



LUIS ENRIQUE BENITES SÁNCHEZ

MODELOS BIRNBAUM-SAUNDERS BIVARIADOS

CAMPINAS  
2014





UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE MATEMÁTICA, ESTATÍSTICA  
E COMPUTAÇÃO CIENTÍFICA

LUIS ENRIQUE BENITES SÁNCHEZ

MODELOS BIRNBAUM-SAUNDERS BIVARIADOS

Dissertação apresentada ao Instituto de  
Matemática, Estatística e Computação Científica da Universidade Estadual  
de Campinas como parte dos requisitos exigidos para  
obtenção do título de Mestre em estatística.

**Orientador: Prof. Dr. Filidor Edilson Vilca Labra**

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO  
DEFENDIDA PELO ALUNO LUIS ENRIQUE BENITES SÁNCHEZ,  
E ORIENTADA PELO PROF. DR. FILIDOR EDILFONSO VILCA LABRA.

Assinatura do Orientador

A handwritten signature in black ink, written over a horizontal line. The signature is cursive and appears to read "Filidor Edilson Vilca Labra".

CAMPINAS  
2014

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Matemática, Estatística e Computação Científica  
Maria Fabiana Bezerra Muller - CRB 8/6162

B437m Benites Sánchez, Luis Enrique, 1983-  
Modelos Birnbaum-Saunders bivariados / Luis Enrique Benites Sánchez. –  
Campinas, SP : [s.n.], 2014.

Orientador: Filidor Edilfonso Vilca Labra.  
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de  
Matemática, Estatística e Computação Científica.

1. Distribuição (Probabilidades). 2. Modelos lineares (Estatística). 3.  
Inferência. I. Vilca Labra, Filidor Edilfonso, 1964-. II. Universidade Estadual de  
Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Models bivariate Birnbaum-Saunders

**Palavras-chave em inglês:**

Distribution (Probability theory)

Linear models (Statistics)

Inference

**Área de concentração:** Estatística

**Titulação:** Mestre em Estatística

**Banca examinadora:**

Filidor Edilfonso Vilca Labra [Orientador]

Caio Lucidius Naberezny Azevedo

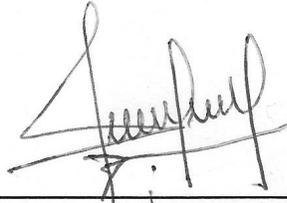
Camila Borelli Zeller

**Data de defesa:** 07-03-2014

**Programa de Pós-Graduação:** Estatística

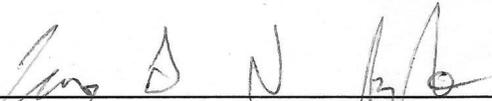
**Dissertação de Mestrado defendida em 07 de março de 2014 e aprovada**

**Pela Banca Examinadora composta pelos Profs. Drs.**



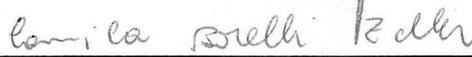
---

**Prof(a). Dr(a). FILIDOR EDILFONSO VILCA LABRA**



---

**Prof(a). Dr(a). CAIO LUCIDIUS NABEREZNY AZEVEDO**



---

**Prof(a). Dr(a). CAMILA BORELLI ZELLER**



## Abstract

Several works have been done on the univariate  $\mathcal{BS}$  distribution and its extensions, the bivariate version of this model was presented only recently by Kundu *et al.* (2010). They proposed a bivariate  $\mathcal{BS}$  distribution with dependence structure and established several attractive properties of that bivariate distribution. It possesses a close relationship with the bivariate normal distribution just as the univariate  $\mathcal{BS}$  distribution has with the univariate normal.

In this work provides a study some aspect of inference, analysis of diagnostics and lifetime analysis based on the failure rate function of bivariate  $\mathcal{BS}$  distributions: Hypotheses test studies are considered using the Wald, Score and Likelihood Ratios statistics, the analysis of diagnostics is based on the Cook (1986) approach, and the discussion on lifetime analysis is based on the idea of Basu (1971).

Finally, numerical examples are given to illustrate our methodology and the properties of the statistics are investigated through Monte Carlo simulations.

## Resumo

Vários trabalhos tem sido feitos sobre a distribuição  $\mathcal{BS}$  univariada e suas extensões. A versão bivariada deste modelo foi apresentada recentemente por Kundu *et al.* (2010). Eles propuseram uma distribuição  $\mathcal{BS}$  bivariada com estrutura de dependência e estabeleceram várias propriedades atrativas para a distribuição  $\mathcal{BS}$  bivariada, que possui uma estreita relação com a distribuição normal bivariada; assim como a distribuição  $\mathcal{BS}$  univariada tem com a distribuição normal univariada.

Este trabalho apresenta um estudo de alguns aspectos de inferência, análise de diagnóstico e análise de tempo de vida baseada na função taxa de falha da distribuição  $\mathcal{BS}$  bivariada: aspectos de inferência serão através de testes de hipótese considerando as estatísticas de Wald, Razão de Verossimilhança e Escore; o análise de diagnóstico será baseada na metodologia de Cook (1986) e a discussão sobre análise sobrevivência será baseada na idéia de Basu (1971).

Finalmente, exemplos numéricos serão apresentados para ilustrar as metodologias propostas e as propriedades das estatísticas serão investigadas por meio de simulações de Monte Carlo .



# Sumário

<b>Lista de Figuras</b>	<b>xviii</b>
<b>Lista de Tabelas</b>	<b>xx</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Introdução . . . . .	1
1.2 Conceitos básicos de sobrevivência . . . . .	2
1.2.1 Função de Sobrevivência . . . . .	2
1.2.2 Função de taxa de falha . . . . .	3
1.3 O modelo Birnbaum-Saunders . . . . .	3
1.4 Objetivos do Trabalho . . . . .	6
1.5 Suporte computacional . . . . .	6
1.6 Organização do Trabalho . . . . .	6
<b>2 O modelo Birnbaum-Saunders Bivariado</b>	<b>9</b>
2.1 Introdução . . . . .	9
2.2 Distribuição Birnbaum-Saunders Bivariada . . . . .	10
2.2.1 Propriedades . . . . .	11
2.3 Matriz de Informação de Fisher Observada . . . . .	21
2.4 Matriz de Informação de Fisher Esperada . . . . .	22
2.5 Estimação dos parâmetros . . . . .	24
2.5.1 Estimação por Máxima Verossimilhança . . . . .	24
2.5.2 Estimação pelo Método dos Momentos . . . . .	28
2.6 Testes de Hipóteses . . . . .	31
2.6.1 Estatística da razão de verossimilhança . . . . .	32
2.6.2 Estatística de Escore . . . . .	32
2.6.3 Estatística de Wald . . . . .	32
2.7 Curva TTT . . . . .	33
2.8 Estudo de simulação . . . . .	34
2.9 Aplicação 1 . . . . .	36
2.10 Aplicação 2 . . . . .	40

<b>3</b>	<b>Análise de influência local no modelo BS Bivariado</b>	<b>43</b>
3.1	Introdução . . . . .	43
3.2	Método de eliminação de casos . . . . .	44
3.3	Influência local . . . . .	46
3.4	Aplicação no Modelo Birnbaum-Saunders Bivariado . . . . .	49
3.4.1	Perturbação de Ponderação de Casos . . . . .	49
3.5	Ilustração numérica . . . . .	51
3.5.1	Exemplo 1 . . . . .	51
<b>4</b>	<b>Análise de sobrevivência para um modelo BS bivariado</b>	<b>61</b>
4.1	Introdução . . . . .	61
4.2	A fda da distribuição normal bivariada . . . . .	62
4.2.1	Aproximação de Tsay and Ke (2009) . . . . .	63
4.2.2	Aproximação de Albers e Kallenberg (1994) . . . . .	67
4.2.3	Aproximação de Cox e Wermuth (1991) . . . . .	67
4.3	Dados de tempo de vida bivariado . . . . .	68
4.3.1	Função de sobrevivência bivariada . . . . .	68
4.3.2	Função de taxa de falha bivariada . . . . .	70
4.4	Estudo de Simulação . . . . .	78
<b>5</b>	<b>Considerações Finais</b>	<b>83</b>
5.1	Conclusões . . . . .	83
<b>A</b>	<b>Código R da aproximação feita por Tsay Ke (2009)</b>	<b>85</b>
<b>B</b>	<b>Código C++, para obter as tabelas finais do Capítulo 4</b>	<b>87</b>
<b>C</b>	<b>Matriz de Informação de Fisher</b>	<b>91</b>
<b>I</b>	<b>Conjuntos de dados</b>	<b>97</b>
<b>II</b>	<b>Código C++ da aproximação feita por Meyer (2013)</b>	<b>99</b>
<b>III</b>	<b>Licença</b>	<b>103</b>
III.1	Sobre a licença dessa obra . . . . .	103
	<b>Referência Bibliográfica</b>	<b>110</b>

*A Jesus, meu Senhor e Salvador. . . .*



Disse-lhe Jesus: Eu sou o caminho, e a verdade e a vida; ninguém vem ao Pai, senão por mim  
(João 14:6)



# Agradecimentos

A Jesus, meu Senhor e Salvador, pois reconheço que, se não fosse por ele, não tinha feito nada.

- A minha esposa Rocío, por seu apoio, compreensão e paciência.
- A minha mãe, Delia, que sempre me apoio em tudo e dava força para lutar.
- Aos meus irmãos, Christian, Marcos e Susan, por seu constante apoio.
- A minha avó, Isabel, porque sei que sempre ora por mim.
- Aos meus tios, Luis e Miro, pelos conselhos que me ajudaram na minha vida.
- Ao meu orientador, Dr. Filidor Edilfonso Vilca Labra , pela amizade, oportunidade, animo, motivação e paciência.
- Ao Prof. Víctor Hugo Lachos Dávila, porque sem me conhecer, quando cheguei neste país, me mostrou sua amabilidade, apoio, amizade.
- Ao senhor Reginaldo Ramos Mendes, pela sua amável colaboração neste trabalho.
- À CAPES, pelo apoio financeiro recebido na elaboração deste trabalho.



# Lista de Figuras

1.1	(a) Fdp e (b) Fda da distribuição $\mathcal{BS}$ para diferentes valores de $\alpha$ e fixando $\beta = 1.0$ .	4
1.2	(a) Funções de sobrevivência e (b) risco do modelo BS considerando diferentes valores para $\alpha$ e fixando $\beta = 1$ .	5
2.1	Fdp e contorno do modelo BS bivariado quando $\alpha_1 = \alpha_2 = 1$ , $\beta_2 = \beta_2 = 1$ : (a) e (b) $\rho = 0.5$ , (c) e (d) $\rho = -0.5$	17
2.2	Fdp e contorno do modelo BS bivariado quando $\alpha_1 = \alpha_2 = 1$ , $\beta_2 = \beta_2 = 1$ : (a) e (b) $\rho = 0.9$ , (c) e (d) $\rho = -0.9$	18
2.3	Fdp e contorno do modelo BS bivariado quando $\alpha_1 = 0.5$ , $\alpha_2 = 1$ , $\beta_2 = \beta_2 = 1$ : (a) e (b) $\rho = 0.5$ , (c) e (d) $\rho = -0.5$	19
2.4	Fdp e contorno do modelo BS bivariado quando $\alpha_1 = 0.5$ , $\alpha_2 = 1$ , $\beta_2 = \beta_2 = 1$ : (a) e (b) $\rho = 0.9$ , (c) e (d) $\rho = -0.9$	20
2.5	Curvas das transformadas TTT empíricas. As curvas A, B, C, D e E, respectivamente, indicam que os dados possuem taxas de risco crescente, decrescente, constante, banheira e unimodal.	34
2.6	(Esquerda) Q-Q plots e envelope simulado e (Direita) Contorno do gráfico da função de densidade estimada para $t_1$ e $t_2$ .	37
2.7	(Esquerda) Diagrama de dispersão entre $T_1$ e $T_2$ e (Direita) Diagrama de caixas entre $T_1$ e $T_2$ .	38
2.8	(Esquerda) TTT-Plot para o $T_1$ e (Direita) TTT-Plot para o $T_2$ .	38
2.9	Histograma para os dados e as curvas estimados para as funções de densidade marginais de $T_1$ e $T_2$ .	39
2.10	(Esquerda) Q-Q plots e envelope simulado e (Direita) Contorno do gráfico da função de densidade estimada para $T_5$ e $T_6$ .	40
3.1	Representação gráfica do enfoque de influência local.	47
3.2	Distância Generalizada de Cook para os diferentes parâmetros	52
3.3	Gráfico de $LD_i$ sob o modelo BS bivariado	53
3.4	Gráfico de $M(0)$ sob perturbação de ponderação de casos.	54
3.5	Gráfico das medidas de influência local total (a) $C_i(\boldsymbol{\theta})$ vs observações (b) $C_i(\boldsymbol{\alpha})$ vs observações (c) $C_i(\boldsymbol{\beta})$ vs observações e (d) $C_i(\rho)$ vs observações	55
3.6	Ponderação de casos (a) Gráfico de influência total $C_i$ vs observações. (b) Gráfico das medidas de influência local $C_i(\boldsymbol{\alpha})$ vs $C_i(\boldsymbol{\beta})$	56

3.7	Gráfico das medidas de influência local (a) $C_i(\alpha_1)$ vs observações (b) $C_i(\alpha_2)$ vs observações . . . . .	56
3.8	Gráfico das medidas de influência local (a) $C_i(\beta_1)$ vs observações e (b) $C_i(\beta_2)$ vs observações . . . . .	57
3.9	Gráfico das medidas de influência local (a) $C_i(\alpha_1)$ vs $C_i(\beta_1)$ (b) $C_i(\alpha_2)$ vs $C_i(\beta_2)$ (c) $C_i(\alpha_2)$ vs $C_i(\beta_1)$ e (d) $C_i(\alpha_1)$ vs $C_i(\beta_2)$ . . . . .	58
3.10	(a) Gráfico das medidas de influência local $C_i(\alpha_1)$ vs $C_i(\alpha_2)$ (b) Gráfico das medidas de influência local $C_i(\beta_1)$ vs $C_i(\beta_2)$ . . . . .	59
3.11	(a) Gráfico das medidas de influência local $C_i(\rho)$ vs $C_i(\alpha)$ (b) Gráfico das medidas de influência local $C_i(\rho)$ vs $C_i(\beta)$ . . . . .	59
4.1	Função de sobrevivência bivariada conjunta de $(T_1, T_2)$ quando $\alpha_1 = \alpha_2 = 1$ , $\beta_2 = \beta_2 = 1$ para (a) $\rho = 0.5$ , (b) $\rho = -0.5$ , (c) $\rho = 0.9$ , (d) $\rho = -0.9$ . . . . .	69
4.2	Função de sobrevivência bivariada conjunta de $(T_1, T_2)$ quando $\alpha_1 = 0.5$ , $\alpha_2 = 1$ , $\beta_2 = \beta_2 = 1$ para (a) $\rho = 0.5$ , (b) $\rho = -0.5$ , (c) $\rho = 0.9$ , (d) $\rho = -0.9$ . . . . .	70
4.3	(a) Função de sobrevivência condicional $T_1$ , dado $T_2 = t_2 = 1$ , quando $\alpha_1 = \alpha_2 = 1$ , $\beta_2 = \beta_2 = 2$ e (a) $\rho = 0.9$ , (b) $\rho = 0.5$ , (c) $\rho = 0$ , (d) $\rho = -0.9$ . . . . .	71
4.4	Função de taxa de falha bivariada quando $\beta_1 = \beta_2 = 1$ , $\rho = 0$ e (a) $\alpha_1 = \alpha_2 = 1.75$ , (b) $\alpha_1 = \alpha_2 = 1.5$ , (c) $\alpha_1 = \alpha_2 = 1$ , (d) $\alpha_1 = \alpha_2 = 0.75$ . . . . .	73
4.5	Função de taxa de falha bivariada conjunta de $(T_1, T_2)$ quando $\alpha_1 = \alpha_2 = 1$ , $\beta_2 = \beta_2 = 1$ para (a) $\rho = 0.5$ , (b) $\rho = -0.5$ , (c) $\rho = 0.9$ , (d) $\rho = -0.9$ . . . . .	74
4.6	Função de taxa de falha condicional de $T_1$ , dado $T_2 = t_2 = 1$ , quando $\rho = 0.9$ com (a) $t_2 = 0.8$ , (b) $t_2 = 1$ , (c) $t_2 = 2$ , (d) $t_2 = 8$ . . . . .	75
4.7	Função de taxa de falha condicional de $T_1$ , dado $T_2$ , quando $\rho = 0.5$ com (a) $t_2 = 0.8$ , (b) $t_2 = 1$ , (c) $t_2 = 2$ , (d) $t_2 = 8$ . . . . .	76

# Lista de Tabelas

2.1	Médias das estimativas e erro quadrático médio (entre parênteses) de as EMV, quando $\alpha_1 = \alpha_2 = 1$ , $\beta_1 = \beta_2 = 2$ e para diferentes valores de $\rho$ . . . . .	25
2.2	Nível de significância empírica do teste baseado na estatística de Wald, Escore e Razão de Verossimilhança para $\rho = 0.5$ . . . . .	35
2.3	Nível de significância empírica do teste baseado na estatística de Wald, Escore e Razão de Verossimilhança para $\rho = -0.5$ . . . . .	35
2.4	Nível de significância empírica do teste baseado na estatística de Wald, Escore e Razão de Verossimilhança para $\rho = 0.9$ . . . . .	36
2.5	Nível de significância empírica do teste baseado na estatística de Wald, Escore e Razão de Verossimilhança para $\rho = -0.9$ . . . . .	36
2.6	Conjunto de dados de ossos: Estatísticas descritivas para $T_1$ e $T_2$ . . . . .	37
2.7	EMV para os parâmetros. Os valores dos EP, entre parênteses, e 95% intervalo de confiança (IC) . . . . .	37
2.8	Parâmetros EMV da distribuição BS bivariada para os dados apresentados na Tabela acima . . . . .	39
2.9	Estatísticas do teste Wald, Razão de Verossimilhança e Escore . . . . .	39
2.10	EMV para os parâmetros. Os valores dos EP, entre parênteses, e 95% intervalo de confiança (IC) . . . . .	41
2.11	Parâmetros EMV da distribuição BS bivariada para os dados apresentados na Tabela acima . . . . .	41
2.12	Estatísticas do teste Wald, Razão de Verossimilhança e Escore . . . . .	41
3.1	Mudança nas estimativas dos parâmetros (em %), excluindo as observações potencialmente influentes. . . . .	53
79table.4.1		
4.2	$\rho = 0.4$ . . . . .	79
4.3	$\rho = 0.6$ . . . . .	80
4.4	$\rho = 0.8$ . . . . .	80
81table.4.5		
4.6	$\rho = 0.4$ . . . . .	81
4.7	$\rho = 0.6$ . . . . .	82
4.8	$\rho = 0.8$ . . . . .	82

I.1	Densidade mineral dos ossos . . . . .	97
I.2	Contenido de mineral em ossos . . . . .	98

# Capítulo 1

## Introdução

### 1.1 Introdução

A distribuição Birnbaum-Saunders ( $\mathcal{BS}$ ), proposta por Birnbaum & Saunders (1969a) e Birnbaum & Saunders (1969b), baseia-se num argumento físico de dano cumulativo que produz fadiga no material. Este argumento é conhecido como a “regra de Miner”, para mais detalhes veja Miner (1945). Birnbaum & Saunders (1968) obteve uma interpretação probabilística desta regra. Assim, Birnbaum & Saunders (1969a) descreveu um modelo de distribuição de vida que pode ser derivado de um processo de fadiga física, onde o crescimento da rachadura causa a falha. Desmond (1985) reforçou a justificativa física para o uso desta distribuição e algumas hipóteses antecipadas por Birnbaum & Saunders (1969a). Além disso, uma derivação mais geral da distribuição  $\mathcal{BS}$  foi fornecida por Desmond (1985), com base em um modelo biológico introduzido por Cramér (1946). Para mais detalhes sobre as distribuições de vida, o leitor poder consultar Johnson *et al.* (1995) e Marshall & Olkin (2007). Li & From (2006) apresentaram alternativas para os estimadores de máxima verossimilhança dos parâmetros da distribuição  $\mathcal{BS}$ . Cysneiro *et al.* (2008) realizaram inferência com a distribuição  $\mathcal{BS}$ .

