cada uma ajustamos o modelo de regressão log-BS com censura intervalar e à direita. Assim, para cada modelo ajustado, os resíduos r_{C_i} , r_{M_i} e r_{D_i} foram calculados e armazenados. As Figuras de (3.1) à (3.6) apresentam os resíduos versus os valores esperados das estatísticas de ordem da distribuição normal padrão. Esse gráfico corresponde ao gráfico normal de probabilidade e serve para avaliar o afastamento da suposição de normalidade para esses resíduos (Weisberg, 1985).

Dos gráficos apresentados temos as seguintes interpretações:

- i)* Nota-se que a distribuição empírica do resíduo componente do desvio modificado concorda com a distribuição normal padrão.
- ii)* As distribuições empíricas dos resíduos de Cox-Snell ajustado e do resíduo martingale apresentam um comportamento semelhante, sendo que ambos apresentam uma acentuada assimetria.
- iii)* Quanto menor a censura mais próxima da distribuição normal se torna a distribuição empírica do resíduo componente do desvio modificado.
- iv)* Quanto maior o tamanho da amostra, melhor a concordância entre a distribuição empírica do resíduo componente do desvio modificado e a distribuição normal padrão.

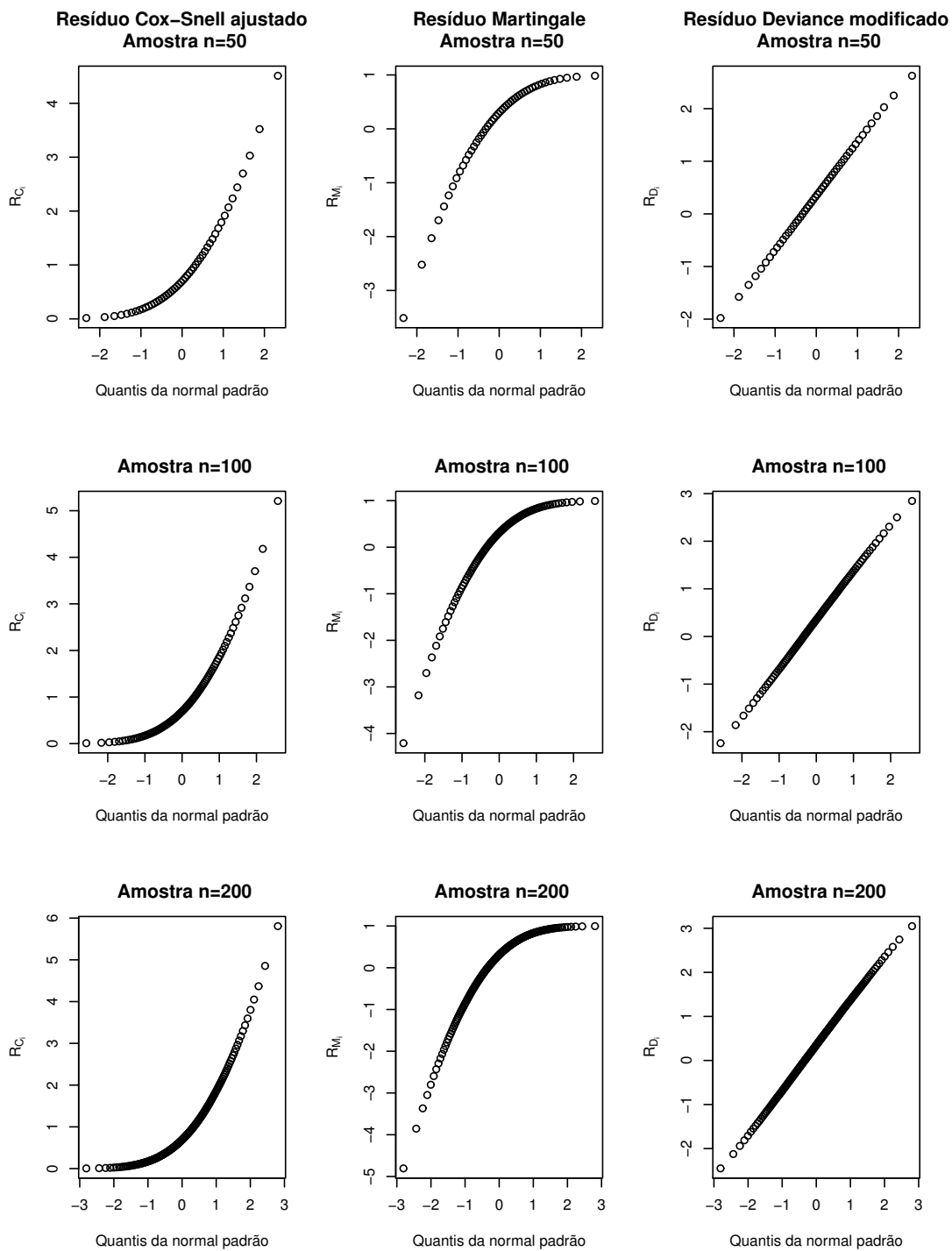


Figura 3.1: Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado (r_{C_i}), resíduo Martingale (r_{M_i}) e resíduo componente do desvio modificado (r_{D_i}) com porcentagem de censura à direita de 0% e $\alpha = 1$.

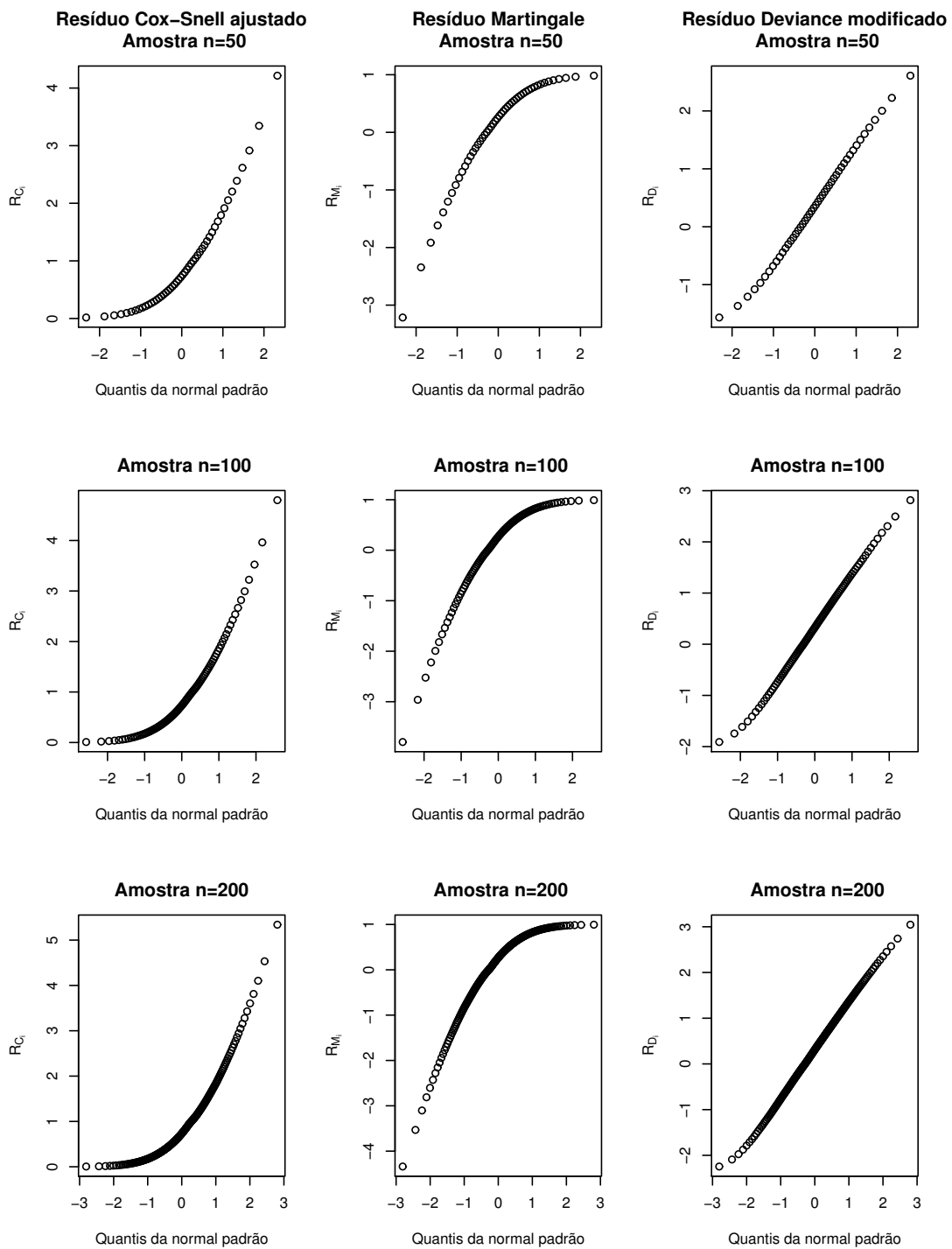


Figura 3.2: Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado (r_{C_i}), resíduo Martingale (r_{M_i}) e resíduo componente do desvio modificado (r_{D_i}) com porcentagem de censura à direita de 10% e $\alpha = 1$.

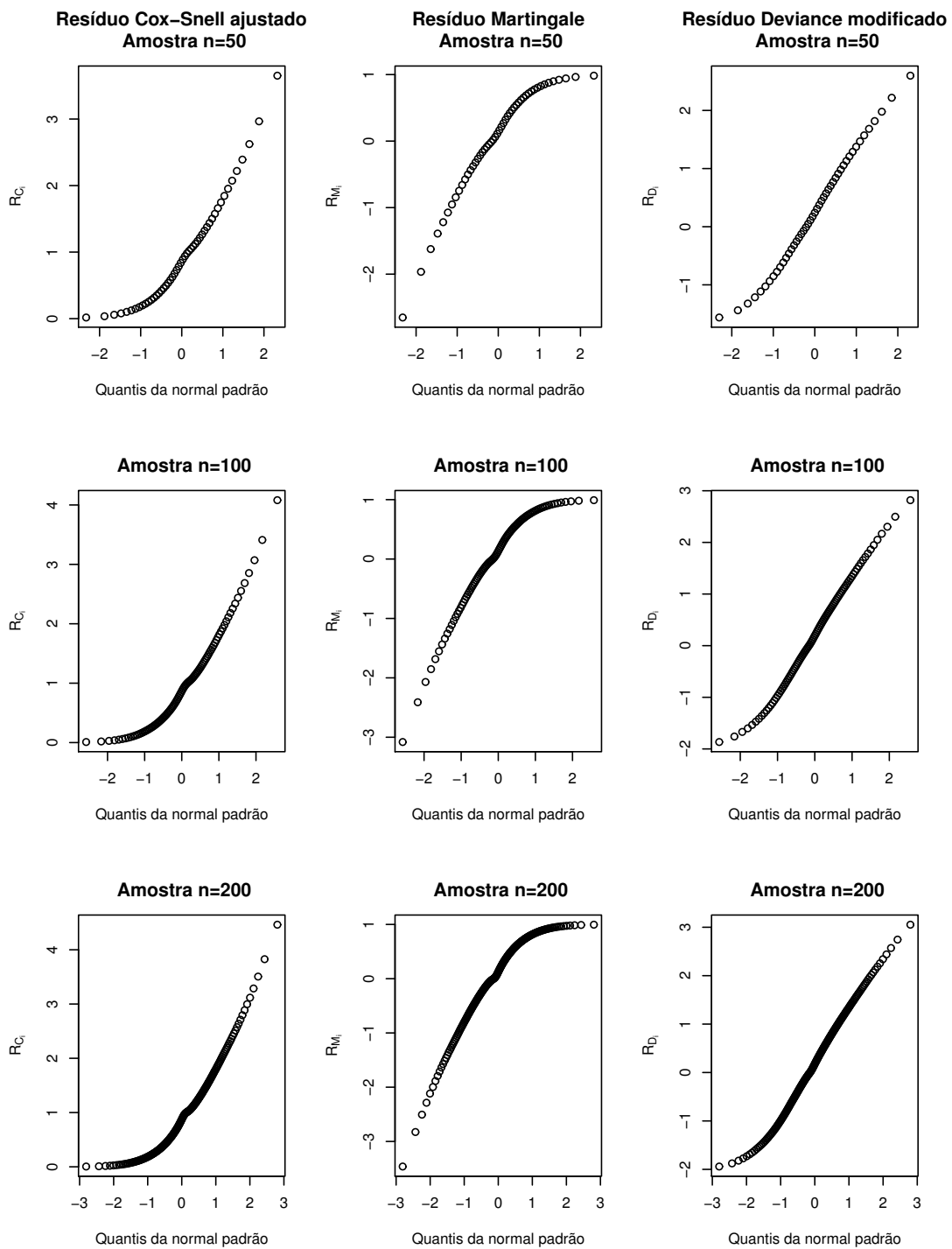


Figura 3.3: Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado (r_{C_i}), resíduo Martingale (r_{M_i}) e resíduo componente do desvio modificado (r_{D_i}) com porcentagem de censura à direita de 30% e $\alpha = 1$.

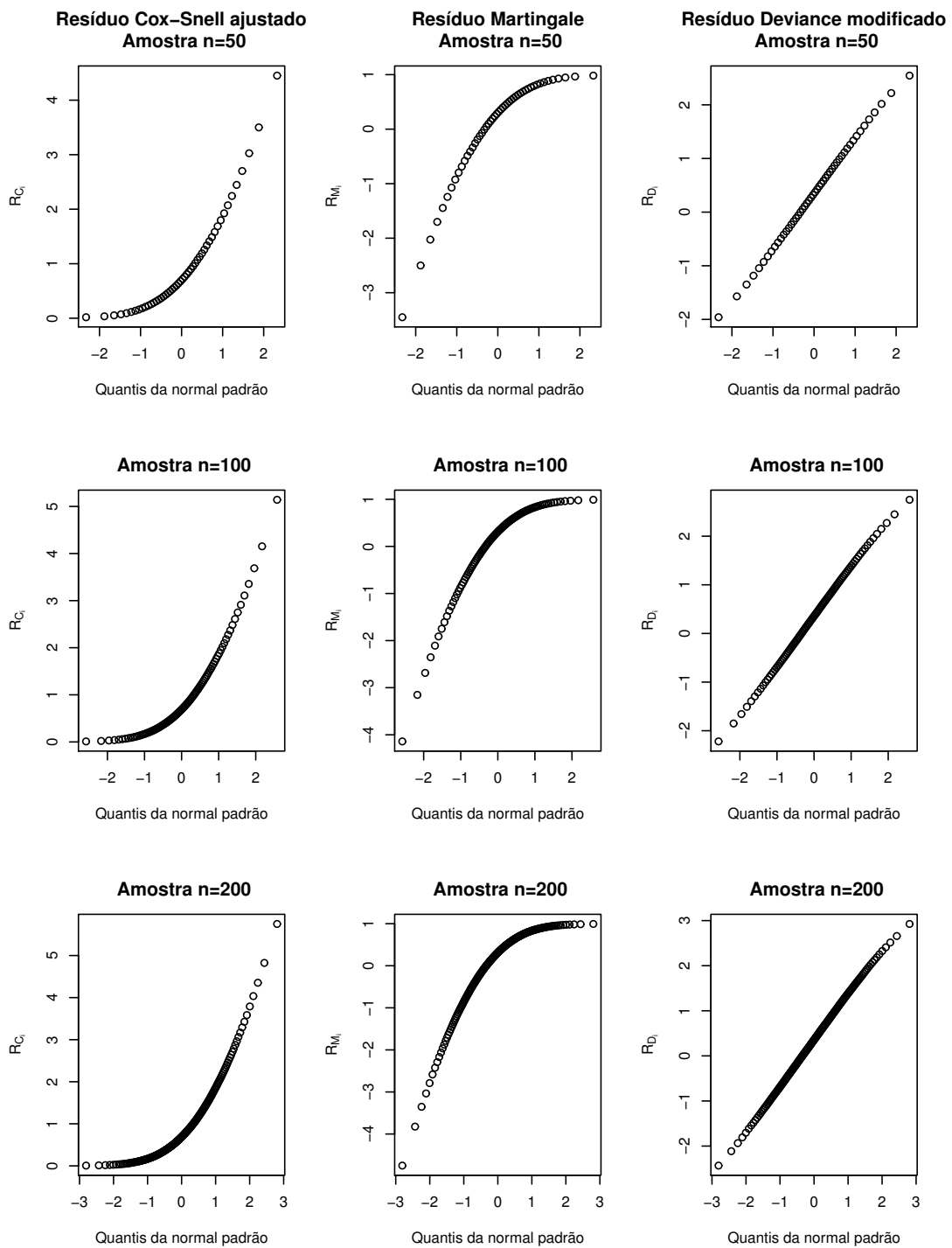


Figura 3.4: Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado (r_{C_i}), resíduo Martingale (r_{M_i}) e resíduo componente do desvio modificado (r_{D_i}) com porcentagem de censura à direita de 0% e $\alpha = 2$.

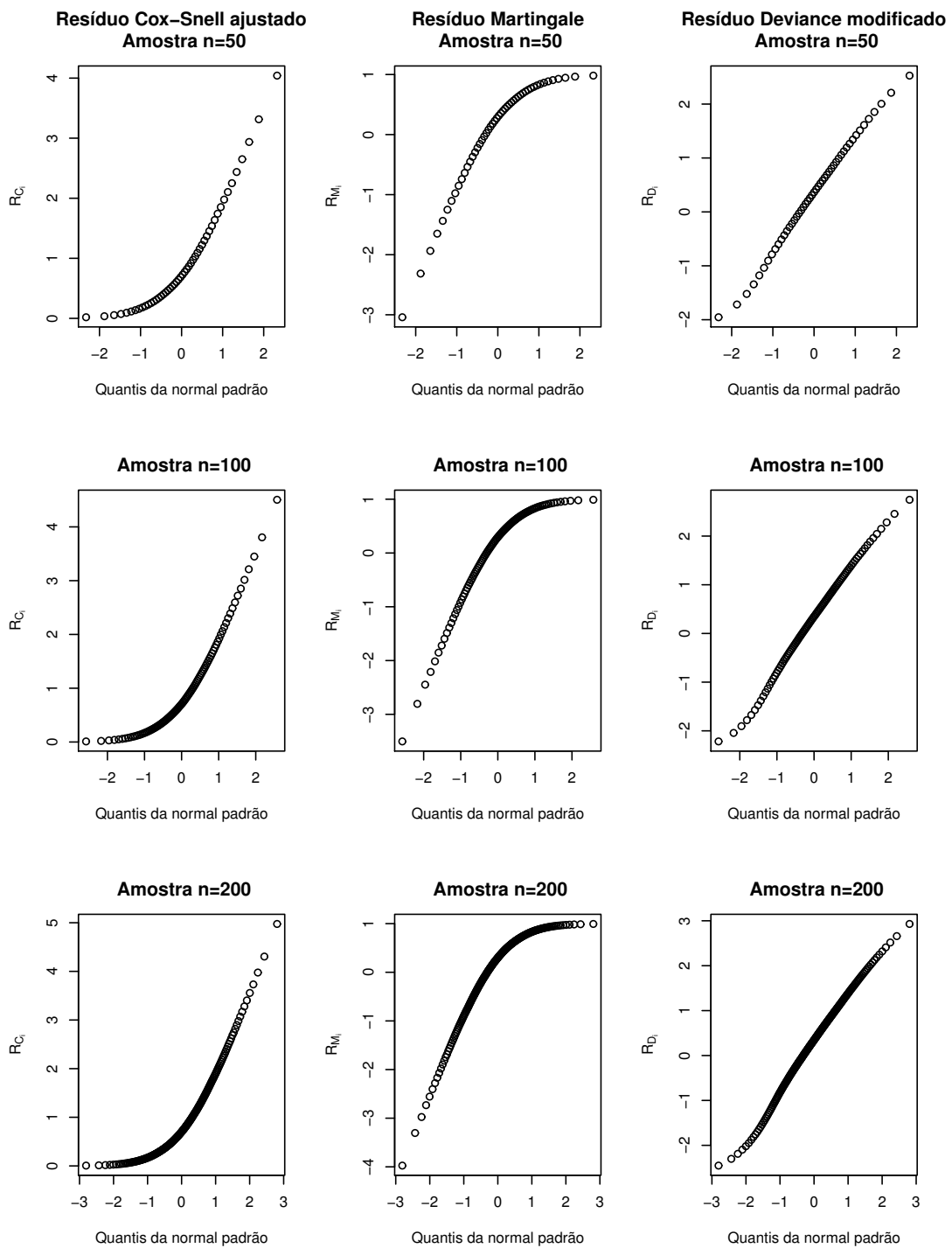


Figura 3.5: Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado (r_{C_i}), resíduo Martingale (r_{M_i}) e resíduo componente do desvio modificado (r_{D_i}) com porcentagem de censura à direita de 10% e $\alpha = 2$.

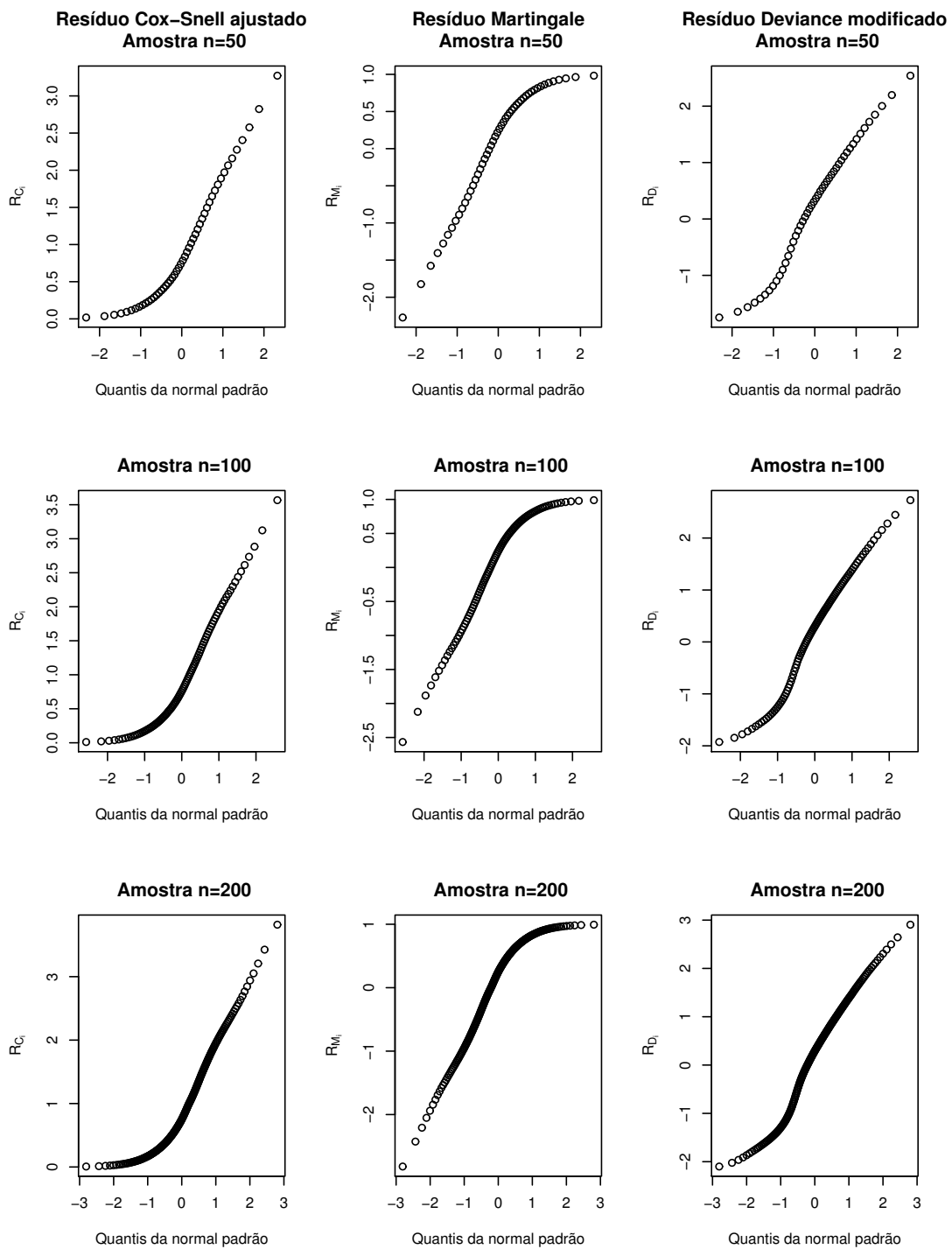


Figura 3.6: Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado (r_{C_i}), resíduo Martingale (r_{M_i}) e resíduo componente do desvio modificado (r_{D_i}) com porcentagem de censura à direita de 30% e $\alpha = 2$.

Caso 2 - Estudo de simulação para o modelo log-BS- t de student

Geramos os logaritmos dos tempos de sobrevivência y_1, y_2, \dots, y_n por meio de simulações de Monte Carlo considerando o modelo de regressão dado em (3.1), em que f corresponde a densidade de uma distribuição t_ν , ou seja, o modelo de regressão log-BS- t de student. Os y_i foram gerados a partir relação

$$y_i = x_0\beta_0 + x_1\beta_1 + \log \left\{ \left[\alpha U/2 + \sqrt{((\alpha^2 U^2)/4) + 1} \right]^2 \right\}$$

em que $U \sim t(\nu)$. Fixamos $\alpha = 1$, $\beta_0 = 4$ e $\beta_1 = 2$ e geramos os x_i com $i = 1, \dots, n$ de uma uniforme(0, τ), onde τ foi ajustado até que fossem obtidas as porcentagens de censura à direita de 0%, 10% e 30% para amostras de tamanho $n = 200$. Os logaritmos dos tempos com censura intervalar $\log(u_i)$ e $\log(v_i)$ foram gerados de forma que $y_i \in (\log(u_i), \log(v_i))$ com probabilidade um. Variamos o grau de liberdade ν e a porcentagem de censura, sendo que geramos 1000 amostras e para cada uma ajustamos o modelo regressão log-BS- t com censura intervalar. Com o ajuste feito, calculamos e armazenamos os resíduos de Cox-Snell, Martingale e componente do desvio modificado conforme é mostrado nas Figuras de (3.7)-(3.12).

Dos gráficos apresentados temos as seguintes interpretações:

- i)* Nota-se que a distribuição empírica do resíduo componente do desvio modificado concorda com a distribuição normal padrão.
- ii)* As distribuições empíricas dos resíduos de Cox-Snell ajustado e do resíduo martingale apresentam um comportamento semelhante, sendo que ambos apresentam uma acentuada assimetria.
- iii)* Quanto menor a censura mais próxima da distribuição normal se torna a distribuição empírica do resíduo componente do desvio modificado.
- iv)* Quanto maior o tamanho da amostra, melhor a concordância entre a distribuição empírica do resíduo componente do desvio modificado e a distribuição normal padrão.
- v)* Quanto maior o grau de liberdade da distribuição t de student mais próximo a distribuição empírica dos resíduos componente do desvio modificado fica da distribuição normal padrão.

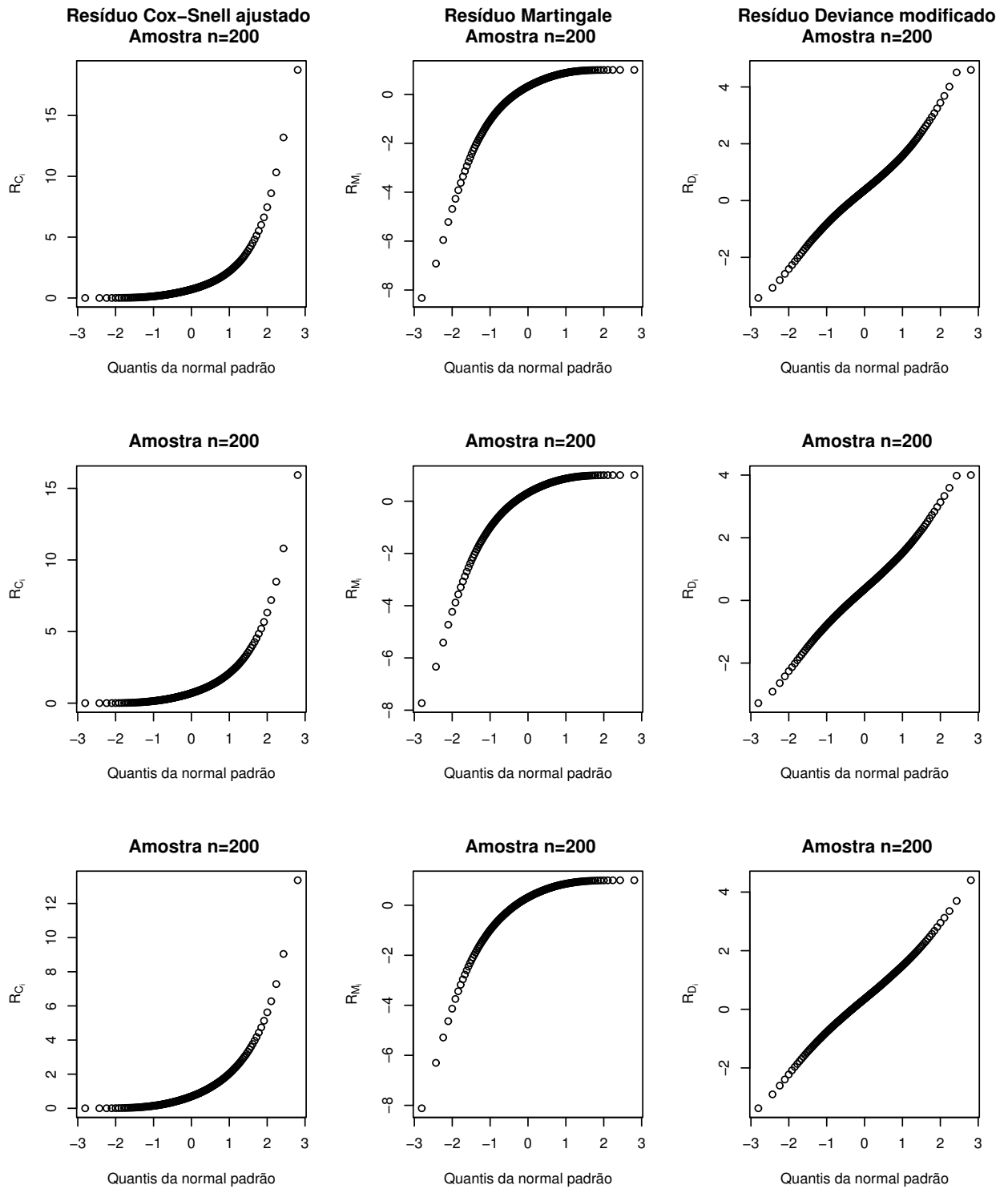


Figura 3.7: Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado, Martingale e componente de desvio modificado com $\nu = 3, 4$ e 5 nas três primeiras linhas, respectivamente e 0% de censura à direita sendo $\alpha = 1$.

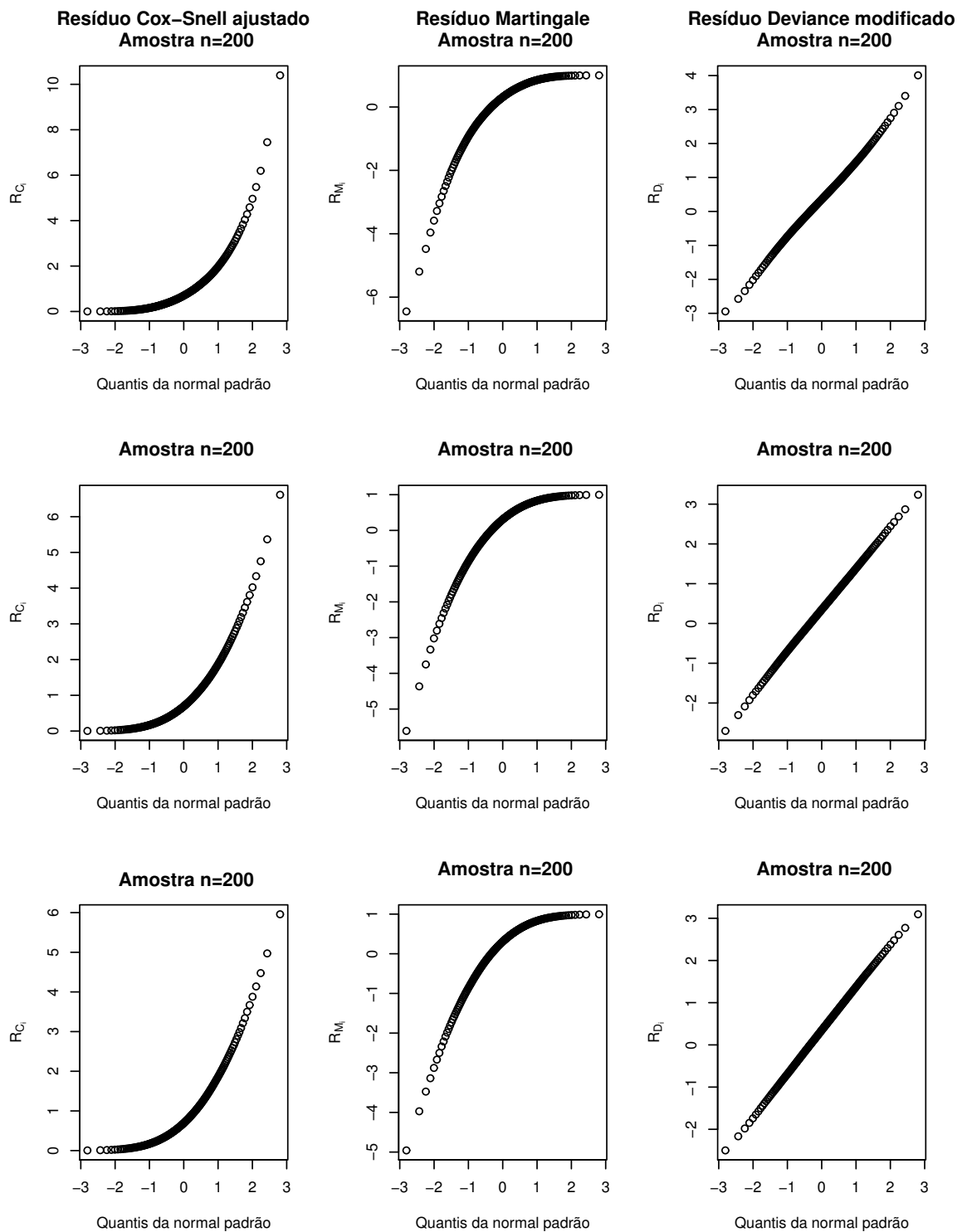


Figura 3.8: Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado, Martingale e componente de desvio modificado com $\nu = 7, 30$ e 100 nas três primeiras linhas, respectivamente e 0% de censura à direita sendo $\alpha = 1$.

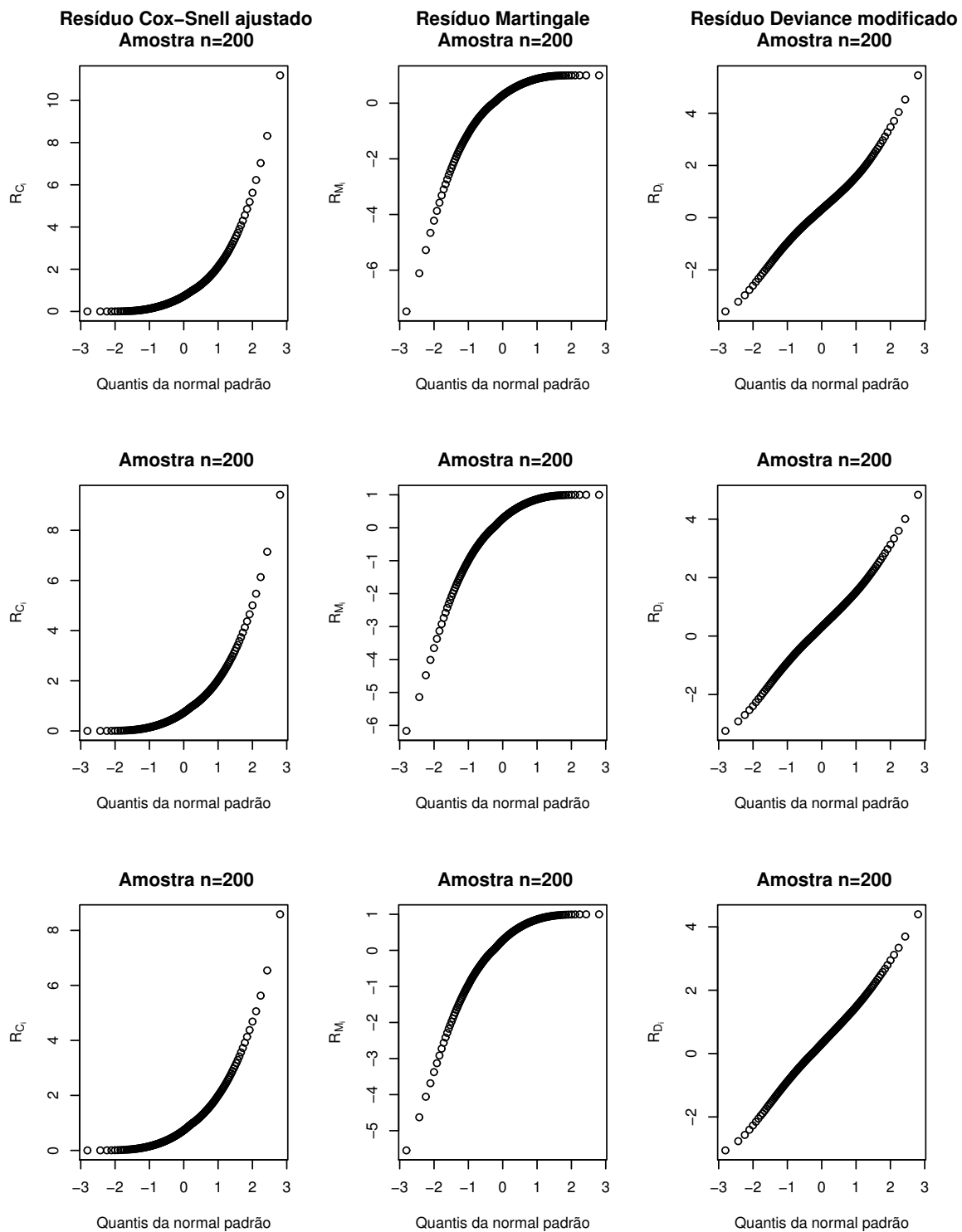


Figura 3.9: Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado, Martingale e componente de desvio modificado com $\nu = 3, 4$ e 5 nas três primeiras linhas, respectivamente e 10% de censura à direita sendo $\alpha = 1$.