A distribuição  $\mathcal{BS}$ , conhecida também como a distribuição de vida, tem sido aplicada na indústria, negócios, confiabilidade, análise de sobrevivência, ciências ambientais e ciências médicas. Recentemente, tem sido aplicada também em empresas e meio ambiente; por exemplo, veja Leiva *et al.* (2007), Leiva *et al.* (2008), Barros *et al.* (2008) e Paula *et al.* (2012). No contexto de negócios, Paula *et al.* (2011) propuseram o uso de modelagem estatística robusta baseada nos modelos  $\mathcal{BS}$ .

Birnbaum & Saunders (1969a) obtiveram originalmente os estimadores de máxima verossimilhança para os parâmetros do modelo. Mann *et al.* (1974) mostraram que a distribuição  $\mathcal{BS}$  é unimodal. Engelhardt *et al.* (1981) propuseram intervalos de confiança para o parâmetro de forma ( $\alpha$ ).

Por otro lado diversos fenômenos do mundo real podem ser descritos por distribuições multivariadas, sendo um tema amplamente estudado pois muitas distribuições univariadas generalizadas para formar distribuições bivariadas. Inúmeras distribuições discretas (Binomial, Hipergeométrica

e Poisson) e contínuas (Uniforme, Normal, Exponencial, Beta, Weibull e Gama), ver Johnson *et al.* (1994) e Johnson *et al.* (1995) para distribuições univariadas e Johnson *et al.* (1997), Kotz *et al.* (2000), Balakrishnan & Lai (2009) para as distribuições bivariadas. A distribuição  $\mathcal{BS}$  está sendo amplamente considerada, devido suas boas propriedades e sua estreita relação com a distribuição normal; veja Johnson *et al.* (1994).

Recentemente, Kundu *et al.* (2010) apresentaram uma extensão da distribuição  $\mathcal{BS}$  univariada para o caso bivariado. Uma extensão para o caso multivariado pode ser encontrada em Kundu *et al.* (2013), seguindo os estudos já considerados para outras distribuições contínuas como as distribuições Beta, exponencial, Gama, Uniforme e Weibull, entre outras; veja Johnson *et al.* (1994) para distribuições univariadas, e Kotz *et al.* (2000) para distribuições bivariadas.

Uma observação influente produz alterações relevantes no resultado da análise quando por exemplo é extraída. Uma técnica que pode ajudar na detecção de observações influentes é o diagnóstico de influência. O mesmo apresenta basicamente duas abordagens. A primeira baseia-se na deleção de observações através da distância de Cook, veja Cook & Weisberg (1982). A segunda metodologia de diagnóstico de influência, chamada de Influência Local foi proposta por Cook (1986), baseada na medida de afastamento de verossimilhança. Na literatura, a abordagem predominante na análise de influência local para conjuntos de dados se baseia na simples análise exploratória. Poucos trabalhos fornecem marcas de referência para determinar estatisticamente as observações são influentes. Dentre essas exceções, temos Poon & Poon (1999).

## 1.2 Conceitos básicos 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.

### 1.2.1 Função de Sobrevivência

A função de sobrevivência da variável aleatória  $T$ , denotada por  $S_T(t)$ , é fundamental em análise de sobrevivência. Segundo Kalbfleisch & Prentice (2002) a probabilidade de uma observação não falhar, ou seja, sobreviver até o tempo  $t$ , denominada de função de sobrevivência é dada por

$$S_T(t) = P(T > t) = 1 - F_T(t), \quad t \in \mathbb{R}_+, \quad (1.2.1)$$

em que  $F_T(t)$  representa a função de distribuição acumulada (fda) da variável aleatória  $T$ . Esta é uma das principais funções probabilísticas usada para descrever estudos de sobrevivência. Algumas propriedades da função de sobrevivência são:

(i)  $S_T(\cdot)$  é monótona não-crescente.

(ii)  $S_T(0) = 1$  e  $\lim_{t \rightarrow \infty} S_T(t) = 0$ .

(iii)  $S_T(\cdot)$  é contínua.

### 1.2.2 Função de taxa de falha

Sendo  $T$  uma variável aleatória contínua não negativa, a função taxa de falha instantânea, também conhecida como função de risco, de uma variável aleatória não negativa,  $T$ , é dada por

$$h_T(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}. \quad (1.2.2)$$

Pode-se mostrar facilmente que

$$h_T(t) = -\frac{\partial \ln[S_T(t)]}{\partial t}.$$

## 1.3 O modelo Birnbaum-Saunders

O interesse na distribuição  $\mathcal{BS}$  decorre de suas propriedades atrativas e sua estreita relação com a distribuição normal. De fato, a distribuição  $\mathcal{BS}$  está relacionada com a distribuição normal através da representação estocástica a seguir,

$$T = \beta \left[ \frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1} \right]^2, \quad (1.3.1)$$

em que  $Z \sim N(0, 1)$ ,  $\alpha > 0$  e  $\beta > 0$ . A variável aleatória  $T$  em (1.3.1) é dita ter uma distribuição  $\mathcal{BS}$  com parâmetros de forma e escala,  $\alpha$  e  $\beta$ , respectivamente, e é geralmente denotada por  $T \sim \mathcal{BS}(\alpha, \beta)$ ; veja Birnbaum & Saunders (1969a). A fda de  $T$  é dada por

$$F_T(t) = \Phi(a(t; \alpha, \beta)), \quad t \in \mathbb{R}_+, \quad (1.3.2)$$

em que  $\Phi(\cdot)$  é a fda da distribuição normal padrão e  $a(t; \alpha, \beta) = (\sqrt{t/\beta} - \sqrt{\beta/t})/\alpha$ .

Considerando a fda da variável aleatória  $T$ , definida em (1.3.2), a sua correspondente fdp é dada por

$$f_T(t) = \phi(a(t; \alpha, \beta))A(t; \alpha, \beta), \quad (1.3.3)$$

em que

$$A(t; \alpha, \beta) = \frac{\partial a(t; \alpha, \beta)}{\partial t} = \frac{t^{-3/2}(t + \beta)}{2\alpha\beta^{1/2}}.$$

Algumas propriedades da distribuição  $\mathcal{BS}$ ,  $T \sim \mathcal{BS}(\alpha, \beta)$ , são

1.  $cT \sim \mathcal{BS}(\alpha, c\beta) \quad \forall$  constante  $c$  positiva;
2.  $T^{-1} \sim \mathcal{BS}(\alpha, \beta^{-1})$ .

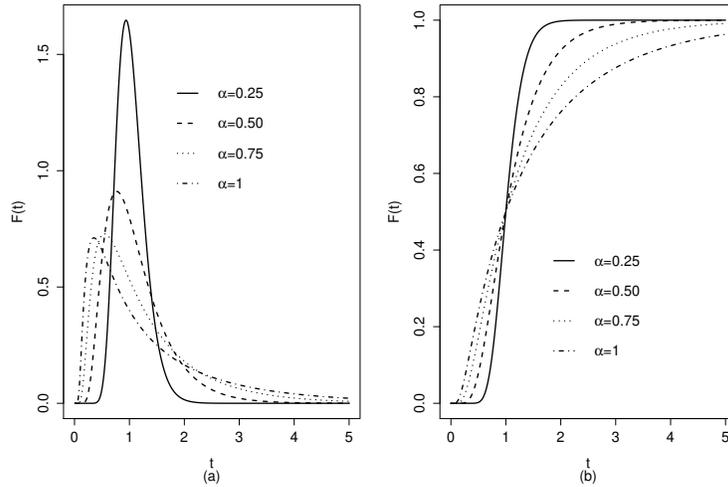


Figura 1.1: (a) Fdp e (b) Fda da distribuição  $\mathcal{BS}$  para diferentes valores de  $\alpha$  e fixando  $\beta = 1.0$ .

Na Figura 1.1(a) e 1.1(b) apresentamos a função de densidade de probabilidade (fdp) e fda do modelo  $\mathcal{BS}$ , respectivamente, para alguns valores de  $\alpha$  considerando  $\beta = 1$ . Na Figura 1.1(a) note-se que à medida que o valor de  $\alpha$  decresce, a curva tende a ficar mais simétrica em torno de  $\beta$ , que é a mediana da distribuição, ou seja,  $F(\beta) = 0.5$ . Da mesma forma, nota-se que a variância também decresce com  $\alpha$ .

Agora, suponhamos que  $T \sim \mathcal{BS}(\alpha, \beta)$ , então de (1.3.1) temos que a variável aleatória  $Z$  definida por

$$Z = \frac{1}{\alpha} \left[ \sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right], \quad (1.3.4)$$

tem distribuição  $N(0, 1)$  ou  $X = \frac{\alpha}{2}Z \sim N(0, \frac{\alpha^2}{4})$ . Além disso, a média, a variância, e coeficiente de variação e coeficiente de assimetria ( $\mu_3$ ) e curtose ( $\mu_4$ ) da distribuição  $\mathcal{BS}$  são, respectivamente,

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

$$\mu_3 = \frac{16\alpha^2(11\alpha^2 + 6)}{(5\alpha^2 + 4)^3} \quad \text{e} \quad \mu_4 = 3 + \frac{6\alpha^2(93\alpha^2 + 41)}{5 + \alpha^2 + 4}.$$

Apresentamos a função de sobrevivência e risco e respectivamente e na Figura 1.2(a) e 1.2(b) a função de sobrevivência e risco para o modelo  $\mathcal{BS}$ .

$$S_T(t) = 1 - \Phi(a(t; \alpha, \beta)),$$

e

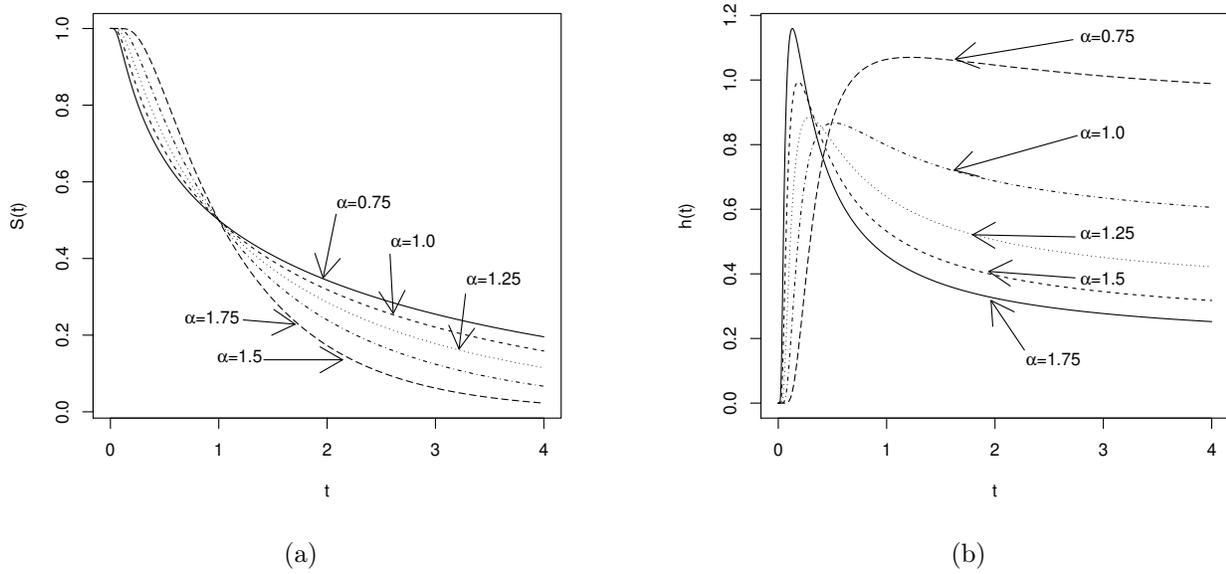


Figura 1.2: (a) Funções de sobrevivência e (b) risco do modelo BS considerando diferentes valores para  $\alpha$  e fixando  $\beta = 1$ .

$$h_T(t) = \frac{\phi(a(t; \alpha, \beta))A(t; \alpha, \beta)}{1 - \Phi(a(t; \alpha, \beta))}.$$

A função risco  $h_T(t)$  é zero em  $t = 0$ , cresce até um máximo para algum  $t_0$  e finalmente decresce até um valor finito. De fato, o comportamento assintótico de  $h_T(t)$  é apresentado na Figura 1.2(b), e de acordo a Kundu *et al.* (2008), temos que

$$\lim_{t \rightarrow \infty} h_T(t) = \frac{1}{2\alpha^2\beta}.$$

Finalmente, considere as variáveis aleatórias  $T_1, T_2 \sim BS(\alpha, \beta)$  independentes e as variáveis aleatórias  $T_{(1)}$  e  $T_{(2)}$  definidas, respectivamente por

$$T_{(1)} = \min(T_1, T_2) \text{ e } T_{(2)} = \max(T_1, T_2).$$

Então, suas respectivas fdp são dadas por

$$f_{T_{(1)}}(t) = 2\phi(a(t; \alpha, \beta))\Phi(-a(t; \alpha, \beta))A(t; \alpha, \beta), \quad t > 0,$$

e

$$f_{T_{(2)}}(t) = 2\phi(a(t; \alpha, \beta))\Phi(-a(t; \alpha, \beta))A(t; \alpha, \beta), \quad t > 0.$$

## 1.4 Objetivos do Trabalho

Os principais objetivos deste trabalho são estudar alguns aspectos de inferência, análise de diagnóstico e análise de sobrevivência no modelo  $\mathcal{BS}$  bivariado. O estudo de diagnóstico é desenvolvido seguindo a metodologia de Cook (1986). Métodos de aproximação serão apresentados para aproximar a fda da normal bivariada. Os objetivos específicos podem ser resumidos como segue:

1. Descrever o modelo  $\mathcal{BS}$  bivariado, obtendo a matriz de informação de Fisher esperada;
2. Propor o teste de Wald, Escore e o teste da Razão de Verossimilhança para verificar se os parâmetros são iguais e se não há presença de correlação.
3. Realizar um estudo de diagnóstico no modelo  $\mathcal{BS}$  bivariado, seguindo a metodologia Cook (1986);
4. Apresentar e descrever a função de sobrevivência, função de risco no modelo  $\mathcal{BS}$  bivariado, usando a aproximação de Tsay & Ke (2009).

Finalmente, dois exemplos numéricos serão apresentados para ilustrar a metodologia e as propriedades da estatística de Razão de Verossimilhança, Wald e Escore serão investigados através de simulações de Monte Carlo.

## 1.5 Suporte computacional

As avaliações numéricas realizadas ao longo desta dissertação foram feitas através do R e C++ para sistemas operacionais Windows para o caso do C++ foi usado o compilador online, que pode ser encontrado no endereço [http://www.compileonline.com/compile\\_cpp11\\_online.php](http://www.compileonline.com/compile_cpp11_online.php). Para a geração dos gráficos e aplicações, foi utilizado o software R na sua versão 2.15.3 para sistema operacionais Windows e que se encontra disponível de forma gratuita no endereço [www.r-project.org/](http://www.r-project.org/). Para a realização dos cálculos das derivadas parciais, foi utilizado o software Matlab (licenciado pelo Instituto de Matemática, Estatística e Ciências da Computação da UNICAMP) na sua versão 2010 para sistemas operacionais Windows. Por fim, o trabalho foi digitado usando o sistema de tipografia  $\text{\LaTeX}$ . Maiores detalhes podem ser encontrados em Mittelbach *et al.* (2004), ou através do site [www.tex.ac.uk/CTAN/latex](http://www.tex.ac.uk/CTAN/latex)

## 1.6 Organização do Trabalho

A presente dissertação encontra-se dividida em cinco capítulos. Neste primeiro capítulo foi apresentada uma introdução do modelo  $\mathcal{BS}$  e algumas de suas propriedades, uma breve descrição para a função de sobrevivência e taxa de falha.

No Capítulo 2, é feita uma revisão dos conceitos básicos da distribuição  $\mathcal{BS}$  bivariada destacando suas propriedades e representação estocástica. São obtidas as matrizes de informação observada e esperada de Fisher. Estimções de parâmetros pelo método de máxima verossimilhança

e momentos foram também consideradas, assim como suas propriedades assintóticas, e testes de hipótese foram discutidos, baseados nas estatísticas de Wald, Escore e Razão de Verossimilhança. Um pequeno estudo de simulação é apresentado para estudar o comportamento dos três testes. Ainda, duas aplicações foram desenvolvidas.

No Capítulo 3, é feito um estudo dos métodos de influência local propostos por Cook (1986). Consideramos o tipo de perturbação de ponderação de casos. Além disso, apresentamos uma aplicação dos resultados obtidos a um conjunto de dados reais encontrado na literatura.