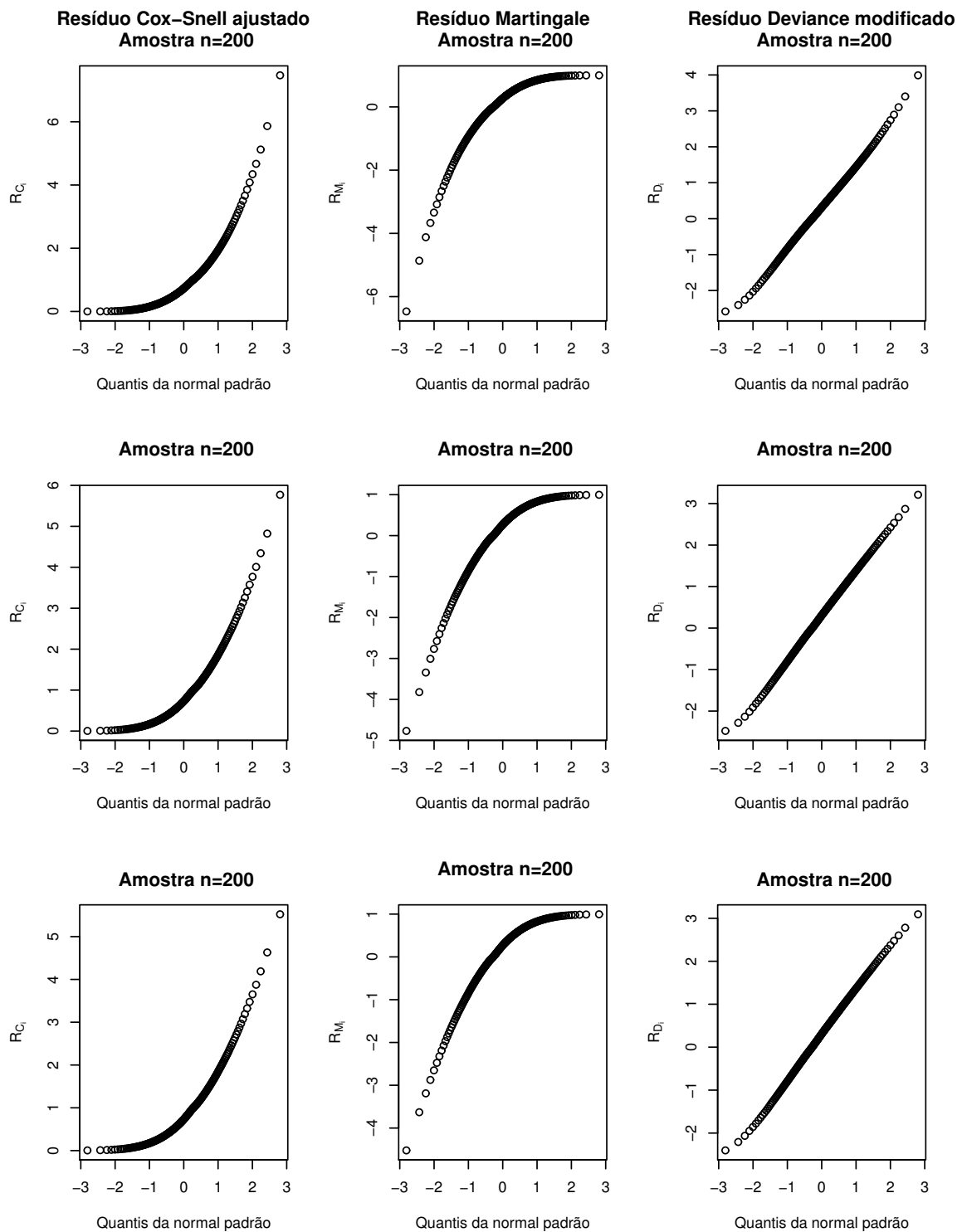


Figura 3.10: Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado, Martingale e componente de desvio modificado com $\nu = 7, 30$ e 100 nas três primeiras linhas, respectivamente e 10% de censura à direita sendo $\alpha = 1$.

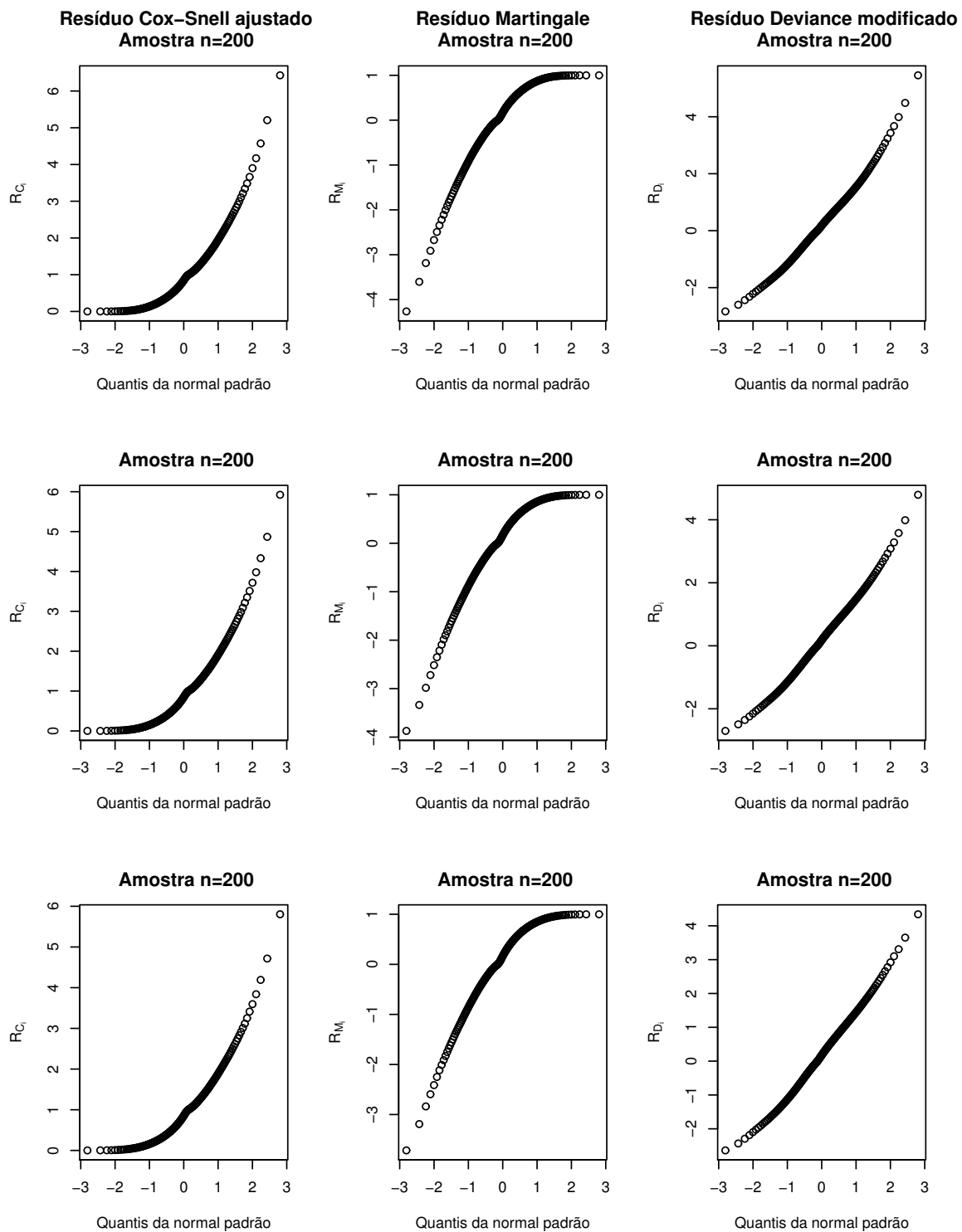


Figura 3.11: Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado, Martingale e componente de desvio modificado com $\nu = 3, 4$ e 5 nas três primeiras linhas, respectivamente e 30% de censura à direita sendo $\alpha = 1$.

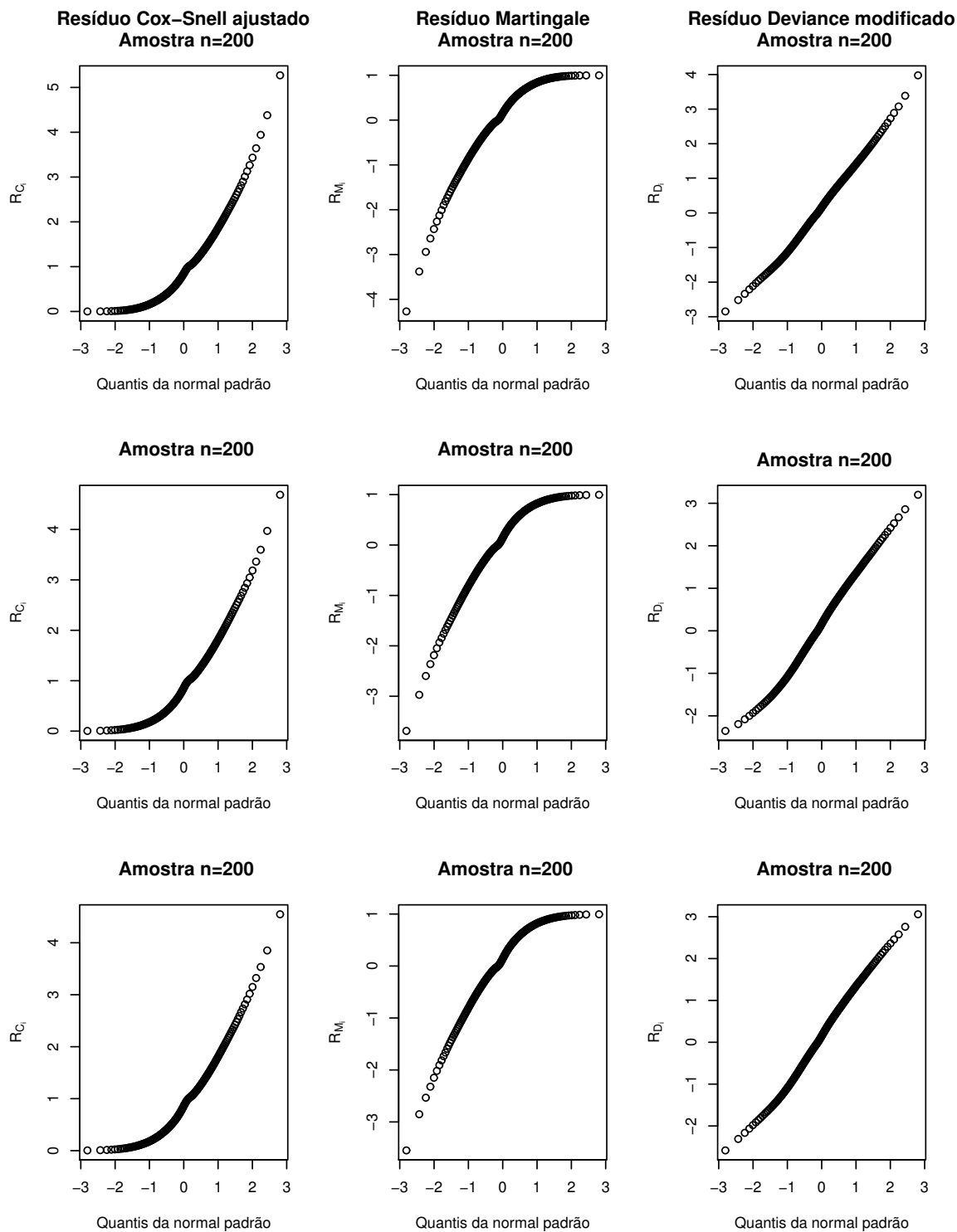


Figura 3.12: Gráficos normais de probabilidade para os resíduos Cox-Snell ajustado, Martingale e componente de desvio modificado com $\nu = 7, 30$ e 100 nas três primeiras linhas, respectivamente e 30% de censura à direita sendo $\alpha = 1$.

3.2.3 Influência local

O objetivo do estudo da influência local é buscar a existência de observações (pontos) que sob pequenas modificações no modelo causam variações desproporcionais, ou seja, pontos que possivelmente estejam fazendo com que o modelo esteja sendo mal ajustado. Existem algumas técnicas desenvolvidas nesse intuito, entre elas, a mais conhecida é a deleção de casos introduzida por Cook (1977) que avalia o impacto da retirada de uma observação ou de um conjunto de observações. Cook (1977) introduziu essa técnica inicialmente em modelos normais lineares e depois a mesma foi adaptada para outras classes de modelos. Moolgavkar et al. (1984), por exemplo, estenderam para modelos de regressão não-linear com aplicações em estudos emparelhados enquanto Paula e Peres (1988), discutiram a deleção de pontos em modelos lineares generalizados com parâmetros restritos na forma $C\beta \geq \mathbf{0}$.

O estudo inicial da influência local baseou-se na análise do afastamento pela verossimilhança $LD(\omega) = 2\{l(\hat{\theta}) - l(\hat{\theta}_\omega)\}$, em que $\hat{\theta}_\omega$ denota o estimador de máxima verossimilhança sob o modelo perturbado e $\omega = (\omega_1, \dots, \omega_n)^\top$ é o vetor de perturbações aplicados ao modelo e/ou aos dados. Assim, se a distância entre $\hat{\theta}$ e $\hat{\theta}_\omega$ permanecer “pequena” quando ω varia em $\Omega \subseteq \mathbb{R}^q$, com $q \leq n$, temos a indicação de que o modelo ajustado é estável no que diz respeito ao esquema de perturbação utilizado. Cook (1986) propôs uma abordagem mais moderna que se baseia na análise das curvaturas das seções normais da superfície,

$$\alpha(\omega) = \begin{pmatrix} \omega \\ LD(\omega) \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_q \\ LD(\omega) \end{pmatrix}, \quad \omega \in \Omega,$$

denominada de *gráfico de influência*, numa vizinhança de um ponto ω_0 , chamado de vetor de não perturbação, que satisfaz $l(\hat{\theta}) = l(\hat{\theta}_{\omega_0})$.

Para isso, Cook (1986) considerou uma direção arbitrária em \mathbb{R}^q a qual representou por um vetor unitário \mathbf{d} , e uma linha reta em Ω , na direção de \mathbf{d} , ou seja,

$$\omega(a) = \omega_0 + a\mathbf{d}, \quad \text{com } a \in \mathbb{R},$$

em seguida considerou o gráfico de $LD\{\omega(a)\}$ contra a , com $a \in \mathbb{R}$. Esse gráfico é

denominado de *linha projetada* e cada linha projetada pode ser caracterizada pela curvatura normal $C_l(\hat{\boldsymbol{\theta}})$ em torno de $a = 0$. (para mais detalhes ver Souza, 1999).

A sugestão de Cook (1986) é considerar a direção \mathbf{l}_{max} correspondente à maior curvatura $C_{l_{max}}$. O gráfico de \mathbf{l}_{max} contra o índice das observações pode mostrar os elementos que, sob pequenas perturbações, exercem notável influência sobre $LD(\boldsymbol{\omega})$. Cook (1986) mostrou que a curvatura normal pode ser expressa na forma geral como

$$C_l(\boldsymbol{\theta}) = 2|\mathbf{l}^\top \boldsymbol{\Delta}^\top \ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1} \boldsymbol{\Delta} \mathbf{l}| \quad (3.7)$$

em que $\boldsymbol{\Delta}$ é uma matriz $(p+q) \times n$ que depende do esquema de perturbação usado, cujos elementos são $\Delta_{ij} = \partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial \theta_j \partial \omega_i$, $i = 1, \dots, p+q$ e $j = 1, \dots, n$, com todas as quantidades sendo avaliadas em $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ e $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. Assim, $C_{l_{max}}$ é o maior autovalor da matriz

$$\mathbf{B} = -\boldsymbol{\Delta}^\top \ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1} \boldsymbol{\Delta},$$

e \mathbf{l}_{max} o correspondente autovetor. Quando o interesse é avaliar a influência parcial em subconjunto de $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)$, $\boldsymbol{\theta}_1$ por exemplo, temos que a curvatura normal na direção do vetor \mathbf{l} é dada por

$$C_l(\boldsymbol{\theta}_1) = 2|\mathbf{l}^\top \boldsymbol{\Delta}^\top (\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1} - \mathbf{B}_1) \boldsymbol{\Delta} \mathbf{l}|,$$

com

$$\mathbf{B}_1 = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{\mathbf{L}}_{22}^{-1} \end{bmatrix}$$

e $\ddot{\mathbf{L}}_{22} = \{\partial l(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial \boldsymbol{\theta}_2^\top \partial \boldsymbol{\theta}_2\} |_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}}$.