No Capítulo 4, é feito um estudo da função de sobrevivência e taxa de falha para a distribuição  $\mathcal{BS}$  bivariada, são apresentadas definições básicas e um pequeno estudo de simulação da função de risco.

O Capítulo 5 é dedicado às conclusões e alguns direcionamentos para estudos subsequentes.

Finalmente, nos apêndices, apresentamos alguns resultados da literatura para obter a matriz de informação esperada, além de um pacote implementado no R utilizado para obter as estimativas para os parâmetros do modelo e outras opções que serão detalhadas no decorrer do trabalho.



# Capítulo 2

## O modelo Birnbaum-Saunders Bivariado

### 2.1 Introdução

Diversos fenômenos do mundo real podem ser descritos por distribuições multivariadas, sendo um tema amplamente estudado, pois muitas distribuições univariadas generalizadas para formar distribuições bivariadas. Temos distribuições discretas (binomial, hipergeométrica e Poisson) e contínuas (uniforme, normal, exponencial, beta, Weibull e gama), ver Johnson *et al.* (1994) e Johnson *et al.* (1995) para distribuições univariadas e Johnson *et al.* (1997), Kotz *et al.* (2000), Balakrishnan & Lai (2009) para as distribuições bivariadas. A distribuição  $\mathcal{BS}$  está sendo amplamente considerada, devido às suas boas propriedades e uma estreita relação com a distribuição normal; veja Johnson *et al.* (1994).

Particularmente, devido à sua aplicabilidade, modelos bidimensionais são frequentemente usados; veja Johnson & Wichern (1998) e Kotz *et al.* (2000), pois eles tem sido aplicados em empresas e indústrias, meio ambiente e ciências médicas, confiabilidade e análise de sobrevivência; por exemplo, veja Leiva *et al.* (2007), Leiva *et al.* (2008), Barros *et al.* (2008) e Paula *et al.* (2012).

Recentemente, Kundu *et al.* (2010) apresenta uma extensão da distribuição  $\mathcal{BS}$  univariada para o caso bivariado. Seguindo os estudos já considerados para outras distribuições assimétricas contínuas como as distribuições Beta, exponencial, Gama, Uniforme e Weibull, entre outras; veja Johnson *et al.* (1994) para distribuições univariadas, e Kotz *et al.* (2000) para distribuições bivariadas.

A estimação por máxima verossimilhança é amplamente utilizada e tem sido também considerada para distribuições  $\mathcal{BS}$ , veja Birnbaum & Saunders (1969b) e Kundu *et al.* (2010). Estimação pelo método de momentos é mais simples do que a estimação por máxima verossimilhança, mas em algumas a obtenção das estimativas torna-se difícil. Uma forma alternativa para estimar estes parâmetros é usar o método de momento modificado (MM), que foi discutido em Ng *et al.* (2003), Leiva *et al.* (2008) e Kundu *et al.* (2010).

Neste capítulo, investigamos aspectos de inferência para a distribuição bivariada. Sobre a esti-

mação dos parâmetros do modelo, foi considerado o método dos momentos modificado (MM) e o método de máxima verossimilhança. Provamos que os estimadores obtidos via o método MM são consistentes e distribuídos assintoticamente normais. Estatísticas do tipo Wald foram consideradas assintoticamente distribuídas de acordo com a distribuição qui-quadrado. Esta estatística é baseada no método de MM, que são mostrados para ser consistente e assintoticamente normalmente distribuída.

Sob a suposição de normalidade é apresentada a matriz de informação de Fisher esperada em forma matricial que é apropriada para facilitar a implementação das estatísticas Wald e Escore. Estatísticas de teste com base em estimadores momento são mais simples para cálculos, seu desempenho melhora significativamente à medida que  $n$  aumenta. Comparações entre eles são também relatadas. Por fim, apresentamos os resultados de um estudo de simulação, juntamente com aplicações.

Foi desenvolvido o pacote `bsbiv()` (incluído no software R) para o modelo  $\mathcal{BS}$  bivariado, o qual pode ser obtido em CRAN (<http://CRAN.R-project.org/>) na sua versão 0.1.

## 2.2 Distribuição Birnbaum-Saunders Bivariada

A distribuição  $\mathcal{BS}$  univariada está relacionada com a distribuição normal univariada através da variável aleatória

$$T = \beta(\alpha Z/2 + \sqrt{(\alpha Z/2)^2 + 1})^2, \quad (2.2.1)$$

em que  $Z \sim N(0, 1)$ ,  $\alpha > 0$  e  $\beta > 0$ . A variável aleatória  $T$  é dita ter uma distribuição  $\mathcal{BS}$  com parâmetros forma e escala,  $\alpha$  e  $\beta$ , respectivamente, que é denotada por  $T \sim \mathcal{BS}(\alpha, \beta)$ . A fdp de  $T$  é dado por

$$F_T(t) = \Phi(a(t; \alpha, \beta)), t > 0,$$

em que  $\Phi(\cdot)$  é a fda normal padrão e  $a(t; \alpha, \beta) = (\sqrt{t/\beta} - \sqrt{\beta/t})/\alpha$ .

De Kundu *et al.* (2010) temos que um vetor aleatório bivariado  $\mathbf{T} = (T_1, T_2)^\top$  tem uma distribuição  $\mathcal{BS}$  bivariada com parâmetros  $\alpha_1 > 0$ ,  $\alpha_2 > 0$ ,  $\beta_1 > 0$ ,  $\beta_2 > 0$ , e  $-1 < \rho < 1$ , se sua fda conjunta de  $T_1$  e  $T_2$  pode ser expressa como

$$F_{\mathbf{T}}(\mathbf{t}) = \Phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho), \quad \mathbf{t} = (t_1, t_2)^\top \in \mathbb{R}_+^2,$$

em que  $\Phi_2(\cdot; \rho)$  é a fda conjunta de  $\mathbf{Z} = (Z_1, Z_2) \sim N_2(\mathbf{0}, \boldsymbol{\Sigma})$ , com  $\sigma_{11} = \sigma_{22} = 1$  e o coeficiente de correlação  $\rho$ ,  $a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = (a(t_1; \alpha_1, \beta_1), a(t_2; \alpha_2, \beta_2))^\top$ , com

$$a(t_j; \alpha_j, \beta_j) = (1/\alpha_j)(\sqrt{t_j/\beta_j} - \sqrt{\beta_j/t_j}), \quad \text{para } j = 1, 2, \quad \boldsymbol{\alpha} = (\alpha_1, \alpha_2)^\top \text{ e } \boldsymbol{\beta} = (\beta_1, \beta_2)^\top.$$

Neste caso, usamos a notação  $\mathbf{T} \sim \mathcal{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ . De Kundu *et al.* (2010),  $\mathbf{T}$  tem fdp conjunta dada por

$$f_{\mathbf{T}}(\mathbf{t}) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp(-d(\mathbf{t})/2) A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho) A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}), \quad \mathbf{t} \in \mathbb{R}_+^2, \quad (2.2.2)$$

em que  $\phi_2(\cdot; \rho)$  é a fdp da  $\mathbf{Z}$  dado por

$$\phi_2(\mathbf{z}; \rho) = \phi_2(z_1, z_2; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(z_1^2 + z_2^2 - 2\rho z_1 z_2)\right),$$

$$d(\mathbf{t}) = d(\mathbf{t}, \boldsymbol{\theta}) = a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}), \quad (2.2.3)$$

e

$$A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = A(t_1; \alpha_1, \beta_1) A(t_2; \alpha_2, \beta_2),$$

com

$$A(t_j; \alpha_j, \beta_j) = t_j^{-3/2} (t_j + \beta_j) / (2\alpha_j \beta_j^{1/2}), \quad \text{para } j = 1, 2.$$

### 2.2.1 Propriedades

Algumas propriedades da distribuição  $\mathcal{BS}$  bivariada podem ser derivadas usando a relação estreita entre as distribuições de  $\mathbf{T}$  e  $\mathbf{Z}$ ; veja Kundu *et al.* (2010).

**Teorema 2.2.1.** *Seja  $\mathbf{T} \sim \mathcal{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ . Então*

- (i)  $T_j \sim \mathcal{BS}(\alpha_j, \beta_j)$ , para  $j = 1, 2$ ;
- (ii)  $\mathbf{c} \odot \mathbf{T} \sim \mathcal{BS}_2(\boldsymbol{\alpha}, \mathbf{c} \odot \boldsymbol{\beta}, \rho)$ , em que  $\mathbf{c} = (c_1, c_2)^\top \in \mathbb{R}_+^2$  e  $\odot$  denota o produto de Hadamard;
- (iii)  $\mathbf{T}^{-1} = (T_1^{-1}, T_2^{-1})^\top \sim \mathcal{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}^{-1}, \rho)$ , em que  $\boldsymbol{\beta}^{-1} = (1/\beta_1, 1/\beta_2)^\top$ ;
- (iv)  $\mathbf{T}_1^{-1} = (T_1^{-1}, T_2)^\top \sim \mathcal{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}_1^{-1}, -\rho)$ , em que  $\boldsymbol{\beta}_1^{-1} = (1/\beta_1, \beta_2)^\top$ ;
- (v)  $\mathbf{T}_2^{-1} = (T_1, T_2^{-1})^\top \sim \mathcal{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}_2^{-1}, -\rho)$ , em que  $\boldsymbol{\beta}_2^{-1} = (\beta_1, 1/\beta_2)^\top$ ;
- (vi)  $T_1$  e  $T_2$  são independente se e somente se  $\rho = 0$ ;
- (vii) A fdp condicional de  $T_1$ , dado  $T_2 = t_2$ , é

$$f_{T_1|T_2}(t_1|t_2) = \phi(a(t_1; \alpha_{1\rho}, \beta_1) - \mu_1(t_2)) A(t_1; \alpha_{1\rho}, \beta_1), \quad (2.2.4)$$

em que  $\phi(\cdot)$  é a fdp da distribuição  $N(0, 1)$ ,  $\alpha_{1\rho} = \sqrt{1-\rho^2} \alpha_1$  e  $\mu_1(t_2) = \rho a_{t_2}(\alpha_{2\rho}, \beta_2)$ , com  $\alpha_{2\rho} = \sqrt{1-\rho^2} \alpha_2$ . A partir da fdp dada em (2.2.4), note que a distribuição condicional de  $T_1$  dado  $T_2 = t_2$  é uma distribuição univariada não-central  $\mathcal{BS}$ , que é indicada por  $T_1|(T_2 = t_2) \sim \text{NBS}(\alpha_{1\rho}, \beta_1, \mu_1(t_2))$ ; veja Guiraud *et al.* (2009).

**Demonstração:**

(i) Seja  $\mathbf{T} \sim \text{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ . Então,

$$\begin{aligned}
 f_{T_1}(t_1) &= \int_0^\infty f_{\mathbf{T}}(\mathbf{t}) dt_2 \\
 &= \int_0^\infty \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(a_{t_1}^2 + a_{t_2}^2 - 2\rho a_{t_1} a_{t_2})\right] A_{t_1} A_{t_2} dt_2 \\
 &= \frac{1}{2\pi\sqrt{1-\rho^2}} A_{t_1} \exp\left(-\frac{1}{2(1-\rho^2)} a_{t_1}^2\right) \int_0^\infty \exp\left[-\frac{(a_{t_2}^2 - 2\rho a_{t_1} a_{t_2})}{2(1-\rho^2)}\right] A_{t_2} dt_2 \\
 &= \frac{A_{t_1}}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left(-\frac{a_{t_1}^2}{2(1-\rho^2)}\right) \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{((a_{t_2} - \rho a_{t_1})^2 - \rho^2 a_{t_1}^2)}{2(1-\rho^2)}\right] A_{t_2} dt_2 \\
 &= \frac{A_{t_1}}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left(-\frac{a_{t_1}^2 + \rho a_{t_1}^2}{2(1-\rho^2)}\right) \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2(1-\rho^2)}(a_{t_2} - \rho a_{t_1})^2\right] A_{t_2} dt_2 \\
 &= \frac{A_{t_1}}{\sqrt{2\pi}} \exp\left(-\frac{a_{t_1}^2}{2}\right) \int_0^\infty \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(a_{t_2} - \rho a_{t_1})^2\right] A_{t_2} dt_2 \\
 &= \frac{A_{t_1}}{\sqrt{2\pi}} \exp\left(-\frac{a_{t_1}^2}{2}\right) = \phi(a_{t_1}) A_{t_1}. \quad \blacksquare
 \end{aligned}$$

(ii) Seja  $\mathbf{c} = (c_1, c_2)^\top \in \mathbb{R}_+^2$  e  $\mathbf{Y} = (Y_1, Y_2)$  com  $Y_1 = c_1 T_1$ ,  $Y_2 = c_2 T_2$ . Então  $T_1 = Y_1/c_1$ ,  $T_2 = Y_2/c_2$ . Portanto, o jacobiano é  $|J| = \begin{vmatrix} 1/c_1 & 0 \\ 0 & 1/c_2 \end{vmatrix} = 1/(c_1 c_2)$ . Como

$$f_{\mathbf{T}}(\mathbf{t}) = \phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho) A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}), \quad \mathbf{t} \in \mathbb{R}_+^2,$$

vamos ter que

$$\begin{aligned}
 f_{\mathbf{Y}}(\mathbf{y}) &= f_{\mathbf{T}}(\mathbf{c}^{-1} \odot \mathbf{y}) |J| \\
 &= \phi_2(a(\mathbf{c}^{-1} \odot \mathbf{y}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho) A(\mathbf{c}^{-1} \odot \mathbf{y}; \boldsymbol{\alpha}, \boldsymbol{\beta}) |J| \\
 &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)} \left[ \frac{1}{\alpha_1^2} \left( \sqrt{\frac{y_1/c_1}{\beta_1}} - \sqrt{\frac{\beta_1}{y_1/c_1}} \right)^2 \right. \right. \\
 &\quad \left. \left. + \frac{1}{\alpha_2^2} \left( \sqrt{\frac{y_2/c_2}{\beta_2}} - \sqrt{\frac{\beta_2}{y_2/c_2}} \right)^2 \right] \right\}
 \end{aligned}$$

(2.2.6)

$$\begin{aligned}
 & - \frac{2\rho}{\alpha_1\alpha_2} \left( \sqrt{\frac{y_1/c_1}{c_1\beta_1}} - \sqrt{\frac{\beta_1}{y_1/c_1}} \right) \left( \sqrt{\frac{y_2/c_2}{\beta_2}} - \sqrt{\frac{\beta_2}{y_2/c_2}} \right) \left] \left[ \frac{(y_1/c_1)^{-3/2}(y_1/c_1 + \beta_1)}{2\alpha_1\sqrt{\beta_1}} \right] \right. \\
 & \left. \left[ \frac{(y_2/c_2)^{-3/2}(y_2/c_2 + \beta_2)}{2\alpha_2\sqrt{\beta_2}} \right] \frac{1}{c_1} \frac{1}{c_2} \right. \\
 & = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{1}{\alpha_1^2} \left( \sqrt{\frac{y_1}{c_1\beta_1}} - \sqrt{\frac{c_1\beta_1}{y_1}} \right)^2 + \frac{1}{\alpha_2^2} \left( \sqrt{\frac{y_2}{c_2\beta_2}} - \sqrt{\frac{c_2\beta_2}{y_2}} \right)^2 \right. \right. \\
 & \left. \left. - \frac{2\rho}{\alpha_1\alpha_2} \left( \sqrt{\frac{y_1}{c_1\beta_1}} - \sqrt{\frac{c_1\beta_1}{y_1}} \right) \left( \sqrt{\frac{y_2}{c_2\beta_2}} - \sqrt{\frac{c_2\beta_2}{y_2}} \right) \right] \right\} \left[ \frac{y_1^{-3/2}(y_1 + c_1\beta_1)}{2\alpha_1\sqrt{c_1\beta_1}} \right] \\
 & \left[ \frac{y_2^{-3/2}(y_2 + c_2\beta_2)}{2\alpha_2\sqrt{c_2\beta_2}} \right] \\
 & = \phi_2(a(\mathbf{y}; \boldsymbol{\alpha}, \mathbf{c} \odot \boldsymbol{\beta}); \rho) A(\mathbf{y}; \boldsymbol{\alpha}, \mathbf{c} \odot \boldsymbol{\beta}). \quad \blacksquare
 \end{aligned}$$

(iii) Seja  $\mathbf{Y} = (Y_1, Y_2)$  em que  $Y_1 = T_1^{-1}$ ,  $Y_2 = T_2^{-1}$ . Então  $T_1 = Y_1^{-1}$ ,  $T_2 = Y_2^{-1}$ . Portanto, o jacobiano é  $|J| = \begin{vmatrix} -y_1^{-2} & 0 \\ 0 & -y_2^{-2} \end{vmatrix} = y_1^{-2}y_2^{-2}$ . Como

$$f_{\mathbf{T}}(\mathbf{t}) = \phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho) A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}), \quad \mathbf{t} \in \mathbb{R}_+^2,$$

vamos ter que

$$\begin{aligned}
 f_{\mathbf{Y}}(\mathbf{y}) & = f_{\mathbf{T}}(\mathbf{y}^{-1}) |J| \\
 & = \phi_2(a(\mathbf{y}^{-1}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho) A(\mathbf{y}^{-1}; \boldsymbol{\alpha}, \boldsymbol{\beta}) |J| \\
 & = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{1}{\alpha_1^2} \left( \sqrt{\frac{y_1^{-1}}{\beta_1}} - \sqrt{\frac{\beta_1}{y_1^{-1}}} \right)^2 + \frac{1}{\alpha_2^2} \left( \sqrt{\frac{y_2^{-1}}{\beta_2}} - \sqrt{\frac{\beta_2}{y_2^{-1}}} \right)^2 \right. \right. \\
 & \left. \left. - \frac{2\rho}{\alpha_1\alpha_2} \left( \sqrt{\frac{y_1^{-1}}{\beta_1}} - \sqrt{\frac{\beta_1}{y_1^{-1}}} \right) \left( \sqrt{\frac{y_2^{-1}}{\beta_2}} - \sqrt{\frac{\beta_2}{y_2^{-1}}} \right) \right] \right\} \left[ \frac{(y_1)^{3/2}(y_1^{-1} + \beta_1)}{2\alpha_1\sqrt{\beta_1}} \right] \\
 & \left[ \frac{y_2^{3/2}(y_2^{-1} + \beta_2)}{2\alpha_2\sqrt{\beta_2}} \right] y_1^{-2}y_2^{-2} \\
 & = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{1}{\alpha_1^2} \left( \sqrt{\frac{y_1}{\beta_1^{-1}}} - \sqrt{\frac{\beta_1^{-1}}{y_1}} \right)^2 + \frac{1}{\alpha_2^2} \left( \sqrt{\frac{y_2}{\beta_2^{-1}}} - \sqrt{\frac{\beta_2^{-1}}{y_2}} \right)^2 \right. \right. \\
 & \left. \left. - \frac{2\rho}{\alpha_1\alpha_2} \left( \sqrt{\frac{y_1}{\beta_1^{-1}}} - \sqrt{\frac{\beta_1^{-1}}{y_1}} \right) \left( \sqrt{\frac{y_2}{\beta_2^{-1}}} - \sqrt{\frac{\beta_2^{-1}}{y_2}} \right) \right] \right\} \left[ \frac{y_1^{-3/2}(y_1 + \beta_1^{-1})}{2\alpha_1\sqrt{\beta_1^{-1}}} \right]
 \end{aligned}$$

$$\begin{aligned} & \left[ \frac{y_2^{-3/2}(y_2 + \beta_2^{-1})}{2\alpha_2\sqrt{\beta_2^{-1}}} \right] \\ &= \phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}^{-1}); \rho) A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}^{-1}). \quad \blacksquare \end{aligned}$$