O gráfico do autovetor associado ao maior autovalor da matriz $-\boldsymbol{\Delta}^\top (\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1} - \mathbf{B}_1) \boldsymbol{\Delta}$ contra o índice das observações podem revelar quais observações estão influenciando $\hat{\boldsymbol{\theta}}_1$. Analogamente, quando o interesse está em $\boldsymbol{\theta}_2$, então a curvatura normal na direção do vetor \mathbf{l} é dada por

$$C_l(\boldsymbol{\theta}_2) = 2|\mathbf{l}^\top \boldsymbol{\Delta}^\top (\ddot{\mathbf{L}}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1} - \mathbf{B}_2) \boldsymbol{\Delta} \mathbf{l}|,$$

com

$$\mathbf{B}_2 = \begin{bmatrix} \ddot{\mathbf{L}}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

e $\ddot{\mathbf{L}}_{11} = \{\partial l(\boldsymbol{\theta}|\boldsymbol{\omega}) / \partial \boldsymbol{\theta}_1^\top \partial \boldsymbol{\theta}_1\} |_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}}$.

A influência local das observações em $\hat{\boldsymbol{\theta}}_2$ pode ser avaliada considerando o gráfico l_{max} para a matriz $-\boldsymbol{\Delta}^\top (\ddot{\mathbf{L}}_{\theta\theta}^{-1} - \mathbf{B}_2) \boldsymbol{\Delta}$ contra o índice das observações. Lesaffre e Verbeke (1998) sugerem considerar também a curvatura na direção do i -ésimo vetor da base canônica de \mathbb{R}^{q+1} . Seguindo a sugestão dos mesmos temos que a Expressão (3.7) será dada por $C_i = 2|\boldsymbol{\Delta}_i^\top \ddot{\mathbf{L}}_{\theta\theta}^{-1} \boldsymbol{\Delta}_i|$ em que $\boldsymbol{\Delta}_i^\top$ denota a i -ésima linha da matriz $\boldsymbol{\Delta}$. Verbeke e Molenberghs (2000, Seção 11.3) propuseram considerar como influentes os casos em que

$$C_i \geq 2\bar{C},$$

em que $\bar{C} = \frac{1}{n} \sum_{i=1}^n C_i$.

Para o modelo de regressão log-BSG na de presença de censura intervalar e à direita (3.1) os elementos de $\ddot{\mathbf{L}}_{\theta\hat{\theta}}$ são $\ddot{\mathbf{L}}_{\hat{\beta}\hat{\beta}} = \mathbf{X}^\top \hat{\mathbf{D}} \mathbf{X}$, $\ddot{\mathbf{L}}_{\hat{\beta}\hat{\alpha}} = \mathbf{X}^\top \hat{\mathbf{K}}$ e $\ddot{\mathbf{L}}_{\hat{\alpha}\hat{\alpha}} = \text{tr}(\hat{\mathbf{M}})$, dados em (3.6).

Neste trabalho, derivamos as curvaturas normais para o modelo de regressão linear log-BSG (3.1) com observações censuradas sob sete esquemas de perturbação. São eles: ponderação de casos, dois casos particulares da ponderação de casos, perturbação do limite inferior da variável resposta, perturbação do limite superior da variável resposta, perturbação da variável resposta intervalar e perturbação de uma covariável contínua, seguindo a mesma ideia de Hashimoto et al. (2010).

Cálculo das curvaturas

Doravante, para cada perturbação indicada iremos obter a matriz

$$\boldsymbol{\Delta} = (\boldsymbol{\Delta}_{ji})_{(p+2) \times n} = \left(\frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \theta_j \partial \omega_i} \right), \quad j = 1, \dots, p+2 \text{ e } i = 1, \dots, n$$

avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ de acordo com o modelo definido em (3.1) sendo $l(\boldsymbol{\theta}|\boldsymbol{\omega})$ o logaritmo da função de verossimilhança perturbada correspondente ao respectivo esquema considerado. Como no modelo (3.1) dependemos de um vetor $\boldsymbol{\theta} = (\alpha, \beta_0, \dots, \beta_p)^\top$ temos que a matriz $\boldsymbol{\Delta}$ assume a seguinte forma

$$\begin{pmatrix} \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \omega_1} & \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \omega_2} & \cdots & \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \omega_n} \\ \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_0 \partial \omega_1} & \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_0 \partial \omega_2} & \cdots & \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_0 \partial \omega_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_p \partial \omega_1} & \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_p \partial \omega_2} & \cdots & \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_p \partial \omega_n} \end{pmatrix}.$$

Assim, podemos simplificar a notação da mesma escrevendo-a da seguinte forma $\Delta = (\Delta_\alpha^\top, \Delta_\beta^\top)^\top$, em que Δ_α é uma matriz $1 \times n$ definida por

$$\begin{pmatrix} \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \omega_1} & \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \omega_2} & \cdots & \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \omega_n} \end{pmatrix}$$

e Δ_β é uma matriz $(p+1) \times n$ dada por

$$\begin{pmatrix} \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_0 \partial \omega_1} & \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_0 \partial \omega_2} & \cdots & \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_0 \partial \omega_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_p \partial \omega_1} & \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_p \partial \omega_2} & \cdots & \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_p \partial \omega_n} \end{pmatrix}.$$

Ponderação de casos

Nesse esquema temos o interesse em avaliar se as contribuições das observações com diferentes pesos afetarão as estimativas de máxima verossimilhança de $\boldsymbol{\theta}$. Esse esquema de perturbação é o mais utilizado na avaliação da influência sobre um modelo e, no nosso caso, a função logaritmo da verossimilhança assume a seguinte forma

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in I} \omega_i l_1(\alpha, \xi_{i2u}, \xi_{i2v}) + \sum_{i \in C} \omega_i l_2(\alpha, \xi_{i2u})$$

em que $l_1(\cdot)$ e $l_2(\cdot)$ são como definidos nas expressões (3.4) e (3.5). Onde cada $0 \leq \omega_i \leq 1$, nesse caso, temos que o vetor de não perturbação é $\boldsymbol{\omega}_0 = (1, \dots, 1)^\top$.

Os elementos da matriz $\Delta = (\Delta_\alpha^\top, \Delta_\beta^\top)^\top$ são expressos da seguinte maneira. Os elementos do vetor Δ_α são dados por

$$\Delta_i = \begin{cases} \frac{1}{\hat{\alpha}} \frac{f(\hat{\xi}_{i2u})\hat{\xi}_{i2u} - f(\hat{\xi}_{i2v})\hat{\xi}_{i2v}}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})}, & \text{se } i \in I \\ \frac{1}{\hat{\alpha}} h(\hat{\xi}_{i2u})\hat{\xi}_{i2u}, & \text{se } i \in C \end{cases}$$

e os elementos de Δ_β são dados por

$$\Delta_{ji} = \begin{cases} \frac{1}{2} x_{ij} \left(\frac{\hat{\xi}_{i1u} f(\hat{\xi}_{i2u}) - \hat{\xi}_{i1v} f(\hat{\xi}_{i2v})}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} \right), & \text{se } i \in I \\ \frac{1}{2} x_{ij} \hat{\xi}_{i1u} h(\hat{\xi}_{i2u}), & \text{se } i \in C \end{cases}.$$

Casos especiais da ponderação de casos

Caso 1 - Neste caso a função logaritmo da verossimilhança é dada por

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in I} \omega_i l_1(\alpha, \xi_{i2u}, \xi_{i2v}) + \sum_{i \in C} l_2(\alpha, \xi_{i2u})$$

em que $l_1(\cdot)$ e $l_2(\cdot)$ são como definidos nas expressões (3.4) e (3.5). Onde cada $0 \leq \omega_i \leq 1$, nesse caso, temos que o vetor de não perturbação é $\boldsymbol{\omega}_0 = (1, \dots, 1)^\top$.

Os elementos da matriz $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\alpha^\top, \boldsymbol{\Delta}_\beta^\top)^\top$ são expressos da seguinte maneira. Os elementos do vetor $\boldsymbol{\Delta}_\alpha$ são dados por

$$\Delta_i = \begin{cases} \frac{1}{\hat{\alpha}} \frac{f(\hat{\xi}_{i2u})\hat{\xi}_{i2u} - f(\hat{\xi}_{i2v})\hat{\xi}_{i2v}}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})}, & \text{se } i \in I \\ 0, & \text{se } i \in C \end{cases}$$

e os elementos de $\boldsymbol{\Delta}_\beta$ são dados por

$$\Delta_{ji} = \begin{cases} \frac{1}{2} x_{ij} \left(\frac{\hat{\xi}_{i1u} f(\hat{\xi}_{i2u}) - \hat{\xi}_{i1v} f(\hat{\xi}_{i2v})}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} \right), & \text{se } i \in I \\ 0, & \text{se } i \in C \end{cases}.$$

Caso 2 - Aqui, consideramos a função de log-verossimilhança como sendo

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in I} l_1(\alpha, \xi_{i2u}, \xi_{i2v}) + \sum_{i \in C} \omega_i l_2(\alpha, \xi_{i2u})$$

em que $l_1(\cdot)$ e $l_2(\cdot)$ são como definidos nas expressões (3.4) e (3.5). Onde cada $0 \leq \omega_i \leq 1$, nesse caso, temos que o vetor de não perturbação também é $\boldsymbol{\omega}_0 = (1, \dots, 1)^\top$.

Os elementos da matriz $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\alpha^\top, \boldsymbol{\Delta}_\beta^\top)^\top$ são expressos da seguinte maneira. Os elementos do vetor $\boldsymbol{\Delta}_\alpha$ são dados por

$$\Delta_i = \begin{cases} 0, & \text{se } i \in I \\ \frac{1}{\hat{\alpha}} h(\hat{\xi}_{i2u})\hat{\xi}_{i2u}, & \text{se } i \in C \end{cases}$$

e os elementos de $\boldsymbol{\Delta}_\beta$ são dados por

$$\Delta_{ji} = \begin{cases} 0, & \text{se } i \in I \\ \frac{1}{2} x_{ij} \hat{\xi}_{i1u} h(\hat{\xi}_{i2u}), & \text{se } i \in C \end{cases}.$$

Perturbação do limite inferior da variável resposta $\log(u_i)$

Neste caso, o interesse em fazer essa perturbação é analisar a sensibilidade do modelo quando o limite inferior ($\log(u_i)$) da variável resposta é submetida a uma perturbação aditiva $u_{iw} = \log(u_i) + \omega_i S_u$, $\omega_i \in \mathbb{R}$, em que S_u é um fator de escala que pode ser, por exemplo, a estimativa do desvio padrão da variável aleatória $\log(U)$, ver Hashimoto et al. (2010).

Nessa situação, o logaritmo da função de verossimilhança perturbada é da forma

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in I} l_1(\alpha, \xi_{i2u}^*, \xi_{i2v}) + \sum_{i \in C} l_2(\alpha, \xi_{i2u}^*)$$

em que $\xi_{i1u}^* = \frac{2}{\alpha} \cosh\left(\frac{\log(u_i) + w_i S_u - \mu_i}{2}\right)$, $\xi_{i2u}^* = \frac{2}{\alpha} \sinh\left(\frac{\log(u_i) + w_i S_u - \mu_i}{2}\right)$, $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ é o vetor de não perturbação, $l_1(\cdot)$ e $l_2(\cdot)$ são como definidos nas expressões (3.4) e (3.5).

Os elementos da matriz $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\alpha^\top, \boldsymbol{\Delta}_\beta^\top)^\top$ são expressos com os elementos do vetor $\boldsymbol{\Delta}_\alpha$ dados por

$$\Delta_i = \begin{cases} \frac{S_u \hat{\xi}_{i1u}}{2\hat{\alpha}} \left[\frac{f'(\hat{\xi}_{i2u})\hat{\xi}_{i2u} + f(\hat{\xi}_{i2u})}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} - \frac{f(\hat{\xi}_{i2u})(f(\hat{\xi}_{i2v})\hat{\xi}_{i2v} - f(\hat{\xi}_{i2u})\hat{\xi}_{i2u})}{(F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u}))^2} \right], & \text{se } i \in I \\ \frac{S_u \hat{\xi}_{i1u}}{2\hat{\alpha}} \left[\frac{f'(\hat{\xi}_{i2u})\hat{\xi}_{i2u} + f(\hat{\xi}_{i2u})}{F(-\hat{\xi}_{i2u})} + h^2(\hat{\xi}_{i2u})\hat{\xi}_{i2u} \right], & \text{se } i \in C \end{cases}.$$

Notemos agora que,

$$\begin{aligned} \frac{\partial l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_j} &= \sum_{i \in I} \frac{f(\xi_{i2v}) \left(-\frac{x_{ij}}{2}\right) \xi_{i1v} - f(\xi_{i2u}^*) \left(-\frac{x_{ij}}{2}\right) \xi_{i1u}^*}{F(\xi_{i2v}) - F(\xi_{i2u}^*)} \\ &\quad - \sum_{i \in C} \frac{f(\xi_{i2u}^*) \left(-\frac{x_{ij}}{2}\right) \xi_{i1u}^*}{F(-\xi_{i2u}^*)} \\ &= \sum_{i \in I} \frac{x_{ij}}{2} \left(\frac{f(\xi_{i2u}^*) \xi_{i1u}^* - f(\xi_{i2v}) \xi_{i1v}}{F(\xi_{i2v}) - F(\xi_{i2u}^*)} \right) + \sum_{i \in C} \frac{x_{ij}}{2} h(\xi_{i2u}^*) \xi_{i1u}^*, \\ \frac{\partial (f(\xi_{i2u}^*) \xi_{i1u}^*)}{\partial \omega_i} &= \frac{S_u}{2} (f'(\xi_{i2u}^*) (\xi_{i1u}^*)^2 + f(\xi_{i2u}^*) \xi_{i2u}^*), \end{aligned} \quad (3.8)$$

$$\frac{\partial (h(\xi_{i2u}^*) \xi_{i1u}^*)}{\partial \omega_i} = \frac{S_u (\xi_{i1u}^*)^2}{2} \left(\frac{f'(\xi_{i2u}^*) F(-\xi_{i2u}^*) + f^2(\xi_{i2u}^*)}{F^2(-\xi_{i2u}^*)} \right) + \frac{S_u}{2} \xi_{i2u}^* h(\xi_{i2u}^*), \quad (3.9)$$

e

$$\frac{\partial F(-\xi_{i2u}^*)}{\partial \omega_i} = -f(\xi_{i2u}^*)\xi_{i1u}^* \frac{S_u}{2}. \quad (3.10)$$

Usando as relações de (3.8) – (3.10), obtemos que os elementos da matriz Δ_β são dados por

$$\Delta_{ji} = \begin{cases} \frac{x_{ij}S_u}{4} \left(\frac{f'(\hat{\xi}_{i2u})(\hat{\xi}_{i1u})^2 + f(\hat{\xi}_{i2u})\hat{\xi}_{i2u}}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} + \frac{f(\hat{\xi}_{i2u})\hat{\xi}_{i1u}(f(\hat{\xi}_{i2u})\hat{\xi}_{i1u} - f(\hat{\xi}_{i2v})\hat{\xi}_{i1v})}{(F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u}))^2} \right), \\ \text{se } i \in I \\ \\ \frac{x_{ij}S_u}{4} \left[(\hat{\xi}_{i1u})^2 \left(\frac{f'(\hat{\xi}_{i2u})F(-\hat{\xi}_{i2u}) + f^2(\hat{\xi}_{i2u})}{F^2(-\hat{\xi}_{i2u})} \right) + \hat{\xi}_{i2u}h(\hat{\xi}_{i2u}) \right], \\ \text{se } i \in C \end{cases}$$

Perturbação do limite superior da variável resposta $\log(v_i)$

Similarmente ao caso anterior faremos agora uma perturbação no limite superior, $\log(v_i)$, da variável resposta submetendo-a a seguinte perturbação aditiva $v_{iw} = \log(v_i) + \omega_i S_v$, $\omega_i \in \mathbb{R}$, em que S_v é um fator de escala que pode ser, por exemplo, a estimativa do desvio padrão de $\log(V)$.

O logaritmo da função de verossimilhança perturbada para esse caso é dado por

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in I} l_1(\alpha, \xi_{i2u}, \xi_{i2v}^*) + \sum_{i \in C} l_2(\alpha, \xi_{i2u})$$

em que $\xi_{i1v}^* = \frac{2}{\alpha} \cosh\left(\frac{\log(v_i) + w_i S_v - \mu_i}{2}\right)$, $\xi_{i2v}^* = \frac{2}{\alpha} \sinh\left(\frac{\log(v_i) + w_i S_v - \mu_i}{2}\right)$, $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ é o vetor de não perturbação, $l_1(\cdot)$ e $l_2(\cdot)$ são como definidos nas expressões (3.4) e (3.5).

Os elementos da matriz $\Delta = (\Delta_\alpha^\top, \Delta_\beta^\top)^\top$ são expressos da seguinte maneira. Os elementos do vetor Δ_α são dados por

$$\Delta_i = \begin{cases} -\frac{S_v}{2\hat{\alpha}} \left[\frac{f'(\hat{\xi}_{i2v})\hat{\xi}_{i2v}\hat{\xi}_{i1v} + \hat{\xi}_{i1v}f(\hat{\xi}_{i2v})}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} - \frac{[f(\hat{\xi}_{i2v})\hat{\xi}_{i2v} - f(\hat{\xi}_{i2u})\hat{\xi}_{i2u}]f(\hat{\xi}_{i2v})\hat{\xi}_{i1v}}{[F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})]^2} \right], \\ \text{se } i \in I \\ \\ 0, \\ \text{se } i \in C \end{cases}$$

e os elementos de Δ_β são dados por

$$\Delta_{ji} = \begin{cases} -\frac{x_{ij}S_v}{4} \left[\frac{f'(\hat{\xi}_{i2v})(\hat{\xi}_{i1v})^2 + \hat{\xi}_{i2v}f(\hat{\xi}_{i2v})}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} - \frac{f(\hat{\xi}_{i2v})\hat{\xi}_{i1v}(f(\hat{\xi}_{i2v})\hat{\xi}_{i1v} - f(\hat{\xi}_{i2u})\hat{\xi}_{i1u})}{[F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})]^2} \right], \\ \text{se } i \in I \\ 0, \\ \text{se } i \in C \end{cases}$$

Perturbação da variável resposta intervalar $(\log(u_i), \log(v_i))$

Com o intuito de analisarmos a sensibilidade do modelo quando cada intervalo $(\log(u_i), \log(v_i))$ que contém a variável resposta y_i é submetido simultaneamente aos dois tipos de perturbações expostos anteriormente (perturbação do limite inferior e do limite superior da variável resposta), temos que o logaritmo da função de verossimilhança assume a seguinte forma

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in I} l_1(\alpha, \xi_{i2u}^*, \xi_{i2v}^*) + \sum_{i \in C} l_2(\alpha, \xi_{i2u}^*),$$

em que ξ_{i2u}^* e ξ_{i2v}^* são como definidos anteriormente, $\omega_i \in \mathbb{R}$, $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ é o vetor de não perturbação, $l_1(\cdot)$ e $l_2(\cdot)$ são como definidos nas expressões (3.4) e (3.5), e S_u e S_v são fatores de escala que podem corresponder, respectivamente, as estimativas do desvio padrão das variáveis aleatórias $\log(U)$ e $\log(V)$.

Para determinarmos os elementos da matriz $\Delta = (\Delta_\alpha^\top, \Delta_\beta^\top)^\top$ observemos primeiramente que

$$\frac{\partial l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_i} = \begin{cases} \frac{f(\xi_{i2v}^*)\xi_{i1v}^* \left(\frac{S_v}{2}\right) - f(\xi_{i2u}^*)\xi_{i1u}^* \left(\frac{S_u}{2}\right)}{F(\xi_{i2v}^*) - F(\xi_{i2u}^*)}, & \text{se } i \in I \\ \frac{f(\xi_{i2u}^*)\xi_{i1u}^* \left(\frac{S_u}{2}\right)}{F(-\xi_{i2u}^*)}, & \text{se } i \in C \end{cases}$$

Em seguida, consideremos as relações

$$\frac{\partial F(\xi_{i2v}^*)}{\partial \alpha} = -\frac{1}{\alpha} f(\xi_{i2v}^*)\xi_{i2v}^*, \quad (3.11)$$

$$\frac{\partial (f(\xi_{i2v}^*)\xi_{i1v}^*)}{\partial \alpha} = -\frac{1}{\alpha} (f'(\xi_{i2v}^*)\xi_{i2v}^*\xi_{i1v}^* + \xi_{i1v}^*f(\xi_{i2v}^*)) = -\frac{1}{\alpha}q_{iv}^* \quad (3.12)$$

analogamente,

$$\frac{\partial (f(\xi_{i2u}^*)\xi_{i1u}^*)}{\partial \alpha} = -\frac{1}{\alpha}q_{iu}^*, \quad (3.13)$$

em que $q_{iu}^* = f'(\xi_{i2u}^*)\xi_{i2u}^*\xi_{i1u}^* + \xi_{i1u}^*f(\xi_{i2u}^*)$ e $q_{iv}^* = f'(\xi_{i2v}^*)\xi_{i2v}^*\xi_{i1v}^* + \xi_{i1v}^*f(\xi_{i2v}^*)$.

Utilizando as relações (3.11)-(3.13) obtemos que, se $i \in I$,

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \omega_i} &= -\frac{1}{2\alpha} \left[\frac{q_{iv}^*S_v - q_{iu}^*S_u}{F(\xi_{i2v}^*) - F(\xi_{i2u}^*)} \right] \\ &+ \frac{1}{2\alpha} \left[\frac{[f(\xi_{i2v}^*)\xi_{i2v}^* - f(\xi_{i2u}^*)\xi_{i2u}^*][f(\xi_{i2v}^*)\xi_{i1v}^*S_v - f(\xi_{i2u}^*)\xi_{i1u}^*S_u]}{[F(\xi_{i2v}^*) - F(\xi_{i2u}^*)]^2} \right]. \end{aligned}$$

Caso $i \in C$, temos

$$\frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \omega_i} = \frac{S_u}{2\alpha} \left[\frac{q_{iu}^*}{F(-\xi_{i2u}^*)} + \frac{f^2(\xi_{i2u}^*)\xi_{i1u}^*\xi_{i2v}^*}{F^2(-\xi_{i2u}^*)} \right],$$

onde obtemos que os elementos de $\boldsymbol{\Delta}_\alpha$ são dados por

$$\Delta_i = \begin{cases} -\frac{1}{2\hat{\alpha}} \left[\frac{\hat{q}_{iv}S_v - \hat{q}_{iu}S_u}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} \right] \\ + \frac{1}{2\hat{\alpha}} \left[\frac{[f(\hat{\xi}_{i2v})\hat{\xi}_{i2v} - f(\hat{\xi}_{i2u})\hat{\xi}_{i2u}][f(\hat{\xi}_{i2v})\hat{\xi}_{i1v}S_v - f(\hat{\xi}_{i2u})\hat{\xi}_{i1u}S_u]}{[F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})]^2} \right], & \text{se } i \in I \\ \frac{S_u}{2\hat{\alpha}} \left[\frac{\hat{q}_{iu}}{F(-\hat{\xi}_{i2u})} + \frac{f^2(\hat{\xi}_{i2u})\hat{\xi}_{i1u}\hat{\xi}_{i2v}}{F^2(-\hat{\xi}_{i2u})} \right], & \text{se } i \in C \end{cases}$$

Em seguida determinamos os elementos da matriz $\boldsymbol{\Delta}_\beta$ observando as seguintes relações

$$\frac{\partial (f(\xi_{i2v}^*)\xi_{i1v}^*)}{\partial \beta_j} = -\frac{x_{ij}}{2} (f'(\xi_{i2v}^*)(\xi_{i1v}^*)^2 + \xi_{i2v}^*f(\xi_{i2v}^*)) = -\frac{x_{ij}}{2}p_{iv}^*$$

em que, $p_{iv}^* = f'(\xi_{i2v}^*)(\xi_{i1v}^*)^2 + \xi_{i2v}^*f(\xi_{i2v}^*)$. Analogamente, temos que

$$\frac{\partial (f(\xi_{i2u}^*)\xi_{i1u}^*)}{\partial \beta_j} = -\frac{x_{ij}}{2}p_{iu}^*$$

em que, $p_{iu}^* = f'(\xi_{i2u}^*)(\xi_{i1u}^*)^2 + \xi_{i2u}^* f(\xi_{i2u}^*)$, onde obtemos que, se $i \in I$,

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_j \partial \omega_i} &= -\frac{x_{ij}}{4} \left[\frac{(p_{iv}^* - p_{iu}^*)}{F(\xi_{i2v}^*) - F(\xi_{i2u}^*)} \right] \\ &+ \frac{x_{ij}}{4} \left[\frac{[f(\xi_{i2v}^*)\xi_{i1v}^* - f(\xi_{i2u}^*)\xi_{i1u}^*][f(\xi_{i2v}^*)\xi_{i1v}^* S_v - f(\xi_{i2u}^*)\xi_{i1u}^* S_u]}{[F(\xi_{i2v}^*) - F(\xi_{i2u}^*)]^2} \right] \end{aligned}$$

e que caso $i \in C$, temos

$$\frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_j \partial \omega_i} = \frac{S_u x_{ij}}{4} \left[\frac{p_{iu}^*}{F(-\xi_{i2u}^*)} + \frac{f^2(\xi_{i2u}^*)(\xi_{i1u}^*)^2}{F^2(-\xi_{i2u}^*)} \right].$$

Assim, temos que os elementos da matriz $\Delta_{\boldsymbol{\beta}}$ são dados por

$$\Delta_{ji} = \begin{cases} -\frac{x_{ij}}{4} \left[\frac{(\hat{p}_{iv} S_v - \hat{p}_{iu} S_u)}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} \right] \\ + \frac{x_{ij}}{4} \left[\frac{[f(\hat{\xi}_{i2v})\hat{\xi}_{i1v} - f(\hat{\xi}_{i2u})\hat{\xi}_{i1u}][f(\hat{\xi}_{i2v})\hat{\xi}_{i1v} S_v - f(\hat{\xi}_{i2u})\hat{\xi}_{i1u} S_u]}{[F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})]^2} \right], & \text{se } i \in I \\ \frac{S_u x_{ij}}{4} \left[\frac{f'(\hat{\xi}_{i2u})(\hat{\xi}_{i1u})^2 + \hat{\xi}_{i2u} f(\hat{\xi}_{i2u})}{F(-\hat{\xi}_{i2u})} + \left(\frac{f(\hat{\xi}_{i2u})(\hat{\xi}_{i1u})}{F(-\hat{\xi}_{i2u})} \right)^2 \right], & \text{se } i \in C \end{cases}.$$

Perturbação de uma covariável contínua

O objetivo nesse tipo de perturbação é avaliar a sensibilidade do modelo sob pequenas perturbações em uma covariável contínua, denotada por X_t . Considere uma perturbação aditiva para a covariável, definida por $x_{itw} = x_{it} + \omega_i S_t$, $\omega \in \mathbb{R}$ em que S_t é o fator de escala que pode ser a estimativa do desvio padrão da variável X .

Nesse caso temos que o logaritmo da função de verossimilhança perturbada fica dado por

$$l(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i \in I} l_1(\alpha, \xi_{i2u}^{**}, \xi_{i2v}^{**}) + \sum_{i \in C} l_2(\alpha, \xi_{i2u}^{**})$$

em que

$$\begin{aligned} \xi_{i1u}^{**} &= \frac{2}{\alpha} \cosh \left(\frac{u_i - (\mathbf{x}_i^*)^\top \boldsymbol{\beta}}{2} \right), & \xi_{i2u}^{**} &= \frac{2}{\alpha} \sinh \left(\frac{u_i - (\mathbf{x}_i^*)^\top \boldsymbol{\beta}}{2} \right), \\ \xi_{i1v}^{**} &= \frac{2}{\alpha} \cosh \left(\frac{v_i - (\mathbf{x}_i^*)^\top \boldsymbol{\beta}}{2} \right), & \xi_{i2v}^{**} &= \frac{2}{\alpha} \sinh \left(\frac{v_i - (\mathbf{x}_i^*)^\top \boldsymbol{\beta}}{2} \right), \end{aligned}$$

onde $(\mathbf{x}_i^*)^\top \boldsymbol{\beta} = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_t x_{itw} + \cdots + \beta_p x_{ip}$, $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ é o vetor de não perturbação, $l_1(\cdot)$ e $l_2(\cdot)$ são como definidos nas expressões (3.4) e (3.5).

Para determinarmos os elementos de $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\alpha^\top, \boldsymbol{\Delta}_\beta^\top)^\top$ devemos observar primeiramente as seguintes relações.

$$\frac{\partial F(\xi_{i2v}^{**})}{\partial \omega_i} = -\frac{\beta_t S_t}{2} f(\xi_{i2v}^{**}) \xi_{i1v}^{**},$$

$$\frac{\partial F(\xi_{i2v}^{**})}{\partial \alpha} = -\frac{1}{\alpha} f(\xi_{i2v}^{**}) \xi_{i2v}^{**}.$$

Para $i \in I$,

$$\frac{\partial l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_i} = \frac{\beta_t S_t}{2} \frac{f(\xi_{i2u}^{**}) \xi_{i1u}^{**} - f(\xi_{i2v}^{**}) \xi_{i1v}^{**}}{F(\xi_{i2v}^{**}) - F(\xi_{i2u}^{**})}$$

e para $i \in C$,

$$\frac{\partial l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_i} = \frac{\beta_t S_t}{2} \frac{f(\xi_{i2u}^{**}) \xi_{i1u}^{**}}{F(-\xi_{i2u}^{**})},$$

$$\frac{\partial (f(\xi_{i2v}^{**}) \xi_{i1v}^{**})}{\partial \alpha} = -\frac{1}{\alpha} (f'(\xi_{i2v}^{**}) \xi_{i2v}^{**} \xi_{i1v}^{**} + \xi_{i1v}^{**} f(\xi_{i2v}^{**})) = -\frac{1}{\alpha} q_{iv}^{**} \quad (3.14)$$

em que $q_{iv} = f'(\xi_{i2v}^{**}) \xi_{i2v}^{**} \xi_{i1v}^{**} + \xi_{i1v}^{**} f(\xi_{i2v}^{**})$.

De modo análogo, obtemos também que,

$$\frac{\partial (f(\xi_{i2u}^{**}) \xi_{i1u}^{**})}{\partial \alpha} = -\frac{1}{\alpha} q_{iu}^{**}, \quad (3.15)$$

em que $q_{iu} = f'(\xi_{i2u}^{**}) \xi_{i2u}^{**} \xi_{i1u}^{**} + \xi_{i1u}^{**} f(\xi_{i2u}^{**})$.

Utilizando as relações (3.14) e (3.15) obtemos que, se $i \in I$,

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \omega_i} &= -\frac{\beta_t S_t}{2\alpha} \left[\frac{(q_{iu}^{**} - q_{iv}^{**})}{F(\xi_{i2v}^{**}) - F(\xi_{i2u}^{**})} \right] \\ &+ \frac{\beta_t S_t}{2\alpha} \left[\frac{[f(\xi_{i2v}^{**}) \xi_{i2v}^{**} - f(\xi_{i2u}^{**}) \xi_{i2u}^{**}][f(\xi_{i2u}^{**}) \xi_{i1u}^{**} - f(\xi_{i2v}^{**}) \xi_{i1v}^{**}]}{[F(\xi_{i2v}^{**}) - F(\xi_{i2u}^{**})]^2} \right]. \end{aligned}$$

Caso $i \in C$, temos

$$\frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \omega_i} = -\frac{\beta_t S_t}{2\alpha} \left[\frac{q_{iv}^{**}}{F(-\xi_{i2u}^{**})} + \frac{f^2(\xi_{i2u}^{**}) \xi_{i2u}^{**} \xi_{i1u}^{**}}{F^2(-\xi_{i2u}^{**})} \right],$$

onde obtemos que os elementos de Δ_α são dados por

$$\Delta_i = \begin{cases} -\frac{\hat{\beta}_t S_t}{2\hat{\alpha}} \left[\frac{(\hat{q}_{iu} - \hat{q}_{iv})}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} - \frac{[f(\hat{\xi}_{i2v})\hat{\xi}_{i2v} - f(\hat{\xi}_{i2u})\hat{\xi}_{i2u}][f(\hat{\xi}_{i2u})\hat{\xi}_{i1u} - f(\hat{\xi}_{i2v})\hat{\xi}_{i1v}]}{[F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})]^2} \right], & \text{se } i \in I \\ -\frac{\hat{\beta}_t S_t}{2\hat{\alpha}} \left[\frac{\hat{q}_{iu}}{F(-\hat{\xi}_{i2u})} + \frac{f^2(\hat{\xi}_{i2u})\hat{\xi}_{i2u}\hat{\xi}_{i1u}}{F^2(-\hat{\xi}_{i2u})} \right], & \text{se } i \in C \end{cases}$$

e os elementos da matriz Δ_β para $j \neq t$ são dados por

$$\Delta_{ji} = \begin{cases} -\frac{\hat{\beta}_t S_t x_{ij}}{4} \left[\frac{(\hat{p}_{iu} - \hat{p}_{iv})}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} + \frac{[f(\hat{\xi}_{i2v})\hat{\xi}_{i1v} - f(\hat{\xi}_{i2u})\hat{\xi}_{i1u}]^2}{[F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})]^2} \right], & \text{se } i \in I \\ -\frac{\hat{\beta}_t S_t x_{ij}}{4} \left[\frac{\hat{p}_{iu}}{F(-\hat{\xi}_{i2u})} + \frac{f^2(\hat{\xi}_{i2u})(\hat{\xi}_{i1u})^2}{F^2(-\hat{\xi}_{i2u})} \right], & \text{se } i \in C \end{cases}$$

enquanto os elementos da matriz Δ_β para $j = t$ são dados por

$$\Delta_{ji} = \begin{cases} \frac{S_t f(\hat{\xi}_{i2u})\hat{\xi}_{i1u} - f(\hat{\xi}_{i2v})\hat{\xi}_{i1v}}{2} \frac{1}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} - \frac{\hat{\beta}_t S_t x_{ij}}{4} \left[\frac{(\hat{p}_{iu} - \hat{p}_{iv})}{F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})} + \frac{[f(\hat{\xi}_{i2v})\hat{\xi}_{i1v} - f(\hat{\xi}_{i2u})\hat{\xi}_{i1u}]^2}{[F(\hat{\xi}_{i2v}) - F(\hat{\xi}_{i2u})]^2} \right], & \text{se } i \in I \\ \frac{S_t f(\hat{\xi}_{i2u})\hat{\xi}_{i1u}}{2} \frac{1}{F(-\hat{\xi}_{i2u})} - \frac{\hat{\beta}_t S_t x_{ij}}{4} \left[\frac{\hat{p}_{iu}}{F(-\hat{\xi}_{i2u})} + \frac{f^2(\hat{\xi}_{i2u})(\hat{\xi}_{i1u})^2}{F^2(-\hat{\xi}_{i2u})} \right], & \text{se } i \in C \end{cases}$$

$$\text{com, } \hat{p}_{iu} = f'(\hat{\xi}_{i2u}^{**})(\hat{\xi}_{i1u}^{**})^2 + \hat{\xi}_{i2u}^{**} f(\hat{\xi}_{i2u}^{**}).$$

Capítulo 4

Aplicação

No decorrer desse capítulo faremos uso da teoria estudada para modelar uma situação real de estudo.

4.1 Descrição dos dados

Consideremos o banco de dados de câncer de mama, apresentado na Tabela (4.1), extraído de Colosimo e Giolo (2006) e podendo ser obtido através do pacote “DPpackage”, do software R., de Jara et al. (2011) com o nome de *deterioration*. Este conjunto de dados considera a informação de tempo para deterioração estética da mama de mulheres com câncer de mama no primeiro estágio, que passaram por uma mastectomia e se submeteram a dois tipos de tratamentos, sendo eles: somente radioterapia ou radioterapia em conjunto com quimioterapia. O nosso interesse principal é o impacto estético dos dois tratamentos, em outras palavras, o tempo da primeira retração moderada ou severa de alguma mama, uma vez que ambos os tratamentos são considerados muito eficazes na prevenção da recorrência desse tipo de câncer na fase inicial. Os dados foram obtidos através de um estudo retrospectivo com 94 mulheres das quais 46 delas receberam somente a radioterapia enquanto as demais, 48, receberam a radioterapia juntamente com a quimioterapia. Cada mulher fez uma série de visitas a um médico, o qual diagnosticava se havia, ou não, ocorrido a retração da mama. Caso tivesse ocorrido a retração, o tempo era conhecido apenas por situar-se entre o tempo da visita atual e da visita imediatamente anterior, caracterizando então uma censura intervalar. Também era possível não ocorrer a retração em alguma mama de um paciente no de-

correr do estudo caracterizando, assim, a censura à direita. Os tempos de sobrevivência foram dispostos em meses. Colosimo e Giolo faz uma análise não paramétrica desse banco de dados e uma análise paramétrica usando os modelos logístico e gaussiano. Hashimoto et al. (2010) por sua vez faz uma análise paramétrica, utilizando o modelo de regressão Weibull-Exponenciada.