(iv) Seja  $\mathbf{Y} = (Y_1, Y_2)$  em que  $Y_1 = T_1^{-1}$ ,  $Y_2 = T_2$ . Então  $T_1 = Y_1^{-1}$ ,  $T_2 = Y_2$ . Portanto, o jacobiano é  $|J| = \begin{vmatrix} -y_1^{-2} & 0 \\ 0 & 1 \end{vmatrix} = y_1^{-2}$ . Como

$$f_{\mathbf{T}}(\mathbf{t}) = \phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho) A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}), \quad \mathbf{t} \in \mathbb{R}_+^2,$$

vamos ter que

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= f_{\mathbf{T}}(\mathbf{y}_1^{-1}) |J| \\ &= \phi_2(a(\mathbf{y}_1^{-1}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho) A(\mathbf{y}_1^{-1}; \boldsymbol{\alpha}, \boldsymbol{\beta}) |J| \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{1}{\alpha_1^2} \left( \sqrt{\frac{y_1^{-1}}{\beta_1}} - \sqrt{\frac{\beta_1}{y_1^{-1}}} \right)^2 + \frac{1}{\alpha_2^2} \left( \sqrt{\frac{y_2}{\beta_2}} - \sqrt{\frac{\beta_2}{y_2}} \right)^2 \right. \right. \\ &\quad \left. \left. - \frac{2\rho}{\alpha_1\alpha_2} \left( \sqrt{\frac{y_1^{-1}}{\beta_1}} - \sqrt{\frac{\beta_1}{y_1^{-1}}} \right) \left( \sqrt{\frac{y_2}{\beta_2}} - \sqrt{\frac{\beta_2}{y_2}} \right) \right] \right\} \left[ \frac{(y_1)^{3/2}(y_1^{-1} + \beta_1)}{2\alpha_1\sqrt{\beta_1}} \right] \\ &\quad \left[ \frac{y_2^{-3/2}(y_2 + \beta_2)}{2\alpha_2\sqrt{\beta_2}} \right] y_1^{-2} \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{1}{\alpha_1^2} \left( \sqrt{\frac{y_1}{\beta_1^{-1}}} - \sqrt{\frac{\beta_1^{-1}}{y_1}} \right)^2 + \frac{1}{\alpha_2^2} \left( \sqrt{\frac{y_2}{\beta_2}} - \sqrt{\frac{\beta_2}{y_2}} \right)^2 \right. \right. \\ &\quad \left. \left. - \frac{2\rho}{\alpha_1\alpha_2} \left( \sqrt{\frac{y_1}{\beta_1^{-1}}} - \sqrt{\frac{\beta_1^{-1}}{y_1}} \right) \left( \sqrt{\frac{y_2}{\beta_2}} - \sqrt{\frac{\beta_2}{y_2}} \right) \right] \right\} \left[ \frac{y_1^{-3/2}(y_1 + \beta_1^{-1})}{2\alpha_1\sqrt{\beta_1^{-1}}} \right] \\ &\quad \left[ \frac{y_2^{-3/2}(y_2 + \beta_2)}{2\alpha_2\sqrt{\beta_2^{-1}}} \right] \\ &= \phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}_1^{-1}); \rho) A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}_1^{-1}). \quad \blacksquare \end{aligned}$$

(v) Seja  $\mathbf{Y} = (Y_1, Y_2)$  em que  $Y_1 = T_1$ ,  $Y_2 = T_2^{-1}$ . Então  $T_1 = Y_1$ ,  $T_2 = Y_2^{-1}$ . Portanto, o jacobiano é  $|J| = \begin{vmatrix} 1 & 0 \\ 0 & -y_2^{-2} \end{vmatrix} = y_2^{-2}$ . Então,

$$f_{\mathbf{T}}(\mathbf{t}) = \phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho) A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}), \quad \mathbf{t} \in \mathbb{R}_+^2,$$

vamos ter que

$$\begin{aligned}
 f_{\mathbf{Y}}(\mathbf{y}) &= f_{\mathbf{T}}(\mathbf{y}_2^{-1}) |J| \\
 &= \phi_2(a(\mathbf{y}_2^{-1}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho) A(\mathbf{y}_2^{-1}; \boldsymbol{\alpha}, \boldsymbol{\beta}) |J| \\
 &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{1}{\alpha_1^2} \left( \sqrt{\frac{y_1}{\beta_1}} - \sqrt{\frac{\beta_1}{y_1}} \right)^2 + \frac{1}{\alpha_2^2} \left( \sqrt{\frac{y_2^{-1}}{\beta_2}} - \sqrt{\frac{\beta_2}{y_2^{-1}}} \right)^2 \right. \right. \\
 &\quad \left. \left. - \frac{2\rho}{\alpha_1\alpha_2} \left( \sqrt{\frac{y_1}{\beta_1}} - \sqrt{\frac{\beta_1}{y_1}} \right) \left( \sqrt{\frac{y_2^{-1}}{\beta_2}} - \sqrt{\frac{\beta_2}{y_2^{-1}}} \right) \right] \right\} \left[ \frac{(y_1)^{-3/2}(y_1 + \beta_1)}{2\alpha_1\sqrt{\beta_1}} \right] \\
 &\quad \left[ \frac{y_2^{3/2}(y_2^{-1} + \beta_2)}{2\alpha_2\sqrt{\beta_2}} \right] y_2^{-2} \\
 &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{1}{\alpha_1^2} \left( \sqrt{\frac{y_1}{\beta_1}} - \sqrt{\frac{\beta_1}{y_1}} \right)^2 + \frac{1}{\alpha_2^2} \left( \sqrt{\frac{y_2}{\beta_2^{-1}}} - \sqrt{\frac{\beta_2^{-1}}{y_2}} \right)^2 \right. \right. \\
 &\quad \left. \left. - \frac{2\rho}{\alpha_1\alpha_2} \left( \sqrt{\frac{y_1}{\beta_1}} - \sqrt{\frac{\beta_1}{y_1}} \right) \left( \sqrt{\frac{y_2}{\beta_2^{-1}}} - \sqrt{\frac{\beta_2^{-1}}{y_2}} \right) \right] \right\} \left[ \frac{y_1^{-3/2}(y_1 + \beta_1)}{2\alpha_1\sqrt{\beta_1}} \right] \\
 &\quad \left[ \frac{y_2^{-3/2}(y_2 + \beta_2^{-1})}{2\alpha_2\sqrt{\beta_2^{-1}}} \right] \\
 &= \phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}_2^{-1}); \rho) A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}_2^{-1}). \quad \blacksquare
 \end{aligned}$$

(vi) Se  $\mathbf{T} \sim \text{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$  e  $\rho = 0$ . Como

$$f_{\mathbf{T}}(\mathbf{t}) = \phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho) A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}), \quad \mathbf{t} \in \mathbb{R}_+^2,$$

vamos ter que

$$\begin{aligned}
 f_{\mathbf{T}}(\mathbf{t}) &= \frac{1}{2\pi} \exp \left\{ -\frac{1}{2} \left[ \frac{1}{\alpha_1^2} \left( \sqrt{\frac{y_1}{\beta_1}} - \sqrt{\frac{\beta_1}{y_1}} \right)^2 + \frac{1}{\alpha_2^2} \left( \sqrt{\frac{y_2}{\beta_2}} - \sqrt{\frac{\beta_2}{y_2}} \right)^2 \right] \right\} \\
 &\quad \left[ \frac{(y_1)^{-3/2}(y_1 + \beta_1)}{2\alpha_1\sqrt{\beta_1}} \right] \left[ \frac{y_2^{3/2}(y_2 + \beta_2)}{2\alpha_2\sqrt{\beta_2}} \right] \\
 &= \phi(a(t_1; \alpha_1, \beta_1)) A(t_1; \alpha_1, \beta_1) \phi(a(t_2; \alpha_2, \beta_2)) A(t_2; \alpha_2, \beta_2) \\
 &= f_{T_1}(t_1) f_{T_2}(t_2). \quad \blacksquare
 \end{aligned}$$

além disso as propriedades (i)-(vi) permite a obtenção do seguinte resultado:

$$(\alpha_1 T_1 / \beta_1, \alpha_2 T_2 / \beta_2)^\top \sim \text{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\alpha}, \rho),$$

e

$$(T_1 / \beta_1, T_2 / \beta_2)^\top \sim \text{BS}_2(\boldsymbol{\alpha}, \mathbf{1}_2, \rho),$$

em que  $\mathbf{1}_2 = (1, 1)^\top$ . Conseqüentemente, a distribuição de qualquer função  $T_1 / \beta_1$  e  $T_2 / \beta_2$  não dependem de  $\beta_1$  e  $\beta_2$ , tal como  $T_1 T_2 / (\beta_1 \beta_2)$  e  $\sqrt{T_1 T_2} / \sqrt{\beta_1 \beta_2}$ . Assim, as esperanças, de acordo com Kundu *et al.* (2010), são dadas por

$$\psi_1^*(\boldsymbol{\alpha}, \rho) = E[T_1 T_2 / (\beta_1 \beta_2)] = 1 + (\alpha_1^2 + \alpha_2^2) / 2 + \alpha_1^2 \alpha_2^2 (1 + \rho^2) / 4 + \alpha_1 \alpha_2 I_1(\rho), \quad (2.2.7)$$

$$\psi_1(\boldsymbol{\alpha}, \rho) = E[\sqrt{T_1 T_2 / (\beta_1 \beta_2)}] = E[\sqrt{\beta_1 \beta_2 / (T_1 T_2)}] = \alpha_1 \alpha_2 \rho / 4 + I_2(\rho), \quad (2.2.8)$$

em que  $I_1(\rho) = E[Z_1 Z_2 Z_{12}]$  e  $I_2(\rho) = E[Z_{12}]$ , com  $Z_{12} = \sqrt{(\alpha_1 Z_1 / 2)^2 + 1} \sqrt{(\alpha_2 Z_2 / 2)^2 + 1}$ , e  $(Z_1, Z_2)^\top \sim N_2(\mathbf{0}, \boldsymbol{\Sigma})$ . Pode ver facilmente que  $\psi_1(\boldsymbol{\alpha}, -\rho) = E[\sqrt{\beta_1 T_2 / (T_1 \beta_2)}] = E[\sqrt{\beta_2 T_1 / (T_2 \beta_1)}]$ . Expressões para  $I_1(\rho)$  e  $I_2(\rho)$  são apresentados em Kundu *et al.* (2010), que são dadas, respectivamente, por

$$\begin{aligned} I_1(\rho) &= a_{0,0} + \frac{1}{2^3} a_{0,1} (\alpha_1^2 + \alpha_2^2) + \frac{1}{2^6} \alpha_1^2 \alpha_2^2 a_{1,1} + \sum_{i=2}^{\infty} (-1)^{i-1} \frac{1.3 \dots (2i-3)}{2^{3i} i!} a_{0,i} (\alpha_1^{2i} + \alpha_2^{2i}) \\ &+ \sum_{i=2}^{\infty} (-1)^{i-1} \frac{1.3 \dots (2i-3)}{2^{3i+3} i!} a_{1,i} (\alpha_1^2 \alpha_2^{2i} + \alpha_2^2 \alpha_1^{2i}) \\ &+ \sum_{i=2}^{\infty} \sum_{j=2}^{\infty} (-1)^{i+j} \frac{1.3 \dots (2i-3)}{2^{3i} i!} \frac{1.3 \dots (2j-3)}{2^{3j} j!} \alpha_1^{2i} \alpha_2^{2j} a_{i,j}, \end{aligned}$$

$$\begin{aligned} I_2(\rho) &= 1 + \frac{1}{2^3} (\alpha_1^2 + \alpha_2^2) + \frac{1}{2^6} \alpha_1^2 \alpha_2^2 (1 + 2\rho^2) + \sum_{i=2}^{\infty} (-1)^{i-1} \frac{1.3 \dots (2i-3)}{2^{3i} i!} b_{0,i} (\alpha_1^{2i} + \alpha_2^{2i}) \\ &+ \sum_{i=2}^{\infty} (-1)^{i-1} \frac{1.3 \dots (2i-3)}{2^{3i+3} i!} b_{1,i} (\alpha_1^2 \alpha_2^{2i} + \alpha_2^2 \alpha_1^{2i}) \\ &+ \sum_{i=2}^{\infty} \sum_{j=2}^{\infty} (-1)^{i+j} \frac{1.3 \dots (2i-3)}{2^{3i} i!} \frac{1.3 \dots (2j-3)}{2^{3j} j!} \alpha_1^{2i} \alpha_2^{2j} b_{i,j}, \end{aligned}$$

em que  $a_{m,n} = E[Z_1^{2m+1} Z_2^{2n+1}]$  e  $b_{m,n} = E[Z_1^{2m} Z_2^{2n}]$ , cujas expressões podem ser encontradas em Kundu *et al.* (2010).

**Lema 2.2.2.** *Seja  $\mathbf{T} = (T_1, T_2)^\top \sim \text{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$  e  $a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = (a(t_1; \alpha_1, \beta_1), a(t_2; \alpha_2, \beta_2))^\top$ . Então,*

$$(i) \ E[a^{2m+1}(t_1; 1, \beta_1) a^{2n+1}(t_2; 1, \beta_2)] = \alpha_1^{2m+1} \alpha_2^{2n+1} a_{m,n};$$

(ii)  $E[a^{2m}(t_1; 1, \beta_1)a^{2n}(t_2; 1, \beta_2)] = \alpha_1^{2m}\alpha_2^{2n} b_{m,n}$ ,

em que  $a_{m,n}$  e  $b_{m,n}$  são como na expressão de  $I_1(\rho)$  e  $I_2(\rho)$ , respectivamente.

**Observação 2.2.3.** A parte (ii) do Teorema 2.2.1 afirma que a distribuição  $\mathcal{BS}$  bivariada pertence a uma família tipo escala, enquanto que outras partes (iii) - (v) afirmam que estas distribuições são fechadas sob reciprocidade, pelo menos uma da suas componentes.

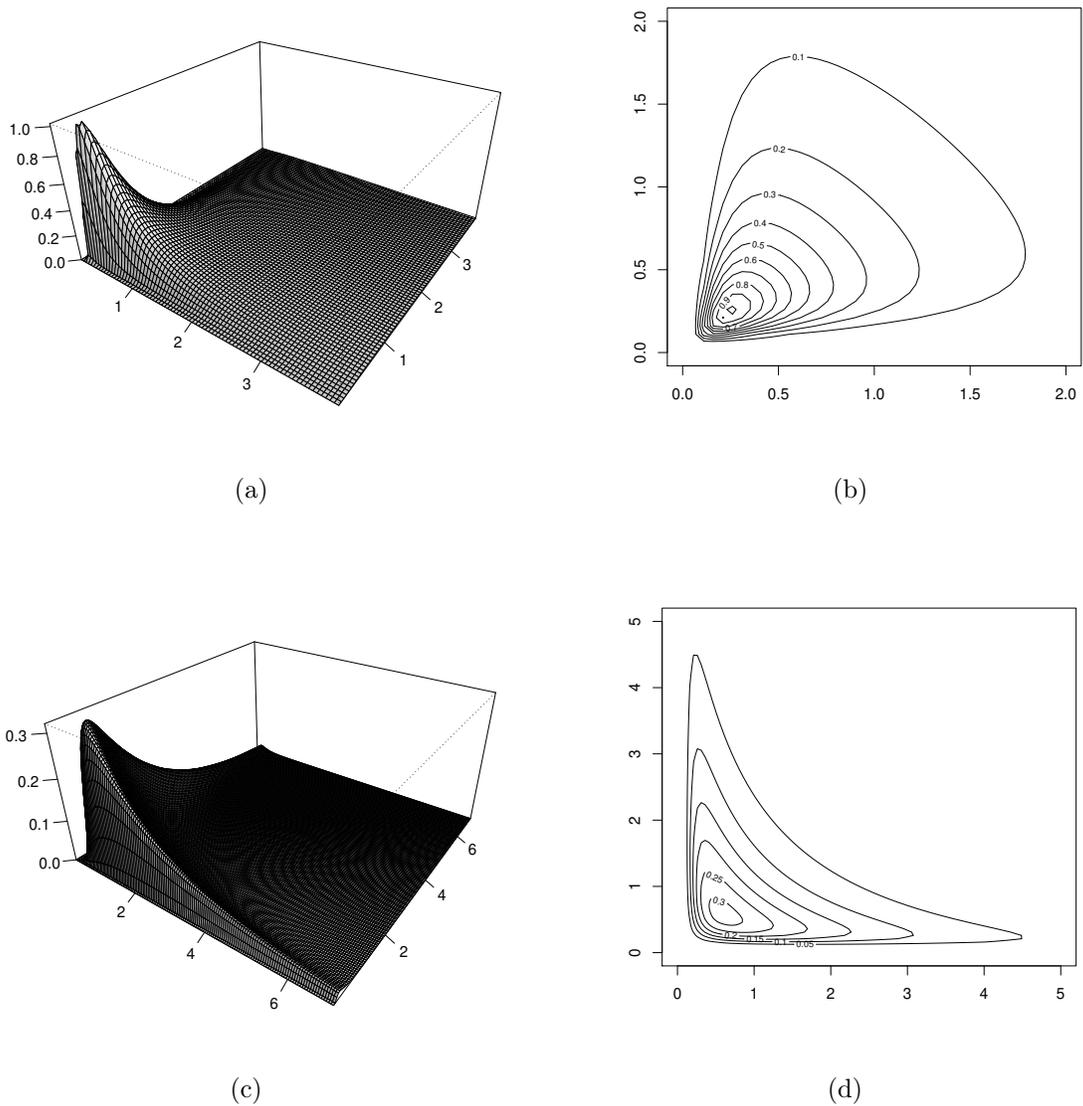


Figura 2.1: Fdp e contorno do modelo BS bivariado quando  $\alpha_1 = \alpha_2 = 1$ ,  $\beta_1 = \beta_2 = 1$ : (a) e (b)  $\rho = 0.5$ , (c) e (d)  $\rho = -0.5$