Tabela 4.1: Tempos (em meses) até a retração de uma das mamas para pacientes com câncer de mama de acordo com dois tratamentos

	(0, 7]; (0, 8]; (0, 5]; (4, 11]; (5, 12]; (5, 11]; (6, 10]; (7, 16]; (7, 14];
	(11, 15], (11, 18]; ≥ 15 ; ≥ 17 ; (17, 25]; (17, 25]; ≥ 18 ; (19, 35];
Radioterapia	(18, 26]; ≥ 22 ; ≥ 24 ; ≥ 24 ; (25, 37]; (26, 40]; (27, 34]; ≥ 32 ; ≥ 33 ;
	≥ 34 ; (36, 44]; (36, 48]; ≥ 36 ; ≥ 36 ; (37, 44]; ≥ 37 ; ≥ 37 ; ≥ 37 ;
	≥ 38 ; ≥ 40 ; ≥ 45 ; ≥ 46 ; ≥ 46 ; ≥ 46 ; ≥ 46 ; ≥ 46 ; ≥ 46 ; ≥ 46 ; ≥ 46 ;
Radioterapia	(0, 22]; (0, 5]; (4, 9]; (4, 8]; (5, 8]; (8, 12]; (8, 21]; (10, 35]; (10, 17];
	(11, 13]; ≥ 11 ; (11, 17]; ≥ 11 ; (11, 20]; (12, 20]; ≥ 13 ; (13, 39];
+	≥ 13 ; ≥ 13 ; (14, 17]; (14, 19]; (15, 22]; (16, 24]; (16, 20]; (16, 24];
Quimioterapia	(16, 60]; (17, 27]; (17, 23]; (17, 26]; (18, 25]; (18, 24]; (19, 32];
	≥ 21 ; (22, 32]; ≥ 23 ; (24, 31]; (24, 30]; (30, 34]; (30, 36]; ≥ 31 ;
	≥ 32 ; (32, 40]; ≥ 34 ; ≥ 34 ; ≥ 35 ; (35, 39]; (44, 48]; ≥ 48

O câncer de mama trata-se de uma doença crônica degenerativa, ou seja, a mesma se desenvolve de forma gradual. Considerando, então, uma pessoa portadora dessa doença é natural assumirmos que a mesma possui uma espécie de dano acumulado causado pela doença no decorrer do tempo, ou seja, temos um processo de fadiga. Assim, faremos uma análise paramétrica propondo dois modelos paramétricos à saber: modelo de regressão log-BS e modelo de regressão log-BS- t de student uma vez que ambos se mostram adequados para analisar tais dados. Faremos também uma análise não paramétrica do modelos propostos na qual estimaremos as curvas de sobrevivência para cada tratamento via algoritmo de Turnbull e em seguida comparamos as mesmas com as dos modelos paramétricos.

4.2 Modelo log-BS

Para ajustar os dados propostos, então, um modelo de regressão linear log-BS, dado por:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, \dots, 94 \quad (4.1)$$

em que $\epsilon_i \sim \text{SN}(\alpha, 0, 2)$ e a covariável x_i , $i = 1, \dots, 94$, assume 1 se o tratamento usado foi apenas radioterapia e 0 se o tratamento utilizado foi radioterapia mais quimioterapia.

Inicialmente vamos fazer uma análise não paramétrica das curvas de sobrevivência estimadas pelo algoritmo de Turnbull para os dois tratamentos assim como apresentado em Colosimo e Giolo (2006). Na Figura (4.1) vemos que no intervalo de tempo de 0 à 18 meses não há diferenças significativas entre os dois tratamentos. Porém, após 18 meses, notamos um rápido decaimento da função de sobrevivência para o tratamento de radioterapia mais quimioterapia indicando que nesse tratamento a retração da mama ocorre de forma mais rápida, sendo portanto um indício de que o tratamento apenas com radioterapia é mais eficaz a longo prazo, no sentido estético.

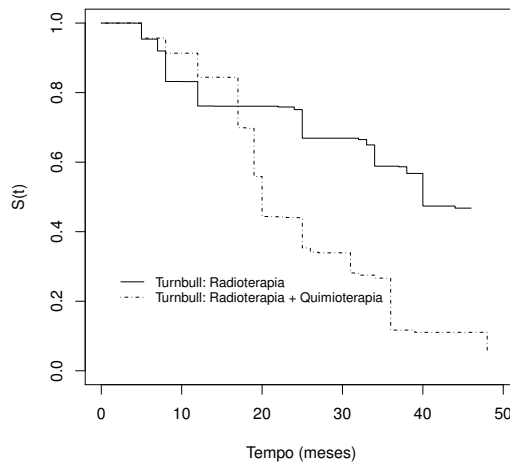


Figura 4.1: Sobrevivência estimada para o Modelo (4.1) via o algoritmo de Turnbull

Na Tabela (4.2) encontram-se as estimativas de máxima verossimilhança, o erro padrão, o p -valor e os intervalos de confiança para cada parâmetro.

Analisando a Tabela (4.2) notamos que fixado o nível de significância em 5%, β_1

Tabela 4.2: Estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-BS.

Parâmetro	Estimativa	Erro padrão	p -valor	Intervalo de confiança
α	0,9347	0,1043	0	(0,7303; 1,1391)
β_0	3,1234	0,1446	0	(2,8400; 3,4068)
β_1	0,3551	0,1994	0,0750	(-0,0357; 0,7459)

não é significativo, ou seja, parece não haver efeito de tratamento.

De acordo com a Figura (4.2) que apresenta a sobrevivência estimada via algoritmo de Turnbull e a sobrevivência ajustada para o modelo BS percebemos que a sobrevivência estimada pelo modelo BS segue a mesma tendência da estimada via algoritmo de Turnbull indicando que a suposição de que os erros assumem distribuição BS pode ser adequada.

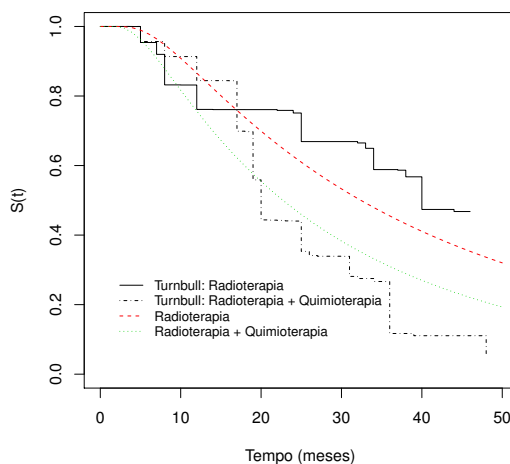


Figura 4.2: Função de sobrevivência empírica e ajustada para o modelo BS.

Com intuito de verificar se há observações extremas e detectar possíveis afastamentos das suposições feitas para o modelo, exibimos na Figura (4.3) o gráfico dos resíduos componente do desvio modificado contra o índice das observações. Como podemos observar destacam-se como possíveis pontos discrepantes as observações #1, #2, #3 e #48. Notemos que tais observações correspondem a pacientes em que a retração da

mama ocorreu até a primeira visita ao médico. Na Figura (4.4) apresentamos o gráfico normal de probabilidade para os resíduos componente do desvio modificado com envelope gerado (Atkinson, 1985). Percebemos, por esse gráfico, algumas ondulações no comportamento dos resíduos bem como alguns pontos fora do envelope sugerindo que talvez um modelo de caudas mais pesadas possa se adequar melhor aos dados, ou seja, o modelo BS não seria o mais adequado.

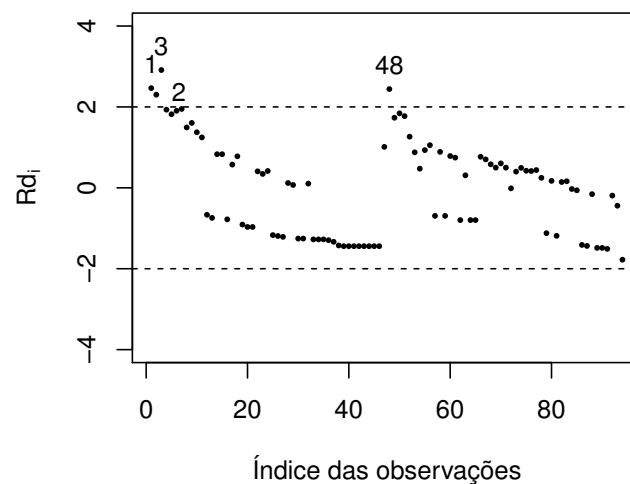


Figura 4.3: Gráfico do resíduo componente de desvio versus índices das observações para o modelo log-BS.

Fizemos uma análise de diagnóstico baseado em influência local com o objetivo de detectar pontos influentes, ou seja, aqueles com influência desproporcional nas estimativas dos coeficientes do modelo. Considerando o esquema de perturbação limite superior, notamos pelos gráficos apresentados na Figura (4.5) que as observações #1, #2, #3 e #48 destoam das demais sendo, portanto, possíveis observações influentes.

Uma vez que detectamos como possíveis pontos influentes as observações #1, #2, #3 e #48 iremos agora usar a deleção de casos, conforme apresentada na Subseção (3.2.3), e faremos uma análise do impacto dessas observações sob o ajuste do modelo. Lembremos que o impacto da retirada de cada observação sob o modelo ajustado é

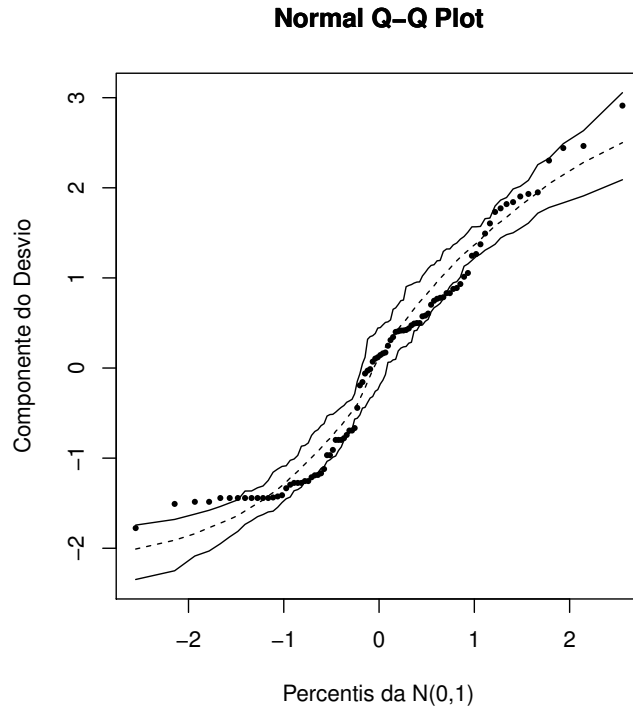


Figura 4.4: Gráfico de envelope dos resíduos componentes do desvio modificado para o modelo de regressão log-BS.

dado por

$$\mathbf{RC}_{\theta_j} = \left| \frac{\hat{\theta}_j - \hat{\theta}_{j^{(i)}}}{\hat{\theta}_j} \right| \times 100,$$

em que $\theta_{j^{(i)}}$ é o estimador de máxima verossimilhança sem a i -ésima observação. Retirando, então, a observação #3 a qual aparenta ser a mais influente e reestimando o modelo, obtemos a Tabela (4.3), em que por meio da mesma vemos que o coeficiente β_1 passou a ser significativo ao nível de 5% ocorrendo, portanto, mudança inferencial. Ou seja, a variável tratamento é importante para explicar o tempo até a retração da mama e devido a um ponto na análise inicial parecia não ser significativa.

Retirando então o conjunto de observações #1, #2, #3 e #48 consideradas influentes obtivemos a Tabela (4.4).

Assim, concluímos que o modelo não é adequado para analisar estes dados, pois o mesmo é sensível a observações atípicas.

Fizemos também uma análise da influência local para o esquema da ponderação de casos e para o primeiro caso especial da ponderação de casos em que obtemos as

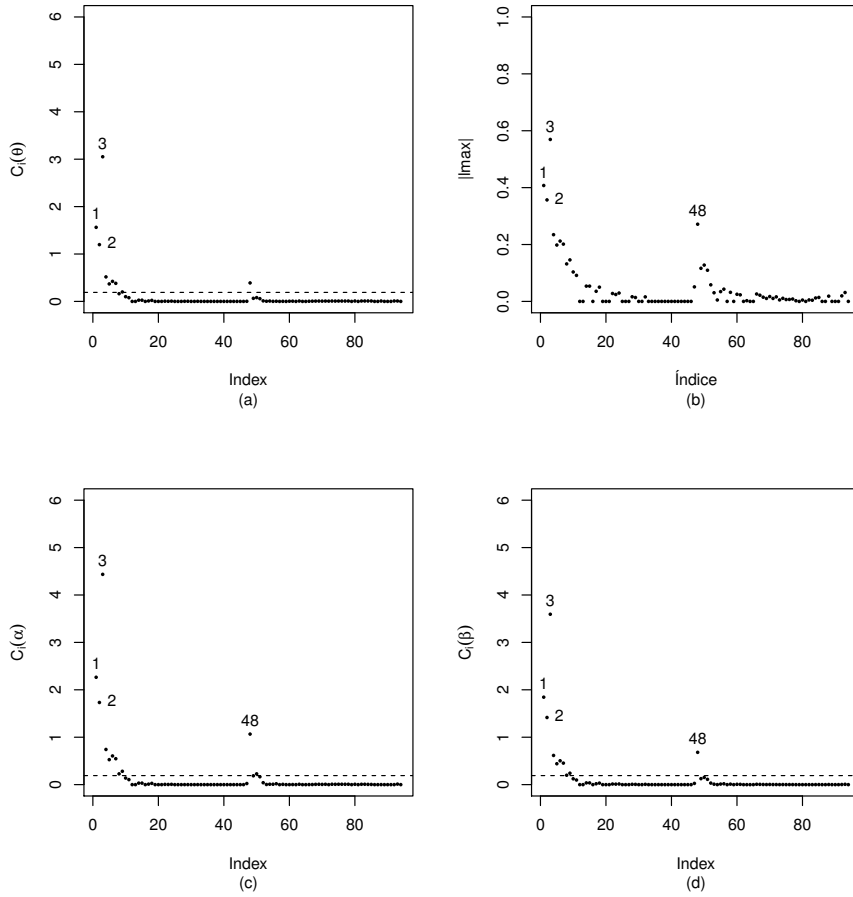


Figura 4.5: Gráficos sob o esquema de perturbação do limite superior: (a) $C_i(\theta)$ versus índices; (b) l_{max} versus índices; (c) $C_i(\alpha)$ versus índices; (d) $C_i(\beta)$ versus índices

mesmas observações citadas anteriormente como influentes, sendo assim suprimimos os gráficos apresentando apenas o da perturbação do limite superior uma vez que na mesma tais observações se destacaram mais. Quanto a perturbação sob o segundo caso especial da ponderação de casos e a perturbação do limite inferior, as mesmas não apresentaram observações atípicas.

4.3 Modelo log-BS t de student

Iremos propor, agora, um modelo de caudas mais pesadas para ajustar os dados de câncer de mama. Consideremos então um modelo log-BS- t de student com 3 graus de liberdade, de forma similar ao apresentado em (4.1) em que $\epsilon_i \sim \text{log-BSG}(\alpha, 0, 2; t_3)$.

Tabela 4.3: Estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-BS para os dados de câncer de mama menos a observação 3.

Parâmetro	Estimativa	Erro padrão	p -valor	Intervalo de confiança	Impacto
α	0,8785	0,0985	0	(0,6854; 1,0716)	6%
β_0	3,1101	0,1346	0	(2,8462; 3,3739)	0,42%
β_1	0,4282	0,1901	0,0250	(0,0539; 0,8025)	20,6%

Tabela 4.4: Estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-BS para os dados de câncer de mama menos as observações 1, 2, 3 e 48.

Parâmetro	Estimativa	Erro padrão	p -valor	Intervalo de confiança	Impacto
α	0,7577	0,0862	0	(0,5887; 0,9266)	19%
β_0	3,1426	0,1188	0	(2,9098; 3,3754)	0,6%
β_1	0,4710	0,1703	0,0057	(0,1371; 0,8048)	32,6%

Os graus de liberdade foram fixados como recomendam Lucas (1997), Barros et al. (2008) e Paula (2011).

Usando um programa feito na linguagem de programação matricial do `0x Doornik` (2011) obtivemos as estimativas dos parâmetros. Os resultados são apresentados na Tabela (4.5). Através da referida tabela notamos que todas as variáveis do nosso modelo são significativas a um nível de significância de 5%. Ou seja, a variável tratamento deve permanecer no modelo.

Tabela 4.5: Estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-BS-t com 3 graus de liberdade para os dados de câncer de mama.

Parâmetro	Estimativa	Erro padrão	p -valor	Intervalo de confiança
α	0,6859	0,0970	$1,5494 \times 10^{-12}$	(0,4957; 0,8760)
β_0	3,1175	0,1185	0	(2,8853; 3,3497)
β_1	0,5182	0,1864	0,0054	(0,1529; 0,8834)

De acordo com a Figura (4.6) que apresenta a sobrevivência estimada pelo algoritmo de Turnbull e a sobrevivência ajustada para o modelo log-BSG ajustado temos

que há indício de um bom ajuste do modelo sendo o mesmo melhor que o ajuste log-BS.

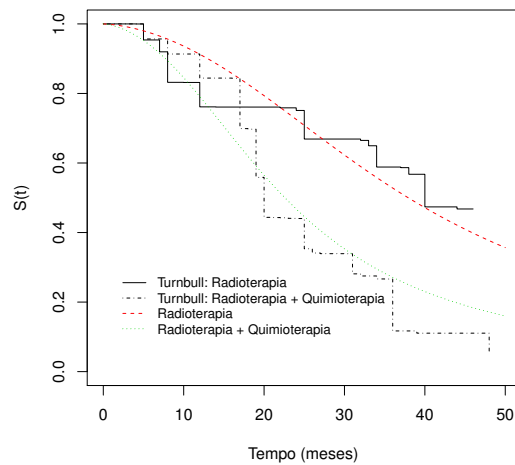


Figura 4.6: Função de sobrevivência empírica e ajustada para o modelo BSG.

No intuito de verificar se há observações extremas e detectar possíveis afastamentos das suposições feitas para o modelo, assim como fizemos na Seção (4.2) exibimos na Figura (4.7) o gráfico dos resíduos componente do desvio modificado contra o índice das observações. Notamos através da Figura (4.7) que também para o modelo log-BSG se destacaram como possíveis pontos discrepantes as observações #1, #2, #3 e #48.