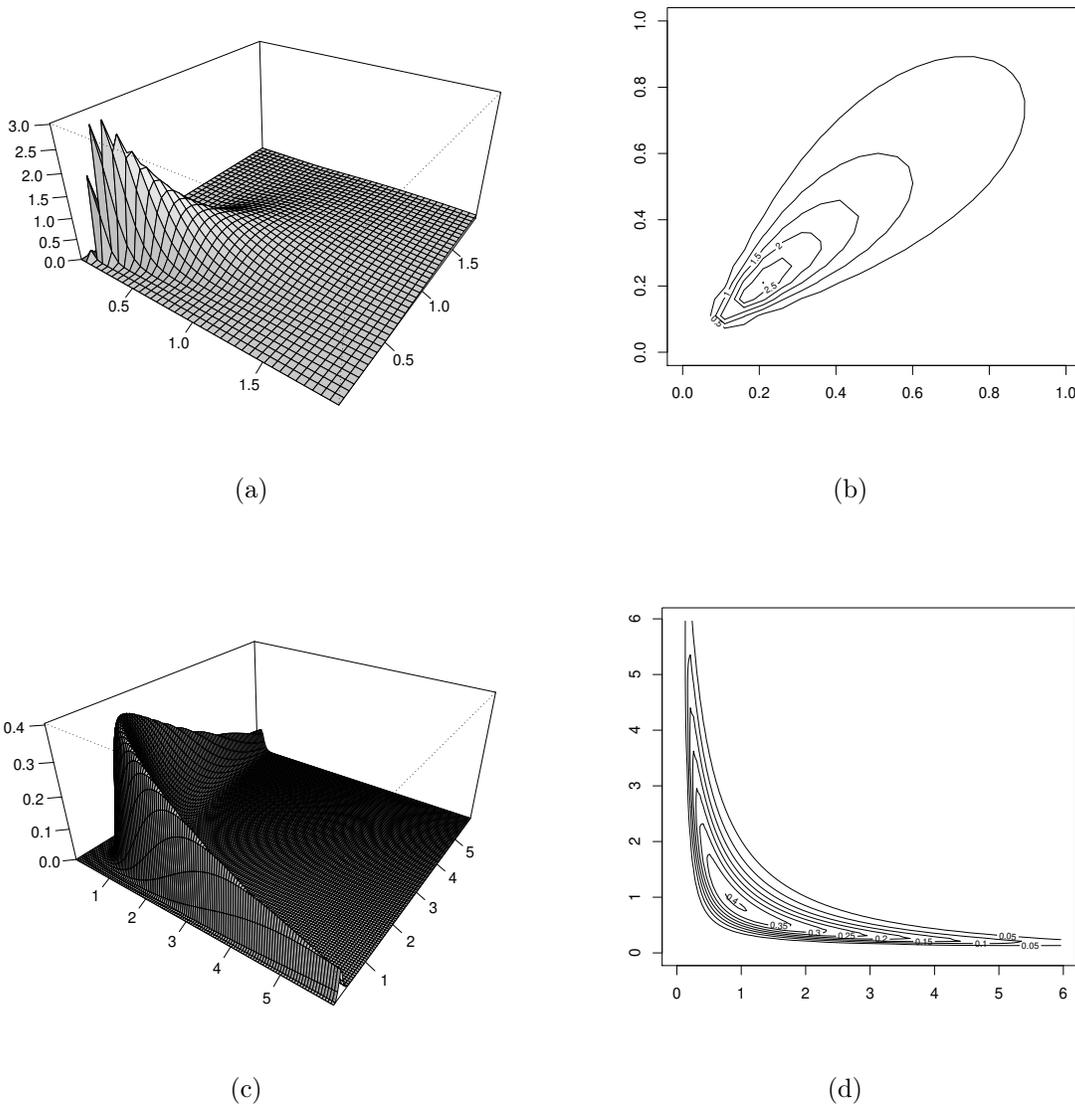


Figura 2.2: Fdp e contorno do modelo BS bivariado quando  $\alpha_1 = \alpha_2 = 1$ ,  $\beta_1 = \beta_2 = 1$ : (a) e (b)  $\rho = 0.9$ , (c) e (d)  $\rho = -0.9$

**Teorema 2.2.4.** *Seja  $\mathbf{T} = (T_1, T_2)^\top \sim BS_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ . Então, os elementos do vetor de médias e matriz de covariância são*

(i)  $\mathbb{E}[T_i] = \frac{\beta_i}{2}(2 + \alpha_i^2);$

(ii)  $Var[T_i] = \frac{\beta_i^2 \alpha_i^2}{4}(4 + 5\alpha_i^2);$

(iii)  $Cov[T_1, T_2] = \beta_1 \beta_2 \alpha_1 \alpha_2 \left[ \frac{1}{4} \alpha_1 \alpha_2 \rho^2 + I_1(\rho) \right];$

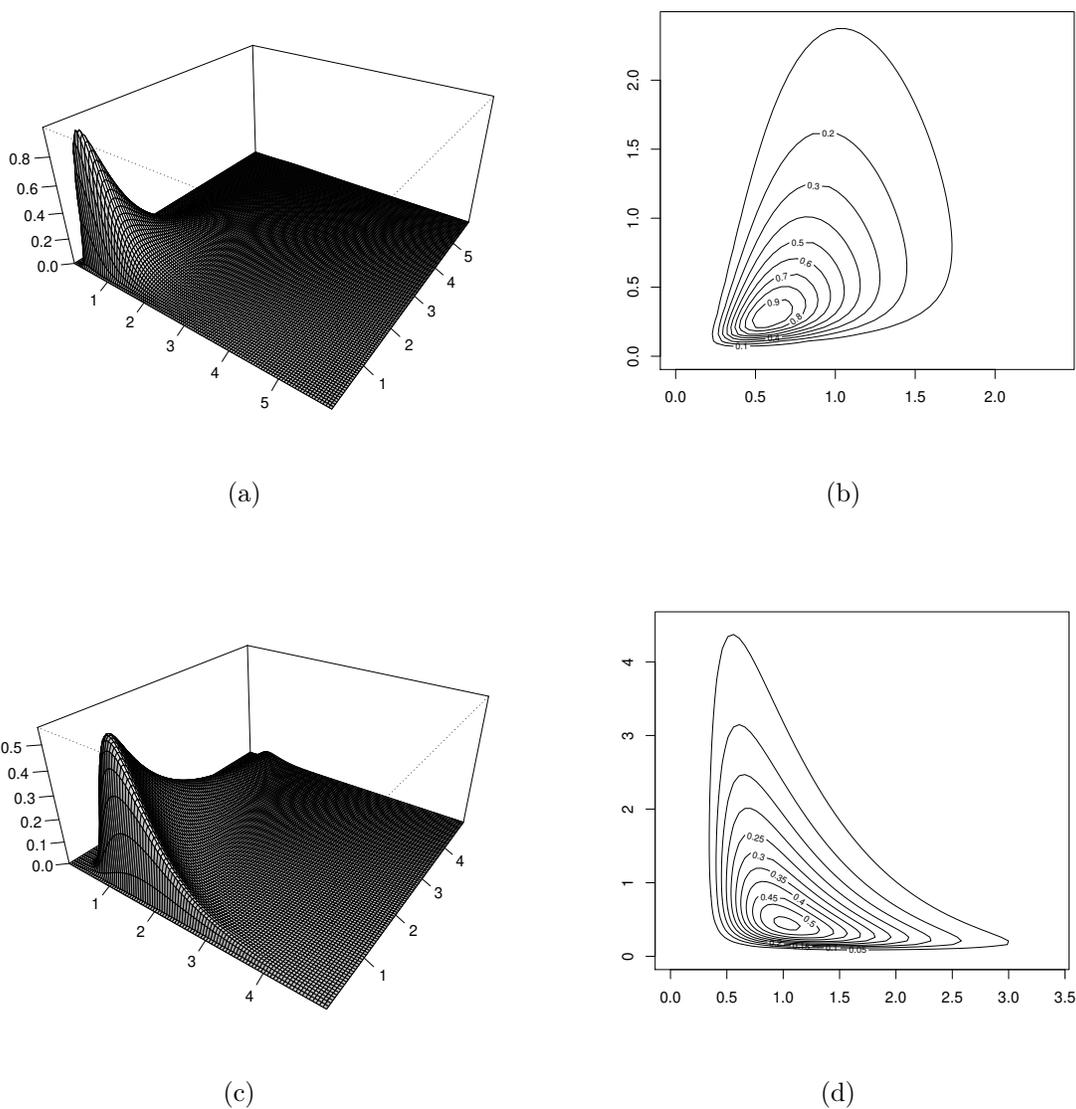


Figura 2.3: Fdp e contorno do modelo BS bivariado quando  $\alpha_1 = 0.5$ ,  $\alpha_2 = 1$ ,  $\beta_2 = \beta_2 = 1$ : (a) e (b)  $\rho = 0.5$ , (c) e (d)  $\rho = -0.5$

em que  $I_1(\rho)$  é definido como em 2.2.7.

**Demonstração.** As demonstrações de (i)-(ii) são diretamente obtidas a partir da definição do modelo  $\mathcal{BS}$ . A demonstração de (iii) é como segue:

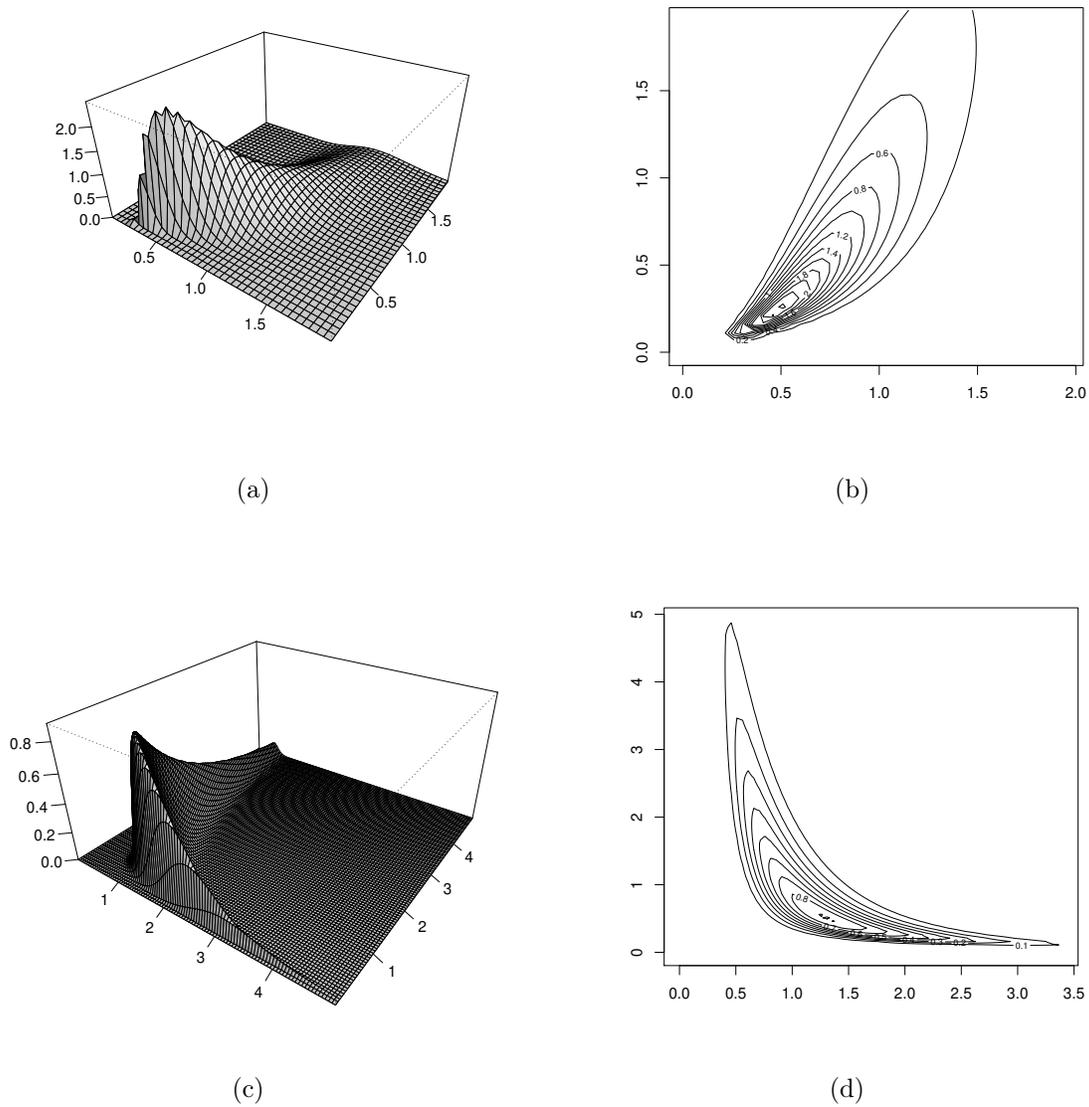


Figura 2.4: Fdp e contorno do modelo BS bivariado quando  $\alpha_1 = 0.5$ ,  $\alpha_2 = 1$ ,  $\beta_1 = \beta_2 = 1$ : (a) e (b)  $\rho = 0.9$ , (c) e (d)  $\rho = -0.9$

$$\begin{aligned}
 Cov[T_1, T_2] &= E(T_1 T_2) - E(T_1)E(T_2) \\
 &= (\beta_1 \beta_2) \left( 1 + (\alpha_1^2 + \alpha_2^2)/2 + \alpha_1^2 \alpha_2^2 (1 + \rho^2)/4 + \alpha_1 \alpha_2 I_1(\rho) \right) \\
 &\quad - \left[ \frac{\beta_1}{2} (2 + \alpha_1^2) \right] \left[ \frac{\beta_2}{2} (2 + \alpha_2^2) \right] \\
 &= \beta_1 \beta_2 \alpha_1 \alpha_2 \left[ \frac{1}{4} \alpha_1 \alpha_2 \rho^2 + I_1(\rho) \right]. \quad \blacksquare
 \end{aligned}$$

**Corolário 2.2.5.** *Seja  $\mathbf{T} = (T_1, T_2)^\top \sim BS_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ . Então o coeficiente de correlação entre  $T_1$  e  $T_2$  é dado por*

$$\rho_{BS_2} = \frac{\alpha_1 \alpha_2 \rho^2 + 4I_1(\rho)}{\sqrt{(4 + 5\alpha_1^2)(4 + 5\alpha_2^2)}},$$

em que  $I_1(\rho)$  é definido como em (2.2.7) e  $\rho$  é o coeficiente de correlação da normal bivariada.

## 2.3 Matriz de Informação de Fisher Observada

Nesta subsecção, vamos apresentar a matriz de informação observada baseada na fdp da  $BS$  bivariada definida em (2.2.2). Esta matriz será utilizada posteriormente para aplicar o método de influência local. Além disso, será utilizada para obter a matriz de informação de Fisher esperada, amplamente usada em procedimentos de inferência estatística baseada no método de máxima verossimilhança.

Suponha que temos  $n$  observações independentes,  $\mathbf{T}_1, \dots, \mathbf{T}_n$ , em que  $\mathbf{T}_i \sim BS_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ , para  $i = 1, \dots, n$ . Então, a função de verossimilhança para  $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \rho)^\top$ , dada a amostra observada  $\mathbf{t} = (\mathbf{t}_1, \dots, \mathbf{t}_n)^\top$  é da forma

$$L(\boldsymbol{\theta}; t_1, \dots, t_n) = \prod_{i=1}^n \phi_2(a(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho) A(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}),$$

e a função de log-verossimilhança pode ser escrita como

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}), \quad (2.3.1)$$

em que

$$\ell_i(\boldsymbol{\theta}) = -\log(2\pi) - \frac{1}{2} \log(1 - \rho^2) + \log(A(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta})) - \frac{1}{2} d(\mathbf{t}_i), \quad (2.3.2)$$

com  $d(\mathbf{t})$  como em (2.2.3).

### a) Derivadas parciais de primeira ordem da $\ell(\boldsymbol{\theta})$

Considerando a log-verossimilhança  $\ell(\boldsymbol{\theta})$  e derivando respeito a  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\beta}$  e  $\rho$ , temos que o vetor de

escore para  $\boldsymbol{\theta}$  é dada por  $U(\boldsymbol{\theta}) = (U_\alpha^\top(\boldsymbol{\theta}), U_\beta^\top(\boldsymbol{\theta}), U_\rho(\boldsymbol{\theta}))^\top$ , cujos elementos são

$$\begin{aligned} U_\alpha(\boldsymbol{\theta}) &= D^{-1}(\boldsymbol{\alpha}) \left\{ -n\mathbf{1}_2 + \sum_{i=1}^n D(a(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta})) \boldsymbol{\Sigma}^{-1} a(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}) \right\}, \\ U_\beta(\boldsymbol{\theta}) &= -\frac{1}{2} D^{-1}(\boldsymbol{\beta}) \sum_{i=1}^n [C(\mathbf{t}_i; \boldsymbol{\beta}) - B(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}) \boldsymbol{\Sigma}^{-1} a(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta})], \\ U_\rho(\boldsymbol{\theta}) &= \frac{n\rho}{1-\rho^2} - \frac{1}{2} \sum_{i=1}^n a^\top(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}) \boldsymbol{\Sigma}_\rho^{-1} a(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}), \end{aligned}$$

em que

$$C(\mathbf{t}_i; \boldsymbol{\beta}) = D^{-1}(\mathbf{t}_i + \boldsymbol{\beta})(\mathbf{t}_i - \boldsymbol{\beta}),$$

e

$$B(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}) = D^{-1/2}(\boldsymbol{\beta}) D^{-1}(\boldsymbol{\alpha}) D^{-1/2}(\mathbf{t}_i) D(\mathbf{t}_i + \boldsymbol{\beta}),$$

com

$$\boldsymbol{\Sigma}_\rho^{-1} = \partial \boldsymbol{\Sigma}^{-1} / \partial \rho = -\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_\rho \boldsymbol{\Sigma}^{-1}.$$

### b) Derivadas parciais de segunda ordem de $\ell(\boldsymbol{\theta})$

Seja  $\ell(\boldsymbol{\theta})$  a log-verossimilhança definida em (2.3.1). Então, a matriz de derivadas parciais de segunda ordem com respeito a  $\boldsymbol{\theta}$  é dada por

$$\ddot{\mathbf{L}} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \sum_{i=1}^n \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}. \quad (2.3.3)$$

As expressões de (2.3.3) pode-se achar no Apêndice C.