Na Figura (4.8) temos o gráfico normal de probabilidade para os resíduos componente do desvio modificado com envelope gerado. Através do mesmo notamos que há um ajuste melhor dos dados para o modelo log-BSG do que para o modelo log-BS apresentado na Figura (4.4). Ou seja, não há indícios de afastamentos sérios da suposição de que os erros possuem distribuição log-BS- t de student.

Quando analisamos os dados com o modelo de regressão log-BS considerando o esquema de perturbação do limite superior tínhamos a existência de observações que destoavam das demais. Esse fato não ocorre quando consideramos o modelo log-BS- t , conforme podemos notar através da Figura (4.9). Analisando a influência local sob os demais esquemas de perturbações os resultados obtidos foram similares ao da análise feita com o modelo log-BS assim optamos por não apresentar os gráficos obtidos.

Detectamos as observações #1, #2, #3 e #48 como potencialmente influentes. Usando a deleção de casos para fazer uma análise do impacto dessas observações no

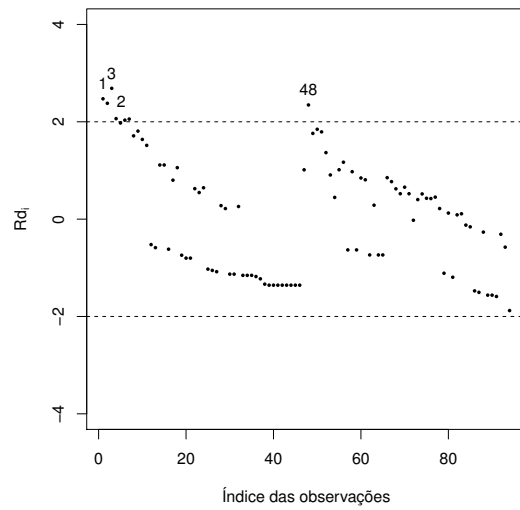


Figura 4.7: Gráfico dos resíduos componentes do desvio modificado contra os índices das observações para o modelo de regressão log-BSG.

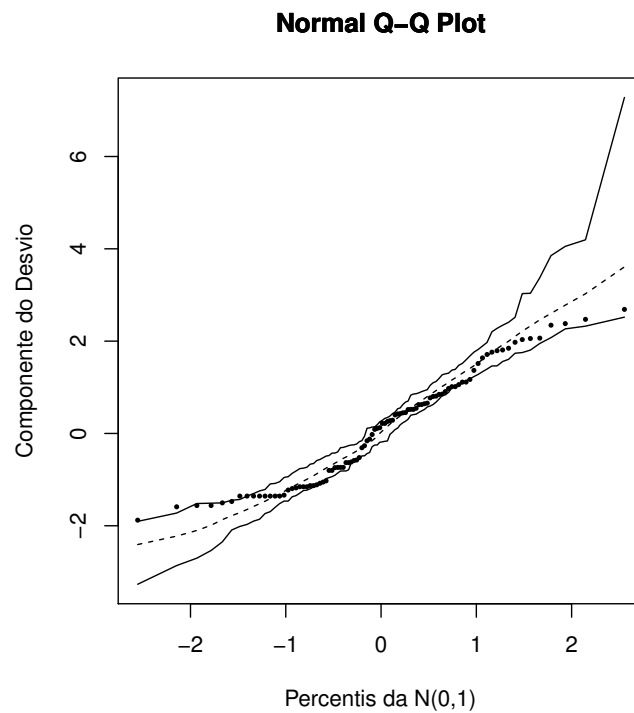


Figura 4.8: Gráfico de envelope dos resíduos componente do desvio modificado para o modelo de regressão log-BSG.

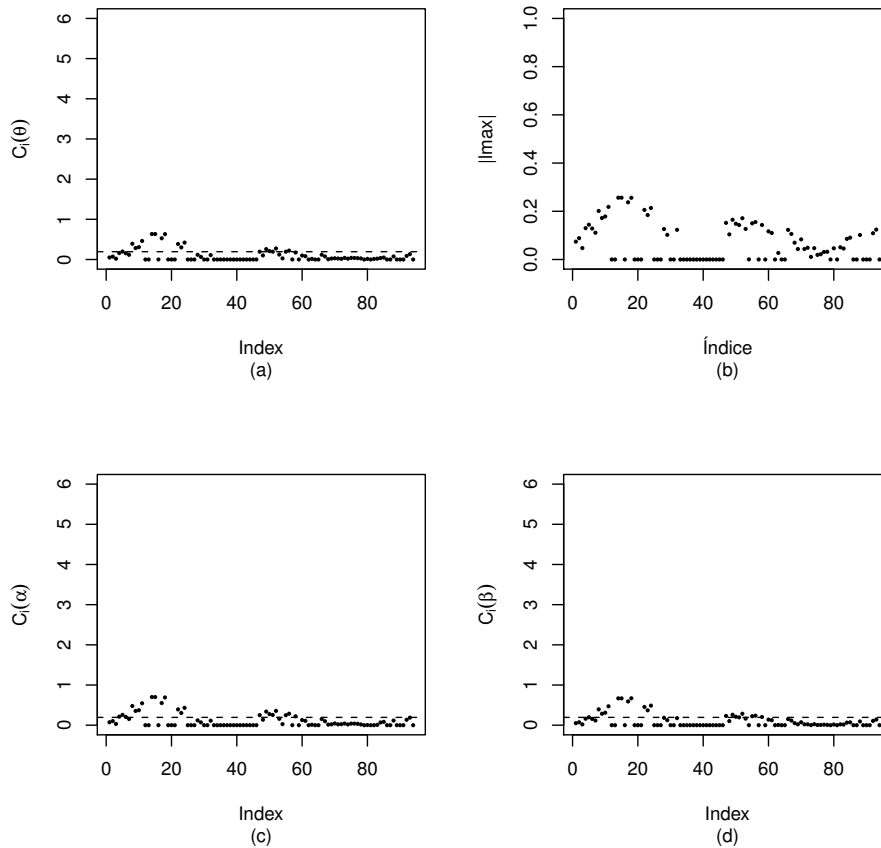


Figura 4.9: Gráficos sob o esquema de perturbação do limite superior: (a) $C_i(\theta)$ versus índices; (b) l_{max} versus índices; (c) $C_i(\alpha)$ versus índices; (d) $C_i(\beta)$ versus índices.

modelo retirando apenas a observação #3 e posteriormente as observações #1, #2, #3 e #48, obtemos respectivamente as Tabelas (4.6) e (4.7) apresentadas à seguir.

De acordo com a Tabela (4.6) percebemos que removendo a observação #3 não ocorreram mais mudanças inferenciais, indicando que o método de máxima verossimilhança para este modelo é menos sensível a observações atípicas. O mesmo ocorre quando retiramos o conjunto formado pelas observações #1, #2, #3 e #48 conforme podemos ver na Tabela (4.7).

Assim, o modelo final selecionado para ajustar os dados foi o modelo log-BS- t com 3 graus de liberdade. Portanto, o modelo ajustado é dado por

$$\hat{y}_i = 3.1175 + 0.51819x_i, \quad i = 1, \dots, 94 \quad (4.2)$$

em que a covariável x_i , $i = 1, \dots, 94$, assume 1 se o tratamento usado foi apenas

Tabela 4.6: Estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-BS-t com 3 graus de liberdade para os dados de câncer de mama menos a observação 3. Nesta tabela I.C denota intervalo de confiança.

Parâmetro	Estimativa	Erro padrão	p -valor	I.C	Impacto
α	0,6539	0,0905	$5,0271 \times 10^{-13}$	(0,4765; 0,8313)	4,66%
β_0	3,1148	0,1132	0	(2,8927; 3,3368)	0,0087%
β_1	0,5517	0,1812	0,0023	(0,1966; 0,9068)	6,46%

Tabela 4.7: Estimativas de máxima verossimilhança dos parâmetros do modelo de regressão log-BS-t com 3 graus de liberdade para os dados de câncer de mama menos as observações 1, 2, 3 e 48. Nesta tabela I.C denota intervalo de confiança.

Parâmetro	Estimativa	Erro padrão	p -valor	I.C	Impacto
α	0,5727	0,0793	$5,2758 \times 10^{-13}$	(0,4172; 0,7282)	16,5%
β_0	3,1287	0,1041	0	(2,9246; 3,3327)	0,359%
β_1	0,5908	0,1661	0,0004	(0,2651; 0,9164)	14%

radioterapia e 0 se o tratamento utilizado foi radioterapia mais quimioterapia. Onde concluímos que no tratamento de radioterapia a taxa do tempo de retração da mama aumenta aproximadamente 68%, uma vez que $\frac{\exp(3.1175 + 0.51819)}{\exp(3.1175)} \simeq 1,68$. Portanto o tratamento usando apenas radioterapia é a melhor opção para o paciente, no que diz respeito a estética.

Capítulo 5

Conclusões

Neste trabalho, propomos um modelo de regressão log-BSG para dados com censura intervalar. Fizemos uso do algoritmo Quase-Newton para obter as estimativas de máxima verossimilhança bem como os respectivos intervalos de confiança baseados na distribuição assintótica dos estimadores de máxima verossimilhança. Efetuamos estudos de simulações para os resíduos de Cox-Snell ajustado, Martingale e componente do desvio modificado, através de simulações de Monte Carlo, em que vimos que o último apresentou alta concordância com a distribuição normal padrão. Desenvolvemos a análise de influência local sob diversos esquemas de perturbação do modelo. Por fim, apresentamos uma aplicação em um conjunto de dados de pacientes com câncer de mama modelando o problema com os modelos de regressão log-BS e log-BS- t onde obtivemos as seguintes conclusões.

- No modelo log-BS a significância do parâmetro β_1 está sendo mascarada pela observação # 3 que corresponde ao paciente que teve o menor tempo até a primeira retração moderada ou severa de uma mama e recebeu apenas radioterapia.
- No modelo log-BS- t foi detectado o efeito do tratamento e não demonstrou-se sensível à observações candidatas à influentes, ou seja, o processo de estimação no modelo log-BS- t é mais robusto.
- No tratamento de radioterapia a taxa do tempo de retração da mama aumenta aproximadamente 68% fazendo com que o mesmo seja a melhor opção para um paciente.

Apêndice A

Cálculo das Segundas derivadas

Adotando as seguintes notações:

$$\begin{aligned}\xi_{i1u} &= \frac{2}{\alpha} \cosh\left(\frac{\log(u_i) - \mu_i}{2}\right), & \xi_{i1v} &= \frac{2}{\alpha} \cosh\left(\frac{\log(v_i) - \mu_i}{2}\right), \\ \xi_{i2u} &= \frac{2}{\alpha} \sinh\left(\frac{\log(u_i) - \mu_i}{2}\right), & \xi_{i2v} &= \frac{2}{\alpha} \sinh\left(\frac{\log(v_i) - \mu_i}{2}\right), \\ h(\xi_{i2u}) &= \frac{\phi(\xi_{i2u})}{\Phi(-\xi_{i2u})},\end{aligned}$$

obtemos as relações,

$$\begin{aligned}\frac{\partial \xi_{i1u}}{\partial \alpha} &= -\frac{1}{\alpha} \xi_{i1u}, & \frac{\partial \xi_{i2u}}{\partial \alpha} &= -\frac{1}{\alpha} \xi_{i2u}, & \frac{\partial \xi_{i1u}}{\partial \beta_j} &= -\frac{x_{ij}}{2} \xi_{i2u}, & \frac{\partial \xi_{i2u}}{\partial \beta_j} &= -\frac{x_{ij}}{2} \xi_{i1u}, \\ \frac{\partial \xi_{i1v}}{\partial \alpha} &= -\frac{1}{\alpha} \xi_{i1v}, & \frac{\partial \xi_{i2v}}{\partial \alpha} &= -\frac{1}{\alpha} \xi_{i2v}, & \frac{\partial \xi_{i1v}}{\partial \beta_j} &= -\frac{x_{ij}}{2} \xi_{i2v}, & \frac{\partial \xi_{i2v}}{\partial \beta_j} &= -\frac{x_{ij}}{2} \xi_{i1v}.\end{aligned}$$

Lembremos que estamos apenas na presença de censura intervalar e de censura à direita, assim nossa função de verossimilhança assume a seguinte forma

$$L(\theta) = \prod_{i \in I} (F(\xi_{i2v}) - F(\xi_{i2u})) \prod_{i \in C} F(-\xi_{i2u})$$

conforme apresentado em (3.2). Além disso, de acordo com (3.3) a log-verossimilhança é

$$l(\theta) = \sum_{i \in I} l_1(\alpha, \xi_{i2u}, \xi_{i2v}) + \sum_{i \in C} l_2(\alpha, \xi_{i2u})$$

em que,

$$l_1(\alpha, \xi_{i2u}, \xi_{i2v}) = \log(F(\xi_{i2v}) - F(\xi_{i2u}))$$

e

$$l_2(\alpha, \xi_{i2u}) = \log(F(-\xi_{i2u}))$$

Daí segue que,

$$\frac{\partial l(\theta)}{\partial \alpha} = \frac{1}{\alpha} \sum_{i \in I} \frac{f(\xi_{i2u})\xi_{i2u} - f(\xi_{i2v})\xi_{i2v}}{F(\xi_{i2v}) - F(\xi_{i2u})} + \frac{1}{\alpha} \sum_{i \in C} h(\xi_{i2u})\xi_{i2u}$$

e

$$\frac{\partial l(\theta)}{\partial \beta_j} = \frac{1}{2} \sum_{i \in I} x_{ij} \left(\frac{\xi_{i1u} f(\xi_{i2u}) - \xi_{i1v} f(\xi_{i2v})}{F(\xi_{i2v}) - F(\xi_{i2u})} \right) + \frac{1}{2} \sum_{i \in C} x_{ij} \xi_{i1u} h(\xi_{i2u})$$

A.1 Cálculo de $\frac{\partial^2 l(\theta)}{\partial \alpha^2}$

Observemos agora que,

$$\begin{aligned} \frac{\partial}{\partial \alpha} (f(\xi_{i2u}) \xi_{i2u}) &= \frac{\partial f(\xi_{i2u})}{\partial \alpha} \xi_{i2u} + f(\xi_{i2u}) \frac{\partial \xi_{i2u}}{\partial \alpha} \\ &= f'(\xi_{i2u}) \left(-\frac{1}{\alpha} \right) \xi_{i2u}^2 + f(\xi_{i2u}) \left(-\frac{1}{\alpha} \right) \xi_{i2u} \\ &= -\frac{1}{\alpha} (f'(\xi_{i2u}) \xi_{i2u}^2 + f(\xi_{i2u}) \xi_{i2u}) \\ &= -\frac{1}{\alpha} b_{iu} \end{aligned}$$

com $b_{iu} = f'(\xi_{i2u}) \xi_{i2u}^2 + f(\xi_{i2u}) \xi_{i2u}$.

Analogamente,

$$\frac{\partial}{\partial \alpha} (f(\xi_{i2v}) \xi_{i2v}) = -\frac{1}{\alpha} b_{iv}$$

em que $b_{iv} = f'(\xi_{i2v}) \xi_{i2v}^2 + f(\xi_{i2v}) \xi_{i2v}$.

Notemos que,

$$\begin{aligned} \frac{\partial (h(\xi_{i2u}) \xi_{i2u})}{\partial \alpha} &= \frac{\partial h(\xi_{i2u})}{\partial \alpha} \xi_{i2u} + h(\xi_{i2u}) \frac{\partial \xi_{i2u}}{\partial \alpha} \\ &= -\frac{\xi_{i2u}}{\alpha} \left[\left(\frac{f'(\xi_{i2u})}{F(-\xi_{i2u})} - \left(\frac{f(\xi_{i2u})}{F(-\xi_{i2u})} \right)^2 \right) \xi_{i2u} + h(\xi_{i2u}) \right] \end{aligned}$$

Observemos também que,

$$\begin{aligned} \frac{\partial}{\partial \alpha} \left(\frac{f(\xi_{i2u}) \xi_{i2u} - f(\xi_{i2v}) \xi_{i2v}}{F(\xi_{i2v}) - F(\xi_{i2u})} \right) &= -\frac{1}{\alpha} \left[\frac{(b_{iu} - b_{iv})}{F(\xi_{i2v}) - F(\xi_{i2u})} \right] \\ &\quad - \frac{1}{\alpha} \left[\frac{f(\xi_{i2u}) \xi_{i2u} - f(\xi_{i2v}) \xi_{i2v}}{F(\xi_{i2v}) - F(\xi_{i2u})} \right]^2 \end{aligned}$$

Assim,

$$\begin{aligned} \frac{\partial^2 l(\theta)}{\partial \alpha^2} &= \sum_{i \in I} \left[-\frac{1}{\alpha^2} \left(\frac{f(\xi_{i2u}) \xi_{i2u} - f(\xi_{i2v}) \xi_{i2v}}{F(\xi_{i2v}) - F(\xi_{i2u})} \right) + \frac{1}{\alpha} \left[\frac{\partial}{\partial \alpha} \left(\frac{f(\xi_{i2u}) \xi_{i2u} - f(\xi_{i2v}) \xi_{i2v}}{F(\xi_{i2v}) - F(\xi_{i2u})} \right) \right] \right] \\ &\quad - \sum_{i \in C} \left[\frac{\xi_{i2u}^2}{\alpha^2} \left(\frac{f'(\xi_{i2u}) F(-\xi_{i2u}) + f^2(\xi_{i2u})}{F^2(-\xi_{i2u})} \right) \right] \end{aligned}$$

Em notação matricial temos,

$$\ddot{\mathbf{L}}_{\alpha\alpha} = \text{tr}(\mathbf{M}), \quad (\text{A.1})$$

em que $\mathbf{M} = \text{diag}\{m_1(\boldsymbol{\theta}), m_2(\boldsymbol{\theta}), \dots, m_n(\boldsymbol{\theta})\}$, com

$$m_i(\boldsymbol{\theta}) = \begin{cases} -\frac{1}{\alpha^2} \left(\frac{f(\xi_{i2u})\xi_{i2u} - f(\xi_{i2v})\xi_{i2v}}{F(\xi_{i2v}) - F(\xi_{i2u})} \right) + \frac{1}{\alpha} \left[\frac{\partial}{\partial \alpha} \left(\frac{f(\xi_{i2u})\xi_{i2u} - f(\xi_{i2v})\xi_{i2v}}{F(\xi_{i2v}) - F(\xi_{i2u})} \right) \right], & \text{se } i \in I \\ \frac{\xi_{i2u}^2}{\alpha^2} \left(\frac{f'(\xi_{i2u})F(-\xi_{i2u}) + f^2(\xi_{i2u})}{F^2(-\xi_{i2u})} \right), & \text{se } i \in C \end{cases}$$