## 2.4 Matriz de Informação de Fisher Esperada

Nesta seção vamos obter a matriz de informação esperada da distribuição bivariada  $\mathcal{BS}$ . A obtenção de matriz de informação esperada será importante para testar as hipóteses de interesse apresentadas na Seção 2.6 e para estudar comportamento assintótico do estimador de máxima verossimilhança de  $\boldsymbol{\theta}$  ou de uma função diferenciável de  $\boldsymbol{\theta}$ .

A matriz de informação esperada de Fisher, que denotamos por  $\mathbf{I}(\boldsymbol{\theta})$ , é definida como

$$\mathbf{I}(\boldsymbol{\theta}) = \mathbb{E} \left( -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right) = -\sum_{i=1}^n \mathbb{E} \left( \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right), \quad (2.4.1)$$

pode ser escrita como

$$\mathbf{I}(\boldsymbol{\theta}) = -n \mathbb{E} \left( \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right),$$

pois as variáveis consideradas são identicamente distribuídas.

Para propósitos de inferência, a matriz de interesse será a matriz  $\mathbf{I}_F(\boldsymbol{\theta})$ , chamada de matriz de informação esperada de Fisher por elementos e definida por

$$\mathbf{I}_F(\boldsymbol{\theta}) = -\mathbb{E} \left( \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right). \quad (2.4.2)$$

Assim, a matriz de informação esperada de Fisher por elementos é dada por

$$\mathbf{I}_F(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{I}_{\alpha\alpha} & \mathbf{0} & \mathbf{I}_{\alpha\rho} \\ \mathbf{0} & \mathbf{I}_{\beta\beta} & \mathbf{0} \\ \mathbf{I}_{\alpha\rho}^\top & \mathbf{0} & \mathbf{I}_{\rho\rho} \end{pmatrix}, \quad (2.4.3)$$

em que os seus componentes podem ser expressos como

$$\begin{aligned} \mathbf{I}_{\alpha\alpha} &= \frac{2}{1-\rho^2} D^{-2}(\boldsymbol{\alpha}) - \frac{\rho^2}{1-\rho^2} D^{-1}(\boldsymbol{\alpha}) \mathbf{1}_2 \mathbf{1}_2^\top D^{-1}(\boldsymbol{\alpha}), \\ \mathbf{I}_{\alpha\rho} &= -\frac{\rho}{1-\rho^2} D^{-1}(\boldsymbol{\alpha}) \mathbf{1}_2, \\ \mathbf{I}_{\rho\rho} &= \frac{1+\rho^2}{(1-\rho^2)^2}, \\ \mathbf{I}_{\beta\beta} &= D^{-2}(\boldsymbol{\beta}) \left\{ \frac{1}{1-\rho^2} \left( \frac{\rho^2}{4} \mathbf{I}_2 + D^{-2}(\boldsymbol{\alpha}) \right) + D(\boldsymbol{\alpha}_J) \right\} \\ &\quad - \frac{\rho}{2(1-\rho^2)\alpha_1\alpha_2\beta_1\beta_2} (\psi_1(\boldsymbol{\alpha}^\top, \rho) + \psi_1(\boldsymbol{\alpha}^\top, -\rho)) (\mathbf{1}_2 \mathbf{1}_2^\top - \mathbf{I}_2), \end{aligned}$$

com  $\psi_1(\boldsymbol{\alpha}, \rho)$  e  $\psi_1(\boldsymbol{\alpha}, -\rho)$  sendo dadas em (2.2.7), respectivamente, e

$$\boldsymbol{\alpha}_J = (J(\alpha_1), J(\alpha_2))^\top,$$

com

$$J(\alpha_j) = \mathbb{E}[T_j/\beta_j + 1]^{-2}, \quad \text{para } j = 1, 2,$$

veja Kundu *et al.* (2010). Então, a inversa da matriz de informação de Fisher pode ser escrita como

$$\boldsymbol{\Omega}(\boldsymbol{\theta}) = \mathbf{I}_F^{-1}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{\Omega}_{\alpha\alpha} & \mathbf{0} & \boldsymbol{\Omega}_{\alpha\rho} \\ \mathbf{0} & \boldsymbol{\Omega}_{\beta\beta} & \mathbf{0} \\ \boldsymbol{\Omega}_{\rho\alpha} & \mathbf{0} & \boldsymbol{\Omega}_{\rho\rho} \end{pmatrix}, \quad (2.4.4)$$

em que

$$\begin{aligned}\Omega_{\alpha\alpha} &= \left( \mathbf{I}_{\alpha\alpha} - \frac{1}{\mathbf{I}_{\rho\rho}} \mathbf{I}_{\alpha\rho} \mathbf{I}_{\rho\alpha} \right)^{-1}, \\ \Omega_{\alpha\rho} &= -\frac{1}{\mathbf{I}_{\rho\rho}} \left( \mathbf{I}_{\alpha\alpha} - \frac{1}{\mathbf{I}_{\rho\rho}} \mathbf{I}_{\alpha\rho} \mathbf{I}_{\rho\alpha} \right)^{-1} \mathbf{I}_{\alpha\rho}, \\ \Omega_{\rho\rho} &= \left( \mathbf{I}_{\rho\rho} - \mathbf{I}_{\rho\alpha} \mathbf{I}_{\alpha\alpha}^{-1} \mathbf{I}_{\alpha\rho} \right)^{-1}.\end{aligned}$$

## 2.5 Estimação dos parâmetros

### 2.5.1 Estimação por Máxima Verossimilhança

Encontrar os EMV de  $\boldsymbol{\theta}$  por métodos convencionais é complexo, dado que as equações resultantes não apresentam solução explícita. Embora existam muitas metodologias para encontrar as EMV de  $\boldsymbol{\theta}$ , dessa forma, devem ser utilizados procedimentos iterativos de maximização numérica da verossimilhança. Nas funções `optimize()`, `optim()`, `nlm()` e `nlminb()` disponíveis em R para maximização numérica e das funções `fitdistr()` e `mle()` disponíveis para estimação por máxima verossimilhança, mais nesta dissertação usaremos a função `optim()`.

#### Estimação de Máxima Verossimilhança Irrestrita

Em seguida, obtém-se as EMV dos parâmetros da distribuição  $\mathcal{BS}$  bivariada sem considerar nenhuma restrição. Seja  $\mathbf{T}_1 = (T_{11}, T_{21})^\top, \dots, \mathbf{T}_n = (T_{1n}, T_{2n})^\top$  seja amostra aleatória de tamanho  $n$  desde  $\mathbf{T} \sim \mathcal{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$  e  $\ell(\boldsymbol{\theta})$  a função de log-verossimilhança para o vetor de parâmetros  $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \rho)^\top$  como determinada em (2.3.1). Denotamos por  $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\alpha}}^\top, \hat{\boldsymbol{\beta}}^\top, \hat{\rho})^\top$  as EMV de  $\boldsymbol{\theta}$ .

Maximizando a função de log-verossimilhança definida em (2.3.1), por meio da função `optim()` do R, obtemos as estimativas de  $\boldsymbol{\beta}$  e logo assim podemos obter as EMV de  $\boldsymbol{\alpha}$  e  $\rho$ ,

$$\hat{\alpha}_j^2 = \left( \frac{s_j}{\hat{\beta}_j} + \frac{\hat{\beta}_i}{r_j} - 2 \right)^{\frac{1}{2}}, \quad j = 1, 2, \quad \hat{\rho} = \frac{\sum_{i=1}^n a(t_{1i}; 1, \hat{\beta}_1) a(t_{2i}; 1, \hat{\beta}_2)}{\sqrt{\sum_{i=1}^n a^2(t_{1i}; 1, \hat{\beta}_1)} \sqrt{\sum_{i=1}^n a^2(t_{2i}; 1, \hat{\beta}_2)}},$$

$$\hat{\beta} = \arg \max_{\boldsymbol{\beta}} \ell(\hat{\boldsymbol{\alpha}}, \boldsymbol{\beta}, \hat{\rho}),$$

em que

$$s_j = \frac{1}{n} \sum_{i=1}^n T_{ji} \quad \text{e} \quad r_j = \left( \frac{1}{n} \sum_{i=1}^n T_{ji}^{-1} \right)^{-1}, \quad j = 1, 2.$$

Na Tabela 2.1, apresentamos uma simulação dos EMV para diferentes tamanhos de amostra, pode-se observar que as estimativas são próximas aos valores teóricos considerados. Além disso para os

diferentes tamanhos de amostra, os EP menores do que aos apresentados na Tabela 2 de Kundu *et al.* (2010).

Tabela 2.1: Médias das estimativas e erro quadrático médio (entre parênteses) de as EMV, quando  $\alpha_1 = \alpha_2 = 1$ ,  $\beta_1 = \beta_2 = 2$  e para diferentes valores de  $\rho$

n	$\rho$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\rho}$
10	0.95	0.9459 (0.0483)	0.9474 (0.0514)	1.9441 (0.1293)	1.9466 (0.1288)	0.9465 (0.0015)
	0.50	0.9232 (0.0490)	0.9434 (0.0478)	1.9422 (0.1315)	1.9452 (0.1285)	0.5132 (0.0654)
	0.25	0.9278 (0.0558)	0.9317 (0.0542)	1.9501 (0.1286)	1.9198 (0.1359)	0.2235 (0.0954)
	0.00	0.9312 (0.0529)	0.9419 (0.0563)	1.9400 (0.1339)	1.9584 (0.1241)	-0.0022 (0.0996)
20	0.95	0.9567 (0.0282)	0.9564 (0.0292)	1.9660 (0.0865)	1.9598 (0.0847)	0.9459 (0.0008)
	0.50	0.9592 (0.0271)	0.9638 (0.0233)	1.9612 (0.0930)	1.9653 (0.0935)	0.4917 (0.0295)
	0.25	0.9531 (0.0266)	0.9618 (0.0287)	1.9602 (0.0922)	1.9733 (0.0916)	0.2579 (0.0473)
	0.00	0.9585 (0.0253)	0.9580 (0.0253)	1.9689 (0.0895)	1.9888 (0.0814)	-0.0093 (0.0534)
50	0.95	0.9856 (0.0110)	0.9849 (0.0109)	1.9916 (0.0450)	1.9999 (0.0456)	0.9498 (0.0002)
	0.50	0.9852 (0.0106)	0.9902 (0.0098)	1.9955 (0.0469)	1.9898 (0.0454)	0.5054 (0.0118)
	0.25	0.9874 (0.0091)	0.9865 (0.0111)	1.9936 (0.0455)	1.9892 (0.0425)	0.2381 (0.0211)
	0.00	0.9894 (0.0107)	0.9875 (0.0096)	1.9997 (0.0482)	1.9940 (0.0421)	0.0039 (0.0196)
100	0.95	0.9947 (0.0056)	0.9959 (0.0055)	2.0096 (0.0294)	2.0098 (0.0299)	0.9496 (0.0001)
	0.50	0.9871 (0.0046)	0.9915 (0.0046)	2.0061 (0.0290)	2.0161 (0.0291)	0.4994 (0.0053)
	0.25	0.9917 (0.0050)	0.9907 (0.0045)	1.9946 (0.0292)	1.9946 (0.0256)	0.2450 (0.0098)
	0.00	0.9966 (0.0049)	0.9953 (0.0053)	2.0173 (0.0265)	1.9976 (0.0270)	0.0022 (0.0109)

### Estimação de Máxima Verossimilhança Restrita

Em seguida, obtém-se as EMV dos parâmetros da distribuição  $\mathcal{BS}$  bivariada considerando as seguintes restrições impostas pelas hipóteses nulas:

$$H_{01}: \alpha_1 = \alpha_2;$$

$$H_{02}: \beta_1 = \beta_2;$$

$$H_{03}: \alpha_1 = \alpha_2, \beta_1 = \beta_2; \text{ e}$$

$$H_{04}: \rho = \rho_0.$$

Seja  $\mathbf{T}_1 = (T_{11}, T_{21})^\top, \dots, \mathbf{T}_n = (T_{1n}, T_{2n})^\top$  uma amostra aleatória de tamanho  $n$  de  $\mathbf{T} \sim \mathcal{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$  e  $\ell(\boldsymbol{\theta})$  a log-verossimilhança para  $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \rho)^\top$ . Denotamos por  $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\alpha}}^\top, \tilde{\boldsymbol{\beta}}^\top, \tilde{\rho})^\top$  as EMV  $\boldsymbol{\theta}$  sob as hipóteses nulas. Então,

(H<sub>01</sub>) Sendo  $\alpha = \alpha_1 = \alpha_2$ , temos que

Neste caso,  $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top, \rho)^\top$ . Logo, segue que o logaritmo da verossimilhança é dado por

$$\ell_i(\boldsymbol{\theta}) = -\log(2\pi) - \frac{1}{2} \log(1 - \rho^2) + \log(A(\mathbf{t}_i; \alpha \mathbf{1}_2, \boldsymbol{\beta})) - \frac{1}{2} d(\mathbf{t}_i),$$

em que

$$d(\mathbf{t}) = d(\mathbf{t}, \boldsymbol{\theta}) = a(\mathbf{t}; \alpha, \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} a(\mathbf{t}; \alpha, \boldsymbol{\beta}),$$

e

$$A(\mathbf{t}; \alpha \mathbf{1}_2, \boldsymbol{\beta}) = A(t_1; \alpha, \beta_1) A(t_2; \alpha, \beta_2),$$

com

$$A(t_j; \alpha, \beta_j) = t_j^{-3/2} (t_j + \beta_j) / (2\alpha\beta_j^{1/2}), \text{ para } j = 1, 2.$$

As equações para estimar  $\boldsymbol{\alpha}$  e  $\rho$  apresentam soluções explícitas, sendo assim maximizando a equação de  $\boldsymbol{\beta}$  por meio da função `optim()` do R, podemos obter as EMV de  $\boldsymbol{\alpha}$  e  $\rho$ , as quais

$$\tilde{\alpha}^2 = \frac{1}{2n} \sum_{i=1}^n \left( a^2(t_{1i}; 1, \tilde{\beta}_1) + a^2(t_{2i}; 1, \tilde{\beta}_2) \right), \quad \tilde{\rho} = \frac{2 \sum_{i=1}^n a(t_{1i}; 1, \tilde{\beta}_1) a(t_{2i}; 1, \tilde{\beta}_2)}{\sum_{i=1}^n (a^2(t_{1i}; 1, \tilde{\beta}_1) + a^2(t_{2i}; 1, \tilde{\beta}_2))},$$

$$\tilde{\boldsymbol{\beta}} = \arg \max_{\boldsymbol{\beta}} \ell(\tilde{\boldsymbol{\alpha}}, \boldsymbol{\beta}, \tilde{\rho}).$$

(H<sub>02</sub>) Sendo  $\beta = \beta_1 = \beta_2$ , temos que

Neste caso,  $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \beta, \rho)^\top$ . Logo, segue que o logaritmo da verossimilhança é dado por

$$\ell_i(\boldsymbol{\theta}) = -\log(2\pi) - \frac{1}{2} \log(1 - \rho^2) + \log(A(\mathbf{t}_i; \boldsymbol{\alpha}, \beta)) - \frac{1}{2} d(\mathbf{t}_i),$$

em que

$$d(\mathbf{t}) = d(\mathbf{t}, \boldsymbol{\theta}) = a(\mathbf{t}; \boldsymbol{\alpha}, \beta \mathbf{1}_2)^\top \boldsymbol{\Sigma}^{-1} a(\mathbf{t}; \boldsymbol{\alpha}, \beta),$$

e

$$A(\mathbf{t}; \boldsymbol{\alpha}, \beta \mathbf{1}_2) = A(t_1; \alpha_1, \beta) A(t_2; \alpha_2, \beta),$$

com

$$A(t_j; \alpha_j, \beta) = t_j^{-3/2} (t_j + \beta) / (2\alpha_j \beta^{1/2}) \quad , \text{ para } j = 1, 2.$$

As equações para estimar  $\boldsymbol{\alpha}$  e  $\rho$  apresentam soluções explícitas, sendo assim maximizando a equação de  $\beta$  por meio da função `optim()` do R, podemos obter as EMV de  $\boldsymbol{\alpha}$  e  $\rho$  para este modelo, dados por

$$\tilde{\alpha}_j^2 = \frac{1}{n} \sum_{i=1}^n a_{t_{ji}}^2(1, \tilde{\beta}), j = 1, 2, \quad \tilde{\rho} = \frac{\sum_{i=1}^n a(t_{1i}; 1, \tilde{\beta}) a(t_{2i}; 1, \tilde{\beta})}{\sqrt{\sum_{i=1}^n a^2(t_{1i}; 1, \tilde{\beta})} \sqrt{\sum_{i=1}^n a^2(t_{2i}; 1, \tilde{\beta})}},$$

$$\tilde{\beta} = \arg \max_{\beta} \ell(\tilde{\boldsymbol{\alpha}}, \beta \mathbf{1}_2, \tilde{\rho}).$$

(H<sub>03</sub>) Sendo  $\alpha = \alpha_1 = \alpha_2$  e  $\beta = \beta_1 = \beta_2$ , temos que

Neste caso,  $\boldsymbol{\theta} = (\alpha, \beta, \rho)^\top$ . Logo, segue que o logaritmo da verossimilhança é dado por

$$l_i(\boldsymbol{\theta}) = -\log(2\pi) - \frac{1}{2} \log(1 - \rho^2) + \log(A(\mathbf{t}_i; \alpha \mathbf{1}_2, \beta \mathbf{1}_2)) - \frac{1}{2} d(\mathbf{t}_i),$$

em que

$$d(\mathbf{t}) = d(\mathbf{t}, \boldsymbol{\theta}) = a(\mathbf{t}; \alpha \mathbf{1}_2, \beta \mathbf{1}_2)^\top \boldsymbol{\Sigma}^{-1} a(\mathbf{t}; \alpha, \beta),$$

e

$$A(\mathbf{t}; \alpha \mathbf{1}_2, \beta \mathbf{1}_2) = A(t_1; \alpha, \beta) A(t_2; \alpha, \beta),$$

com

$$A(t_j; \alpha, \beta) = t_j^{-3/2} (t_j + \beta) / (2\alpha \beta^{1/2}) \quad , \text{ para } j = 1, 2.$$

As equações para estimar  $\boldsymbol{\alpha}$  e  $\rho$  apresentam soluções explícitas, sendo assim maximizando a equação de  $\beta$  por meio da função `optim()` do R, podemos obter as EMV de  $\boldsymbol{\alpha}$  e  $\rho$ , as quais

$$\tilde{\alpha}^2 = \frac{1}{2n} \sum_{i=1}^n \left( a^2(t_{1i}; 1, \tilde{\beta}) + a^2(t_{2i}; 1, \tilde{\beta}) \right), \quad \tilde{\rho} = \frac{2 \sum_{i=1}^n a(t_{1i}; 1, \tilde{\beta}) a(t_{2i}; 1, \tilde{\beta})}{\sum_{i=1}^n (a^2(t_{1i}; 1, \tilde{\beta}) + a^2(t_{2i}; 1, \tilde{\beta}))},$$

$$\tilde{\beta} = \arg \max_{\beta} \ell(\tilde{\boldsymbol{\alpha}} \mathbf{1}_2, \beta \mathbf{1}_2, \tilde{\rho}).$$

(H<sub>04</sub>) Sendo  $\rho = \rho_0$ , nos temos

Neste caso,  $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \rho_0)^\top$ . Logo, segue que o logaritmo da verossimilhança é dado por

$$\ell_i(\boldsymbol{\theta}) = -\log(2\pi) - \frac{1}{2}\log(1 - \rho_0^2) + \log(A(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta})) - \frac{1}{2}d(\mathbf{t}_i),$$

em que

$$d(\mathbf{t}) = d(\mathbf{t}, \boldsymbol{\theta}) = a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta})^\top \boldsymbol{\Sigma}^{-1} a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}),$$

e

$$A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = A(t_1; \alpha_1, \beta_1)A(t_2; \alpha_2, \beta_2),$$

com

$$A(t_j; \alpha_j, \beta_j) = t_j^{-3/2}(t_j + \beta_j)/(2\alpha_j\beta_j^{1/2}), \text{ para } j = 1, 2.$$

A equação para estimar  $\boldsymbol{\alpha}$  é fechada, sendo assim maximizando a equação de  $\boldsymbol{\beta}$  por meio da função `optim()` do R, podemos obter as EMV de  $\boldsymbol{\alpha}$ , os quais

$$\tilde{\alpha}_j^2 = \frac{1}{n} \sum_{i=1}^n a^2(t_{ji}; 1, \tilde{\beta}_j), \quad j = 1, 2, \quad \tilde{\boldsymbol{\beta}} = \arg \max_{\boldsymbol{\beta}} \ell(\tilde{\boldsymbol{\alpha}}, \boldsymbol{\beta}, \rho_0).$$

Em todos os casos, para estimar  $\boldsymbol{\beta}$  é necessário utilizar um processo numérico iterativo, mas os EMV de  $\boldsymbol{\alpha}$  e  $\rho$  tem expressões fechadas para  $\boldsymbol{\beta}$  fixo.

## 2.5.2 Estimação pelo Método dos Momentos

Na caso da distribuição clássica, tem-se que os estimadores obtidos pelo método dos momentos convencional para  $\alpha$  e  $\beta$  nem sempre existem e, quando eles existem podem não ser únicos, veja Ng *et al.* (2003). Diante disto, os autores acima propuseram o método dos momentos modificado. Tal método consiste em igualar  $\mathbb{E}(T)$  e  $\mathbb{E}(T^{-1})$  aos seus respectivos momentos amostrais, para obter os estimadores  $\alpha$  e  $\beta$ . Seguimos essa proposta para o caso do modelo  $\mathcal{BS}$  bivariada.

Sejam  $\mathbf{T}_1, \dots, \mathbf{T}_n$  uma amostra aleatória de tamanho  $n$  de uma distribuição  $\mathcal{BS}$  bivariada com densidade dada em (2.2.2). Também como é bem conhecido, o estimador pelo método dos momentos pode não ser único, e que, além disso, pode não existir. Uma forma alternativa para estimar o parâmetro do modelo  $\mathcal{BS}$  bivariado de distribuição é o método modificado de momento (MM), veja Ng *et al.* (2003), Ng *et al.* (2006), Leiva *et al.* (2008) e Kundu *et al.* (2010). O Teorema (2.5.2) apresenta sua distribuição assintótica, que pode ser utilizada para construir os intervalos de confiança e testes de hipótese para os parâmetros.

Seja  $\mathbf{T}_1, \dots, \mathbf{T}_n$  uma amostra aleatória de  $\mathbf{T} \sim \mathcal{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ . Então, os estimadores pelo método MM para os vetores  $\boldsymbol{\alpha}$  e  $\boldsymbol{\beta}$ , denotados respectivamente por  $\tilde{\boldsymbol{\alpha}}$  e  $\tilde{\boldsymbol{\beta}}$ , possuem elementos obtidos pelas

equações de certos momentos amostrais e populacionais, que é,  $E[T_j]$  a  $S_j$ , e  $E[T_j^{-1}]$  a  $R_j$ , para  $j = 1, 2$ , dado por

$$\tilde{\alpha}_{Mj} = \left(2 \left( (S_j/R_j)^{1/2} - 1 \right)\right)^{1/2} \quad \text{e} \quad \tilde{\beta}_{Mj} = (S_j R_j)^{1/2}, \quad (2.5.1)$$

em que  $S_j = \frac{1}{n} \sum_{i=1}^n T_{ji}$  e  $R_j = \left(\frac{1}{n} \sum_{i=1}^n T_{ji}^{-1}\right)^{-1}$ , para  $j = 1, 2$ . Uma vez que os estimadores obtidos pelo método dos MM são explícitos, eles podem ser utilizados de forma eficaz em situações práticas, bem como valores iniciais do processo iterativo para calcular os EMV. O seguinte resultado fornece uma ferramenta para estudar a distribuição assintótica dos estimadores apresentados no (2.5.1).

**Lema 2.5.1.** *Seja  $\mathbf{T} \sim \text{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ . Então,*

$$\begin{aligned} \text{Var}[\mathbf{T}] &= D^2(\boldsymbol{\alpha})D^2(\boldsymbol{\beta})\left(\mathbf{I}_2 + \frac{5}{4}D^2(\boldsymbol{\alpha})\right) + \beta_1\beta_2A(\boldsymbol{\alpha}, \rho)\mathbf{W}, \\ \text{Var}[\mathbf{T}^{-1}] &= D^2(\boldsymbol{\alpha})D^{-2}(\boldsymbol{\beta})\left(\mathbf{I}_2 + \frac{5}{4}D^2(\boldsymbol{\alpha})\right) + \frac{1}{\beta_1\beta_2}A(\boldsymbol{\alpha}, \rho)\mathbf{W}, \\ \text{Cov}[\mathbf{T}, \mathbf{T}^{-1}] &= -D^2(\boldsymbol{\alpha})\left(\frac{1}{4}D^2(\boldsymbol{\alpha}) + \mathbf{I}_2\right) + A(\boldsymbol{\alpha}, -\rho)D(\boldsymbol{\beta})\mathbf{W}D^{-1}(\boldsymbol{\beta}), \end{aligned}$$

em que  $A(\boldsymbol{\alpha}, \rho) = (\psi_1^*(\boldsymbol{\alpha}, \rho) - (1 + \alpha_1^2/2)(1 + \alpha_2^2/2))$ , com  $\psi_1^*(\boldsymbol{\alpha}, \rho)$  sendo dadas em (2.2.7), e  $\mathbf{W}$  é  $2 \times 2$  matriz simétrica, com  $w_{12} = w_{21} = 1$  e  $w_{11} = w_{22} = 0$ .

O Teorema 2.5.2 fornece a distribuição assintótica dos estimadores  $\tilde{\boldsymbol{\alpha}}$  e  $\tilde{\boldsymbol{\beta}}$  dado em (2.5.1), que, como já foi dito, pode ser usado para construir os intervalos de confiança e testes de hipótese para estes parâmetros.

**Teorema 2.5.2.** *Seja  $\mathbf{T}_1, \dots, \mathbf{T}_n$  uma amostra aleatória de  $\mathbf{T} \sim \text{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ , para  $\rho$  fixo. Então, da distribuição assintótica conjunta dos estimadores pelo método MM  $\tilde{\boldsymbol{\alpha}}_M$  e  $\tilde{\boldsymbol{\beta}}_M$  é dada por*

$$\sqrt{n} \left( \begin{pmatrix} \tilde{\boldsymbol{\alpha}}_M \\ \tilde{\boldsymbol{\beta}}_M \end{pmatrix} - \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} \right) \xrightarrow{d} N_4 \left( \mathbf{0}, \begin{pmatrix} \boldsymbol{\Upsilon}_{\alpha\alpha} & 0 \\ 0 & \boldsymbol{\Upsilon}_{\beta\beta} \end{pmatrix} \right),$$

em que  $\xrightarrow{d}$  denota “convergência na distribuição”,  $\boldsymbol{\Upsilon}_{\alpha\alpha} = \frac{1}{2}(D(\boldsymbol{\alpha})\Sigma_\rho^*D(\boldsymbol{\alpha}) + \{I_1(\rho) + I_1(-\rho)\}\mathbf{W})$ , com  $I_1(\cdot)$  sendo dadas em (2.2.9), e  $\boldsymbol{\Upsilon}_{\beta\beta}$  tem elementos

$$\gamma_{11} = (\alpha_1\beta_1)^2 \frac{(3\alpha_1^2 + 4)}{(\alpha_1^2 + 2)^2}, \gamma_{12} = \frac{2\beta_1\beta_2}{(\alpha_1^2 + 2)(\alpha_2^2 + 2)} (I_1(\rho) - I_1(-\rho)), \gamma_{22} = (\alpha_2\beta_2)^2 \frac{(3\alpha_2^2 + 4)}{(\alpha_2^2 + 2)^2}.$$

**Demonstração.** Seja

$$\mathbf{S} = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \quad \text{e} \quad \mathbf{P} = \begin{pmatrix} 1/R_1 \\ 1/R_2 \end{pmatrix},$$

com  $S_j$  e  $R_j$  sendo dadas em (2.5.1).

Assim, usando o teorema limite central, temos que

$$\sqrt{n} \left( \begin{pmatrix} \mathbf{S} \\ \mathbf{P} \end{pmatrix} - \begin{pmatrix} \mathbf{E}(\mathbf{T}) \\ \mathbf{E}(\mathbf{T}^{-1}) \end{pmatrix} \right) \xrightarrow{d} N_4 \left( \mathbf{0}, \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{pmatrix} \right),$$

em que  $\Psi_{11} = \text{Var}[\mathbf{T}]$ ,  $\Psi_{12} = \text{Cov}[\mathbf{T}, \mathbf{T}^{-1}]$ ,  $\Psi_{21} = \Psi_{12}^\top$  e  $\Psi_{22} = \text{Var}[\mathbf{T}^{-1}]$  dada no Lema 2.5.1. Finalmente, nota-se a partir de (2.5.1) que  $\tilde{\alpha}_{M1} = g_1(S_1, 1/R_1)$ ,  $\tilde{\alpha}_{M2} = g_2(S_2, 1/R_2)$ ,  $\tilde{\beta}_{M1} = g_3(S_1, 1/R_1)$  e  $\tilde{\beta}_{M2} = g_4(S_2, 1/R_2)$  e aplicando o método delta, o teorema está demonstrado. ■

**Corolário 2.5.3.** *A distribuição de  $\tilde{\alpha}$  e  $\tilde{\beta}$  é, respectivamente,*

$$\sqrt{n}(\tilde{\alpha}_M - \alpha) \xrightarrow{d} N_2(\mathbf{0}, \Omega_{\alpha\alpha}) \text{ e } \sqrt{n}(\tilde{\beta}_M - \beta) \xrightarrow{d} N_2(\mathbf{0}, \Omega_{\beta\beta}), \text{ as } n \rightarrow \infty.$$

Corolário 2.5.3 pode ser usado para a construção de intervalos de confiança para os parâmetros  $\alpha$  e  $\beta$ .

Outra maneira de obter um estimador do tipo método dos momentos é usar a relação entre a distribuição  $\mathcal{BS}$  e distribuições normal bivariada. No seguinte teorema é apresentado um estimador do tipo método dos momentos de  $\alpha$  e  $\rho$ , e sua consistência e distribuição assintótica, supondo que se tem estimadores consistentes  $\beta_1$  e  $\beta_2$ .

**Teorema 2.5.4.** *Considere a distribuição  $\mathcal{BS}$  bivariada definida em (2.2.2) e suponha que existem estimadores consistentes  $\check{\beta}_1$  e  $\check{\beta}_2$  de  $\beta_1$  e  $\beta_2$ , respectivamente. Então, os estimadores de  $\alpha$  e  $\rho$  são dados por*

$$\check{\alpha} = \begin{pmatrix} \sqrt{\frac{1}{n} \sum_{i=1}^n a^2(t_{1i}; 1, \check{\beta}_1)} \\ \sqrt{\frac{1}{n} \sum_{i=1}^n a^2(t_{2i}; 1, \check{\beta}_2)} \end{pmatrix},$$

e

$$\check{\rho} = \frac{\sum_{j=1}^n a(t_{1i}; 1, \check{\beta}_1) a(t_{2i}; 1, \check{\beta}_2)}{\sqrt{\sum_{j=1}^n a^2(t_{1i}; 1, \check{\beta}_1) \sum_{j=1}^n a^2(t_{2i}; 1, \check{\beta}_2)}},$$

respectivamente, são consistentes e assintoticamente normal, isto é, para  $n \rightarrow \infty$ ,

(i)  $\check{\alpha} \xrightarrow{P} \alpha$  e  $\check{\rho} \xrightarrow{P} \rho$ ;

(ii)

$$\sqrt{n} \begin{pmatrix} \check{\alpha} - \alpha \\ \check{\rho} - \rho \end{pmatrix} \xrightarrow{d} N_3(\mathbf{0}, \Omega_{\alpha\rho}),$$

em que

$$\Omega_{\alpha\rho} = \begin{pmatrix} \frac{1}{2}\alpha_1^2 & \frac{1}{2}\alpha_1\alpha_2\rho^2 & \frac{1}{2}\alpha_1\rho(1-\rho^2) \\ \frac{1}{2}\alpha_1\alpha_2\rho^2 & \frac{1}{2}\alpha_2^2 & \frac{1}{2}\alpha_2\rho(1-\rho^2) \\ \frac{1}{2}\alpha_1\rho(1-\rho^2) & \frac{1}{2}\alpha_2\rho(1-\rho^2) & (1-\rho^2)^2 \end{pmatrix}.$$

**Demonstração.** Primeiro, consideremos  $\beta_1$  e  $\beta_2$  fixos, e definimos a matriz de covariância  $2 \times 2$  de  $\mathbf{W}_i$  como  $\mathbf{S} = S_n(\beta_1, \beta_2) = \frac{1}{n} \sum_{i=1}^n \mathbf{W}_i \mathbf{W}_i^\top$ , com as suas entradas denotadas por  $S_{11}$ ,  $S_{12}$  e  $S_{22}$ . Podemos estudar as propriedades conjunta da matriz de covariância  $\mathbf{S}$  de ordem  $2 \times 2$ , e através deles as propriedades de  $\check{\alpha}_j = g_j(\mathbf{S}) = \sqrt{S_{jj}}$ , para  $j = 1, 2$ , e  $\check{\rho} = g_3(\mathbf{S}) = S_{12}/\sqrt{S_{11}S_{22}}$ .

Uma vez que a distribuição assintótica conjunta é

$$\sqrt{n} \begin{pmatrix} S_{11} - \alpha_1^2 \\ S_{12} - \alpha_1\alpha_2\rho \\ S_{22} - \alpha_2^2 \end{pmatrix} \xrightarrow{d} N_3(\mathbf{0}, \mathbf{\Upsilon}), \quad \text{onde } \mathbf{\Upsilon} = \begin{pmatrix} 2\alpha_1^4 & 2\alpha_1^3\alpha_2\rho & 2\alpha_1^2\alpha_2^2\rho^2 \\ 2\alpha_1^3\alpha_2\rho & \alpha_1^2\alpha_2^2 + \alpha_1^2\alpha_2^2\rho^2 & 2\alpha_1\alpha_2^3\rho \\ 2\alpha_1^2\alpha_2^2\rho^2 & 2\alpha_1\alpha_2^3\rho & 2\alpha_2^4 \end{pmatrix},$$

a consistência dos estimadores  $\check{\alpha}$  e  $\check{\rho}$  resulta do fato de que  $\mathbf{S}$  tende a  $\mathbf{\Upsilon}$  com probabilidade um. Agora, a aplicação do método delta, obtemos a distribuição assintótica conjunta de  $\check{\alpha}_i$ , para  $i = 1, 2$ , e  $\check{\rho}$ . O resultado necessário, em seguida, segue para  $\beta_1$  e  $\beta_2$  fixo. Finalmente, substituindo  $\beta_1$  e  $\beta_2$  com os seus respectivos estimadores consistentes  $\check{\beta}_1$  e  $\check{\beta}_2$  nos resultados obtidos e, em seguida, aplicar as propriedades de convergência quase certa e o método delta, a demonstração do teorema está completa. ■

## 2.6 Testes de Hipóteses

Nesta seção, usamos os testes de hipótese baseado na Razão de Estatística de Verossimilhança, Escore e Wald, em que  $H_{01}$ ,  $H_{02}$ ,  $H_{03}$  e  $H_{04}$  dados na Seção 2.5. As distribuições exatas dessas estatísticas são desconhecidas, porém vamos trabalhar com as distribuições assintóticas. Devido ao relacionamento entre a  $\mathcal{BS}$  bivariada e a distribuição normal bivariada, a verossimilhança correspondente satisfaz as condições de regularidade padrão; veja Cox & Hinkley (1974). Assim, todas estas estatísticas convergem para uma distribuição qui-quadrado com  $r_k$  graus de liberdade sob a hipótese nula. Denotamos por  $\hat{\boldsymbol{\theta}}$  a EMV de  $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \rho)$  sob o modelo irrestrito e por  $\tilde{\boldsymbol{\theta}}_k$  a EMV sob o modelo restrito  $H_{0k}$ , para  $k = 1, 2, 3, 4$ .