A.2 Cálculo de $\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_k \partial \alpha}$

Notemos que

$$\begin{aligned} \frac{\partial(f(\xi_{i2u})\xi_{i2u})}{\partial \beta_k} &= f'(\xi_{i2u}) \left(-\frac{x_{ik}}{2} \right) \xi_{i1u} \xi_{i2u} + f(\xi_{i2u}) \left(-\frac{x_{ik}}{2} \right) \xi_{i1u} \\ &= -\frac{x_{ik}}{2} \xi_{i1u} (f'(\xi_{i2u})\xi_{i2u} + f(\xi_{i2u})) \\ &= -\frac{x_{ik}}{2} g_{iu} \end{aligned}$$

em que, $g_{iu} = \xi_{i1u} (f'(\xi_{i2u})\xi_{i2u} + f(\xi_{i2u}))$

Analogamente, obtemos que

$$\frac{\partial(f(\xi_{i2v})\xi_{i2v})}{\partial \beta_k} = -\frac{x_{ik}}{2} g_{iv}$$

em que, $g_{iv} = \xi_{i1v} (f'(\xi_{i2v})\xi_{i2v} + f(\xi_{i2v}))$

Além disso,

$$\begin{aligned} \frac{\partial(h(\xi_{i2u})\xi_{i2u})}{\partial \beta_k} &= \frac{\partial h(\xi_{i2u})}{\partial \beta_k} \xi_{i2u} + h(\xi_{i2u}) \frac{\partial \xi_{i2u}}{\partial \beta_k} \\ &= -\frac{x_{ik}}{2} \xi_{i1u} \left[\left(\frac{f'(\xi_{i2u})}{F(-\xi_{i2u})} + \left(\frac{f(\xi_{i2u})}{F(-\xi_{i2u})} \right)^2 \right) \xi_{i2u} + h(\xi_{i2u}) \right] \end{aligned}$$

e

$$\frac{\partial(F(\xi_{i2v}) - F(\xi_{i2u}))}{\partial \beta_k} = f(\xi_{i2v}) \left(-\frac{x_{ik}}{2} \right) \xi_{i1v} - f(\xi_{i2u}) \left(-\frac{x_{ik}}{2} \right) \xi_{i1u}$$

Daí,

$$\begin{aligned} \frac{\partial^2 l(\theta)}{\partial \beta_k \partial \alpha} &= \frac{1}{\alpha} \sum_{i \in I} \left[-\frac{x_{ik}}{2} \left(\frac{g_{iu} - g_{iv}}{F(\xi_{i2v}) - F(\xi_{i2u})} \right) \right] \\ &+ \frac{1}{\alpha} \sum_{i \in I} \left[-\frac{x_{ik}}{2} \left(\frac{(f(\xi_{i2v})\xi_{i1v} - f(\xi_{i2u})\xi_{i1u})(f(\xi_{i2u})\xi_{i2u} - f(\xi_{i2v})\xi_{i2v})}{(F(\xi_{i2v}) - F(\xi_{i2u}))^2} \right) \right] \\ &- \frac{1}{\alpha} \sum_{i \in C} \left\{ \frac{x_{ik}}{2} \xi_{i1u} \left[\left(\frac{f'(\xi_{i2u})}{F(-\xi_{i2u})} + \left(\frac{f(\xi_{i2u})}{F(-\xi_{i2u})} \right)^2 \right) \xi_{i2u} + h(\xi_{i2u}) \right] \right\} \end{aligned}$$

Em notação matricial,

$$\ddot{\mathbf{L}}_{\alpha\beta} = \mathbf{K}^\top \mathbf{X}, \quad (\text{A.2})$$

em que $\mathbf{K} = (k_1(\boldsymbol{\theta}), k_2(\boldsymbol{\theta}), \dots, k_n(\boldsymbol{\theta}))^\top$, com

$$k_i(\boldsymbol{\theta}) = \begin{cases} \frac{g_{iu} - g_{iv}}{F(\xi_{i2v}) - F(\xi_{i2u})} + \frac{(f(\xi_{i2v})\xi_{i1v} - f(\xi_{i2u})\xi_{i1u})(f(\xi_{i2u})\xi_{i2u} - f(\xi_{i2v})\xi_{i2v})}{(F(\xi_{i2v}) - F(\xi_{i2u}))^2} \\ \text{se } i \in I \\ \xi_{i1u} \left[\left(\frac{f'(\xi_{i2u})}{F(-\xi_{i2u})} + \left(\frac{f(\xi_{i2u})}{F(-\xi_{i2u})} \right)^2 \right) \xi_{i2u} + h(\xi_{i2u}) \right] \\ \text{se } i \in C \end{cases}$$

A.3 Cálculo de $\frac{\partial^2 l(\theta)}{\partial \beta_k \partial \beta_j}$

Observemos agora que,

$$\begin{aligned} \frac{\partial(f(\xi_{i2u})\xi_{i1u})}{\partial \beta_k} &= \frac{\partial f(\xi_{i2u})}{\partial \beta_k} \xi_{i1u} + f(\xi_{i2u}) \frac{\partial \xi_{i1u}}{\partial \beta_k} \\ &= f'(\xi_{i2u}) \left(-\frac{x_{ik}}{2} \right) \xi_{i1u} \xi_{i1u} + f(\xi_{i2u}) \left(-\frac{x_{ik}}{2} \right) \xi_{i2u} \\ &= -\frac{x_{ik}}{2} (f'(\xi_{i2u})\xi_{i1u}^2 + f(\xi_{i2u})\xi_{i2u}) \\ &= -\frac{x_{ik}}{2} a_{iu}. \end{aligned}$$

em que $a_{iu} = (f'(\xi_{i2u})\xi_{i1u}^2 + f(\xi_{i2u})\xi_{i2u})$ e $f'(x) = \frac{\partial F(x)}{\partial x}$

Analogamente, temos $\frac{\partial(f(\xi_{i2v})\xi_{i1v})}{\partial \beta_k} = -\frac{x_{ik}}{2} a_{iv}$, em que

$$a_{iv} = (f'(\xi_{i2v})\xi_{i1v}^2 + f(\xi_{i2v})\xi_{i2v}).$$

Notando agora que,

$$\begin{aligned}\frac{\partial(\xi_{i1u}h(\xi_{i2u}))}{\partial\beta_k} &= \frac{\partial\xi_{i1u}}{\partial\beta_k}h(\xi_{i2u}) + \xi_{i1u}\frac{\partial h(\xi_{i2u})}{\partial\beta_k} \\ &= -\frac{x_{ik}}{2}\left[\xi_{i2u}h(\xi_{i2u}) + \xi_{i1u}^2\left(\frac{f'(\xi_{i2u})}{F(-\xi_{i2u})} + \left(\frac{f(\xi_{i2u})}{F(-\xi_{i2u})}\right)^2\right)\right]\end{aligned}$$

obtemos que,

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\beta_k\partial\beta_j} &= -\frac{1}{4}\sum_{i\in I}\left[x_{ij}x_{ik}\left(\frac{a_{iu} - a_{iv}}{F(\xi_{i2v}) - F(\xi_{i2u})} + \frac{(\xi_{i1u}f(\xi_{i2u}) - \xi_{i1v}f(\xi_{i2v}))^2}{[F(\xi_{i2v}) - F(\xi_{i2u})]^2}\right)\right] \\ &\quad - \frac{1}{4}\sum_{i\in C}\left[x_{ij}x_{ik}\left(\xi_{i2u}h(\xi_{i2u}) + \xi_{i1u}^2\left(\frac{f'(\xi_{i2u})}{F(-\xi_{i2u})} + \left(\frac{f(\xi_{i2u})}{F(-\xi_{i2u})}\right)^2\right)\right)\right]\end{aligned}$$

Em notação matricial,

$$\ddot{\mathbf{L}}_{\beta\beta} = \mathbf{X}^\top \mathbf{D} \mathbf{X}, \quad (\text{A.3})$$

em que $\mathbf{D} = \text{diag}\{d_1(\boldsymbol{\theta}), d_2(\boldsymbol{\theta}), \dots, d_n(\boldsymbol{\theta})\}$, com

$$d_i(\boldsymbol{\theta}) = \begin{cases} \frac{a_{iu} - a_{iv}}{F(\xi_{i2v}) - F(\xi_{i2u})} + \frac{(\xi_{i1u}f(\xi_{i2u}) - \xi_{i1v}f(\xi_{i2v}))^2}{[F(\xi_{i2v}) - F(\xi_{i2u})]^2} & \text{se } i \in I \\ \xi_{i2u}h(\xi_{i2u}) + \xi_{i1u}^2\left(\frac{f'(\xi_{i2u})}{F(-\xi_{i2u})} + \left(\frac{f(\xi_{i2u})}{F(-\xi_{i2u})}\right)^2\right) & \text{se } i \in C \end{cases}.$$

A.4 A matriz hessiana

Temos que matriz hessiana é dada por

$$\ddot{\mathbf{L}}_{\theta\theta} = \begin{bmatrix} \ddot{\mathbf{L}}_{\alpha\alpha} & \ddot{\mathbf{L}}_{\alpha\beta} \\ \ddot{\mathbf{L}}_{\beta\alpha} & \ddot{\mathbf{L}}_{\beta\beta} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha\partial\alpha} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\alpha\partial\beta} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\beta^\top\partial\alpha} & \frac{\partial^2 l(\boldsymbol{\theta})}{\partial\beta^\top\partial\beta} \end{bmatrix}, \quad (\text{A.4})$$

em que $l(\boldsymbol{\theta})$ é o logaritmo da função de verossimilhança dado em (3.3). Substituindo (A.1), (A.2) e (A.3) em (A.4) temos que

$$\ddot{\mathbf{L}}_{\theta\theta} = \begin{bmatrix} \ddot{\mathbf{L}}_{\alpha\alpha} & \ddot{\mathbf{L}}_{\alpha\beta} \\ \ddot{\mathbf{L}}_{\beta\alpha} & \ddot{\mathbf{L}}_{\beta\beta} \end{bmatrix} = \begin{bmatrix} \text{tr}(\mathbf{M}) & \mathbf{K}^\top \mathbf{X} \\ \mathbf{X}^\top \mathbf{K} & \mathbf{X}^\top \mathbf{D} \mathbf{X} \end{bmatrix} \quad (\text{A.5})$$

Referências Bibliográficas

- [1] Aalen, O. O. (1978). Nonparametric Inference for a Family of Counting Processes. *Annals of Statistics*, **6**, 701-726.
- [2] Jara, A., Hanson, T., Quintana, F., Mueller, P. e Rosner, G. L. (2011). Bayesian Nonparametric and Semiparametric Analysis. Disponível em: <http://cran.r-project.org/web/packages/DPpackage/DPpackage.pdf>. Acesso 26 de out. 2011
- [3] Atkinson, A. C. (1985). *Plots, Transformations and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*. Clarendon Press, Oxford.
- [4] Barlow, W. E. e Prentice, R. L. (1988). Residuals for relative risk regression. *Biometrika* **75**, 65-74.
- [5] Barros, M. (2007) *Modelos de regressão Birnbaum-Saunders generalizados*. Tese - IME/USP, São Paulo.
- [6] Barros, M., Paula, G.A., Leiva, V. (2008) A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. *Lifetime Data Anal.* **14**, 316-332
- [7] Birnbaum, Z. W. e Saunders, S. C. (1958). A Statistical Model for Life-Length of Materials. *Journal of American Statistical Association*, **72**, 46-53.
- [8] Birnbaum, Z. W. e Saunders, S. C. (1969a). A new family of life distributions. *Journal of Applied Probability* **6**, 319-327.
- [9] Birnbaum, Z. W. e Saunders, S. C. (1969b). Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability* **6**, 328-347.

- [10] Brown, B. W., Flood, M. M. (1947). Tumbler mortality. *Journal of American Statistical Association*, **42**, 562-574.
- [11] Colosimo, E.A. e Giolo, S.R. (2006). *Análise de Sobrevida Aplicada*. Edgard Blücher.
- [12] Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics* **19**, 15-18.
- [13] Cook, R. D. (1986). Assessment of local influence (with discussion). *Journal of the Royal Statistical Society B* **48**, 133-169
- [14] Cox, D. R. e Snell E. J. (1968). A General Definition of Residuals. *Journal of the Royal Statistical Society B*, Oxford, **30**, p.248-275.
- [15] Cox, D.R., Snell, E.J. (1981). *Applied Statistics*, Science Paperbacks. Chapman and Hall, London.
- [16] Díaz-García, J. A. e Domínguez-Molina, J. R. (2006). *Some generalisations of Birnbaum-Saunders and sinh-normal distributions*. *International Mathematical Forum* **1**, 1709- 1727.
- [17] Díaz-García, J. A. e Leiva, V. (2005). A new family of life distributions based on the contoured elliptically distributions. *Journal of Statistical Planning and Inference* **128**, 88-95.
- [18] Díaz-García, J. A. e Leiva, V. (2006). Erratum to a new family of life distributions based on the elliptically contoured distributions [Journal of Statistical Planning and Inference **128**(2) (2005) 445-457]. To appear in *Journal of Statistical Planning and Inference*.
- [19] Doornik, J. A. (2011). *Ox: an Object-Oriented Matrix Language*. fourth ed. Timberlake Consultants Press, London, Oxford. (www.doornik.com).
- [20] Fang, K. T. e Anderson, T. W. (1990). *Statistical Inference in Elliptical Contoured and Related Distributions*. Allerton Press, New York.

- [21] Farrington, C.P. (2000). Residuals for proportional hazards models with interval-censored failure time data. *Biometrics*, Washington, **56**, p.473-482.
- [22] Galea, M., Riquelme, M., and Paula, G. A. (2000). Diagnostics methods in elliptical linear regression models. *Brazilian Journal of Probability and Statistics*, **14**, pp. 167-184.
- [23] Gupta, A. K. e Varga, T. (1993). *Elliptically Contoured Models in Statistics*. Kluwer Academic Publishers.
- [24] Hashimoto, E.M., Ortega, E.M.M., Cancho, V.G., and Cordeiro, G.M. (2010). The log-exponentiated Weibull regression model for interval-censored data. *In Proceedings of Computational Statistics & Data Analysis*, **54**, 1017-1035.
- [25] James, B. R. (2006). *Probabilidade: Um curso em nível intermediário*. Projeto Euclides. IMPA
- [26] Kalbfleisch, J. D. e Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. John Wiley and Sons, New York.
- [27] Kaplan, E. L., Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457-481.
- [28] Klein, J. P., Moeschberger, M. L. (2003). *Survival Analysis: Techniques for censored and Truncated Data*. Springer-Verlag, New York, 2nd edition.
- [29] Lagakos, S.W. (1980). The graphical evaluation of explanatory variables in proportional hazards regression. *Biometrika*, Cambridge, **568**, p.93-98, 1980.
- [30] Lesaffre, E. e Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics* **54**, 570-582.
- [31] Leiva, V.; Barros M.; Paula, G.A. (2009). *Generalized Birnbaum-Saunders Models using R*. Recife: XI Escola de Modelos de Regressão.
- [32] Leiva, V., Barros, M., Paula, G.A., Galea, M. (2007) Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. *Comp. Stat. Data Anal.* **51**, 5694-5707.

- [33] Lucas, A. (1997). Robustness of the student t based M-estimator. *Communications in Statistics – Theory and Methods*, **26**, 1165-1182.
- [34] Mann, N. R., Schafer, R. E., and Singpurwalla, N. D. (1974) *Methods for Statistical Analysis of Reliability and Life Data*. John Wiley e sons, New York.
- [35] Marshall, A. W. and Olkin, I. (2007). *Life Distributions—Structure of Nonparametric, Semiparametric, and Parametric Families*. Springer, New York.
- [36] Meeker W. Q. and Escobar L. A. (1998) *Statistical Methods for Reliability Data*. John Wiley e sons, New York.
- [37] Moolgavkar, S., Lustbader, E. and Venzon, D. (1984). A geometric approach to nonlinear regression diagnostic with application to matched case-control studies. *Annals of Statistics* **12**, 816-826.
- [38] Nelson, W. (1972). Theory and Applications of Hazard Plotting for censored Failure Data. *Technometrics*, **14**, 945-965.
- [39] Ng H. K., Kundu D., and Balakrishnan N. (2003) Modified moment estimation for the two-parameter Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis*, **43**, pp. 283-298.
- [40] Paula, G. A., Leiva, V., Barros, M. and Liu, S. (2011). Robust statistical modeling using the Birnbaum-Saunders-t distribution applied to insurance. *Applied Stochastic Models in Business and Industry* (Print).
- [41] Paula, G. A. (2010) *Modelos de regressão com apoio computacional*. São Paulo, IME/USP. 325p.
- [42] Paula, G. A. e Peres, C. A. (1988). Diagnostics for GLMs with linear inequality parameter constraints. **Communications in Statistics: Theory and Methods** **17**, 4205-4219.
- [43] Podlaski, R. (2008). Characterization of diameter distribution data in near-natural forests using the Birnbaum-Saunders distribution. *Canadian Journal of Forest Research*, **38**, 518-527.

- [44] Pulgar, G. M. I. (2009) *Modelos mistos aditivos semiparamétricos de contornos elípticos*. Tese - IME/USP, São Paulo.
- [45] Rieck, J. R. (1989). *Statistical Analysis for the Birnbaum-Saunders Fatigue Life Distribution*. Ph. D. Thesis, Clemson University, Department of Mathematical Sciences.
- [46] Rieck, J. R. and Nedelman, J. R. (1991). A log-linear model for the Birnbaum-Saunders distribution. *Technometrics* **33**, 51-60.
- [47] Saunders, S. C. (2007). *Reliability, Life Testing and Prediction of Services Lives*. Springer, New York.
- [48] Souza, F. A. M. D. (1999). *Influência Local e Análise de Resíduos em Modelos de Regressão Von Mises*. Tese de Doutorado. IME/USP, São Paulo.
- [49] Turnbull, B. W. (1976). Nonparametric Estimation of a Survivorship Function with doubly Censored Data. *J. R. Statist. Soc. B*, **38**, 290-295.
- [50] Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer, New York.
- [51] Weisberg, S. (1985). *Applied linear regression*. 2nd ed. New York; John Wiley e Sons, 324p.
- [52] Zhao, Q. (2004). *Nonparametric treatment comparisons for interval-censored failure time data*. 2004. 130p. Thesis (Phylosophy Doctor) - Faculty of Graduate School, University of Missouri, Columbia.