Estas estatísticas foram para algumas suposições sobre o vetor de parâmetros  $\boldsymbol{\theta}$  no modelo  $\mathcal{BS}$  bivariado definido em (2.2.2).

$$H_0 : \mathbf{C}\boldsymbol{\theta} = \mathbf{q}_0,$$

em que  $\mathbf{C}$  é uma matriz de dimensão  $l \times (2p + 1)$ ,  $l$  é o posto de  $\mathbf{C}$  e  $\mathbf{q}_0$  é um vetor conhecido de dimensão  $l$ . Logo as estatísticas  $Q$ ,  $E$  e  $W$  podem ser escritas, respectivamente, como

$$Q = 2 \left[ \ell(\hat{\boldsymbol{\theta}}) - \ell(\tilde{\boldsymbol{\theta}}) \right], \quad E = \frac{1}{n} \mathbf{S}(\tilde{\boldsymbol{\theta}})^\top \mathbf{I}_F^{-1}(\tilde{\boldsymbol{\theta}}) \mathbf{S}(\tilde{\boldsymbol{\theta}}),$$

$$W = n (\mathbf{C}\hat{\boldsymbol{\theta}} - \mathbf{q}_0)^\top (\mathbf{C}^\top \mathbf{I}_F^{-1} \mathbf{C}) (\mathbf{C}\hat{\boldsymbol{\theta}} - \mathbf{q}_0),$$

em que  $\ell(\boldsymbol{\theta})$  é a função de log-verossimilhança definida em (2.3.2) e  $\mathbf{I}_F$  é a matriz de informação de Fisher por elementos esperada definida em (2.4.2).

Sob  $H_0$  as três estatísticas definidas acima convergem em distribuição para uma distribuição qui-quadrado com  $q = p - 1$ ,  $p = \text{número do parâmetros}$ , graus de liberdade.

### 2.6.1 Estatística da razão de verossimilhança

Para cada hipótese  $H_{0k}$  a estatística de Razão de Verossimilhança é expressa como

$$Q_{0k} = 2\left(\ell(\hat{\boldsymbol{\theta}}) - \ell(\tilde{\boldsymbol{\theta}}_k)\right) = n \log \left( \frac{1 - \tilde{\rho}_k^2}{1 - \hat{\rho}_0^2} \right) - \sum_{i=1}^n \left( \hat{d}(\mathbf{t}_i) - \tilde{d}_k(\mathbf{t}_i) \right) + 2 \sum_{i=1}^n \log \left( \frac{A(\mathbf{t}_i; \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}})}{A(\mathbf{t}_i; \tilde{\boldsymbol{\alpha}}_k, \tilde{\boldsymbol{\beta}}_k)} \right), \quad (2.6.1)$$

em que  $\hat{d}(\mathbf{t}_i) = d(\mathbf{t}_i, \hat{\boldsymbol{\theta}})$  e  $\tilde{d}_k(\mathbf{t}_i) = d(\mathbf{t}_i, \tilde{\boldsymbol{\theta}}_k)$ . Note que os dois primeiros termos do lado direito de (2.6.1) são análogos ao caso normal bivariado;

### 2.6.2 Estatística de Escore

A estatística de Escore para cada hipótese  $H_{0k}$  é dada por

$$S_{0k} = \frac{1}{n} \tilde{\mathbf{U}}_k^\top \mathbf{I}_F^{-1}(\tilde{\boldsymbol{\theta}}_k) \tilde{\mathbf{U}}_k = \frac{1}{n} \tilde{\mathbf{U}}_k^\top \tilde{\boldsymbol{\Omega}}_k \tilde{\mathbf{U}}_k, \quad k = 1, 2, 3, 4, \quad (2.6.2)$$

em que  $\tilde{\boldsymbol{\Omega}}_k = \boldsymbol{\Omega}(\tilde{\boldsymbol{\theta}}_k)$  e  $\tilde{\mathbf{U}}_k = \mathbf{U}(\tilde{\boldsymbol{\theta}}_k) = (\tilde{U}_{\alpha,k}^\top, \tilde{U}_{\beta,k}^\top, \tilde{U}_{\rho,k}^\top)^\top$ . Pela consideração da estrutura da inversa de matriz de informação esperada por elementos dado em (2.4.4), podemos expressar a estatística dado em (2.6.2) como

$$S_{0k} = \frac{1}{n} \left( \tilde{U}_{\rho,k}^2 \tilde{\boldsymbol{\Omega}}_{\rho\rho,k} + 2\tilde{U}_{\rho,k} \tilde{\boldsymbol{\Omega}}_{\rho\alpha,k} \tilde{U}_{\alpha,k} + \tilde{U}_{\alpha,k}^\top \tilde{\boldsymbol{\Omega}}_{\alpha\alpha,k} \tilde{U}_{\alpha,k} + \tilde{U}_{\beta,k}^\top \tilde{\boldsymbol{\Omega}}_{\beta\beta,k} \tilde{U}_{\beta,k} \right), \quad k = 1, 2, 3, 4.$$

Para a estatística de Escore, é fácil obter os parâmetros estimados sobre o modelo irrestrito  $H_{0k}$ . Outras hipóteses de interesse são

$$H_{05}: \boldsymbol{\alpha} = \boldsymbol{\alpha}_0 \text{ e } H_{06}: \boldsymbol{\beta} = \boldsymbol{\beta}_0,$$

e nestes casos as estatísticas de escore são

$$S_{05} = \frac{1}{n} U_\alpha^\top(\tilde{\boldsymbol{\theta}}_5) \boldsymbol{\Omega}_{\alpha\alpha} U_\alpha(\tilde{\boldsymbol{\theta}}_5) \text{ e } S_{06} = \frac{1}{n} U_\beta^\top(\tilde{\boldsymbol{\theta}}_6) \boldsymbol{\Omega}_{\beta\beta} U_\beta(\tilde{\boldsymbol{\theta}}_6),$$

respectivamente.

### 2.6.3 Estatística de Wald

Note que as quatro hipóteses de interesse podem ser reescritas como  $H_{0k}: \mathbf{A}\boldsymbol{\theta} = \mathbf{q}_{0k}$ , para  $k = 1, 2, 3, 4$ , em que  $\mathbf{A}_k$  e  $\mathbf{q}_{0k}$  são matrizes adequadas. Especificamente, note que para as quatro hipóteses, temos

$$(H_{01}) \quad \mathbf{A}_1 = (1, -1, 0, 0, 0) \text{ e } \mathbf{q}_{01} = 0.$$

$$(H_{02}) \quad \mathbf{A}_2 = (0, 0, 1, -1, 0) \text{ e } \mathbf{q}_{02} = 0.$$

$$(H_{03}) \quad \mathbf{A}_3 = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix} \text{ e } \mathbf{q}_{03} = \mathbf{0}_2.$$

(H<sub>04</sub>)  $A_4 = (0, 0, 0, 0, 1)$  e  $q_{04} = \rho_0$ .

Assim, a estatística de Wald para testar H<sub>0k</sub> é dada por

$$W_{0k} = n(\mathbf{A}_k \hat{\boldsymbol{\theta}})^\top (\mathbf{A}_k \mathbf{I}_F^{-1}(\hat{\boldsymbol{\theta}}) \mathbf{A}_k^\top)^{-1} (\mathbf{A}_k \hat{\boldsymbol{\theta}}), \quad k = 1, 2, 3, 4,$$

em que  $\mathbf{I}_F^{-1}(\boldsymbol{\theta}) = \boldsymbol{\Omega}$  é a inversa da matriz de informação esperada por elementos dado em (2.4.2). Por tanto, a estatística de Wald para testar cada conjunto de hipóteses H<sub>0k</sub> pode ser expressa, respectivamente, por

$$W_{01} = n(\alpha_1 - \alpha_2)^2 (\mathbf{a}^\top \hat{\boldsymbol{\Omega}}_{\alpha\alpha} \mathbf{a})^{-1},$$

$$W_{02} = n(\beta_1 - \beta_2)^2 (\mathbf{a}^\top \hat{\boldsymbol{\Omega}}_{\beta\beta} \mathbf{a})^{-1},$$

$$W_{03} = W_{01} + W_{02},$$

e

$$W_{04} = n(\hat{\rho} - \rho_0)^2 (\hat{\boldsymbol{\Omega}}_{\rho\rho})^{-1}.$$

## 2.7 Curva TTT

A curva TTT (Tempo Total em Teste) possibilita a análise gráfica do comportamento do tempo de sobrevivência. Proposta por Aarset (1987) e estudada por Bergman (1979). A função taxa de falha de  $T$  pode ser caracterizada pelo seu correspondente TTT ou seu TTT escalonado, dado por

$$H_T^{-1}(u) = \int_0^{F_T^{-1}(u)} S_T(t) dy \quad \text{e} \quad W_T(u) = \frac{H_T^{-1}(u)}{H_T^{-1}(1)},$$

para  $0 < u < 1$ , respectivamente, onde  $F_T^{-1}(\cdot)$  é a inversa da fda de  $T$  e  $S_T(\cdot)$  é a função de sobrevivência. Agora,  $W_T(\cdot)$  pode ser aproximado empiricamente por

$$W_n \left( \frac{r}{n} \right) = \sum_{i=1}^r T_{(i)} + (n - r)T_{(r)}, \quad k = 1, \dots, n,$$

em que  $r = 1, \dots, n$  e  $T_{(i)}$ ,  $i = 1, \dots, n$  representam as estatísticas de ordem da amostra. O gráfico de  $r/n$  por  $W_n \left( \frac{r}{n} \right)$  definido em um quadrado de área 1 mostra qual o formato da função de risco correspondente aos dados de sobrevivência, veja a Figura 2.5. Se o gráfico se aproxima de uma linha diagonal temos função de risco constante e, se a curvatura é côncava e depois convexa a função de risco tem forma unimodal. Se o gráfico apresentar curvatura convexa e depois côncava a função de risco é em forma de banheira. O gráfico TTT plot é apenas uma condição suficiente e não necessária para indicar a forma da função de risco e será utilizado como um indicador de seu comportamento. Para mais detalhes do método TTT, veja Aarset (1987).

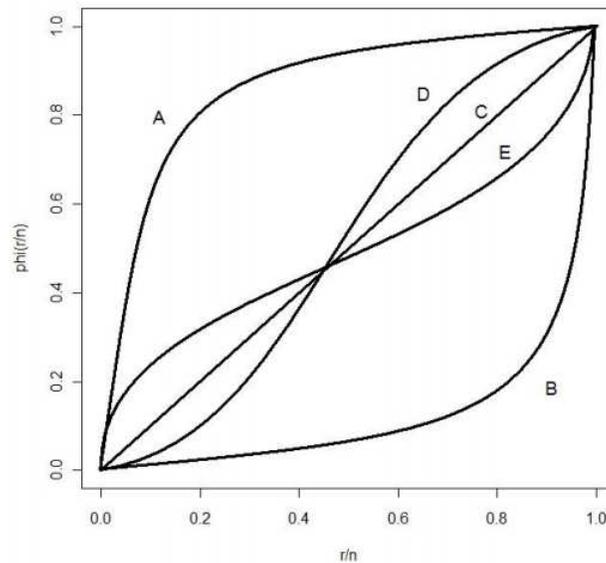


Figura 2.5: Curvas das transformadas TTT empíricas. As curvas A, B, C, D e E, respectivamente, indicam que os dados possuem taxas de risco crescente, decrescente, constante, banheira e unimodal.

## 2.8 Estudo de simulação

Nesta seção, será realizado um estudo de simulação, com o intuito de avaliar numericamente os desempenhos das estatísticas propostas na Seção 5 para testar as hipóteses  $H_{0k}$ , para  $k = 1, 2, 3, 4$  proposto na Seção 2.5 no contexto da distribuição  $\mathcal{BS}$  bivariada. Vamos considerar que um teste tenha bom desempenho caso apresente tamanhos empíricos razoavelmente próximos (ou menores) em nível nominal. Dessa forma, um estudo de simulação, baseado nos níveis de significância dos três testes (Wald, Escore e Razão de Verossimilhança) sob as hipóteses definidas na Seção 2.5, será considerado.

O método de Monte Carlo é utilizado para estimar o nível de significância. Assim, 1000 amostras independentes foram geradas de acordo com o modelo definido em (2.2.2) para tamanhos de amostra  $n = 25, 50$  e  $150$ . Especificamente, note que os quatro cenários para o valor do parâmetro  $\rho$  é,  $\rho = \pm 0,5$  e  $\rho = \pm 0,9$  mais só para  $H_{04}$ ,  $\rho = 0$ . Os outros parâmetros foram baseados no trabalho de Kundu *et al.* (2010) e são:

$$(H_{01}) \quad \alpha_1 = \alpha_2 = 1, \beta_1 = 1, \beta_2 = 2;$$

$$(H_{02}) \quad \alpha_1 = 0.5, \alpha_2 = 1, \beta_1 = \beta_2 = 2;$$

$$(H_{03}) \quad \alpha_1 = \alpha_2 = 1, \beta_1 = \beta_2 = 2;$$

$$(H_{04}) \quad \alpha_1 = 0.5, \alpha_2 = 2, \beta_1 = 1, \beta_2 = 2, \rho = 0 ;$$

O cenário da simulação considerado para o estudo do comportamento de  $W_{0k}$ ,  $S_{0k}$  e  $Q_{0k}$ ,  $k = 1, 2, 3, 4$ , para amostras moderadas e grande  $n \in \{25, 50, 150\}$ . Considerando o nível de significância nominal  $\alpha = 5\%$ , obtivemos os correspondentes níveis de significância empíricos. As estatísticas de Wald (W), Escore (E) e Razão de verossimilhanças (Q) foram calculadas de acordo com a Seção 2.6.

Tabela 2.2: Nível de significância empírica do teste baseado na estatística de Wald, Escore e Razão de Verossimilhança para  $\rho = 0.5$

		$H_{01}: \alpha_1 = \alpha_2$			$H_{02}: \beta_1 = \beta_2$			$H_{03}: \alpha_1 = \alpha_2, \beta_1 = \beta_2$			$H_{04}: \rho = 0$		
$\alpha$	$n$	$W_{01}$	$S_{01}$	$Q_{01}$	$W_{02}$	$S_{02}$	$Q_{02}$	$W_{03}$	$S_{03}$	$Q_{03}$	$W_{04}$	$S_{04}$	$Q_{04}$
5%	25	7.2	5.8	6.6	7.3	5.9	6.2	7.7	5.8	6.6	6.3	5.4	6.0
	50	6.4	5.6	6	5.7	5.4	5.4	5.9	5.3	5.4	5.7	5.3	5.3
	150	5.2	5	5.1	5	5.1	5.1	5.2	4.9	5.1	4.9	5	4.9
10%	25	12.5	10.8	11.7	13.1	10.8	11.9	11.8	10.9	11.6	11.6	10.7	11.2
	50	11	10.5	10.8	12.6	10.5	11	11.5	10.6	10.9	11	10.4	10.7
	150	10.3	10	9.9	10.1	10	10.1	10	10.1	10.2	10.1	10	10

Em todos os casos a característica de interesse  $\mathbf{t}_i = (t_{1i}, t_{1i})$ ,  $i = 1, \dots, n$  foi gerada de acordo a uma distribuição  $BS_2(\boldsymbol{\alpha}_0^\top, \boldsymbol{\beta}_0^\top, \rho_0)$  em que  $\boldsymbol{\alpha}_0$ ,  $\boldsymbol{\beta}_0$  e  $\rho_0$  foram apresentados nas equações.

Tabela 2.3: Nível de significância empírica do teste baseado na estatística de Wald, Escore e Razão de Verossimilhança para  $\rho = -0.5$

		$H_{01}: \alpha_1 = \alpha_2$			$H_{02}: \beta_1 = \beta_2$			$H_{03}: \alpha_1 = \alpha_2, \beta_1 = \beta_2$		
$\alpha$	$n$	$W_{01}$	$S_{01}$	$Q_{01}$	$W_{02}$	$S_{02}$	$Q_{02}$	$W_{03}$	$S_{03}$	$Q_{03}$
5%	25	7.1	5.9	6.9	6.7	5.8	6.3	7.5	5.8	6.7
	50	6.7	5.7	5.8	5.7	5.6	5.5	5.7	5.6	5.5
	150	5.2	4.8	5.2	4.8	4.9	4.8	5.1	5	5
10%	25	12.4	10.9	11.7	12	10.8	11.7	12.5	10.9	11.6
	50	11.5	10.6	10.8	10.9	10.4	10.6	11.6	10.4	10.5
	150	10.3	10	10.1	10.2	10	10.1	10	10	10

Na Tabela 2.2, podemos notar que os testes de Wald, Escore e Razão de Verossimilhança convergem para o níveis nominais de 5% e 10%, sendo o teste de Wald o que apresenta uma convergência um pouco mais lenta para o nível nominal.

Nas Tabelas 2.2 e 2.3, podemos observar que para os coeficientes de correlação  $\rho = 0.5$  e  $\rho = -0.5$  convergem para os níveis nominais de 5% e 10%, ainda para o valor pequeno de  $n = 25$ , sendo que o teste de Score apresenta uma convergência mais rápida para os níveis nominais. Além disso, o teste de Razão de Verossimilhança converge mais rápido para o nível nominal em comparação ao teste de Wald. Também podemos observar que na Tabela 2.2 consideramos  $H_{04}$ , mais considerando  $\rho = 0$  onde o teste de Escore converge mais rápido em comparação aos teste de Wald e Razão de Verossimilhança. Note que para  $n$  suficientemente grande o comportamento dos três testes é otimizado consideravelmente.

Tabela 2.4: Nível de significância empírica do teste baseado na estatística de Wald, Escore e Razão de Verossimilhança para  $\rho = 0.9$ 

$\alpha$	$n$	$H_{01}: \alpha_1 = \alpha_2$			$H_{02}: \beta_1 = \beta_2$			$H_{03}: \alpha_1 = \alpha_2, \beta_1 = \beta_2$		
		$W_{01}$	$S_{01}$	$Q_{01}$	$W_{02}$	$S_{02}$	$Q_{02}$	$W_{03}$	$S_{03}$	$Q_{03}$
5%	25	8	5.7	7.7	8	5.8	6.8	7.6	5.9	7.1
	50	6.2	5.4	6	5.4	5.5	5.7	6.4	5.5	6.1
	150	5.2	5.1	5.2	5.1	5	5.2	5.2	5.1	5.2
10%	25	12	10.7	11.6	12.9	10.7	12.4	11.8	10.8	11.3
	50	11.5	10.5	10.7	11.9	10.4	11.6	11.2	10.5	10.9
	150	10.5	10.1	9.8	10.1	10	10.1	10.2	10.2	10.1

Nas Tabelas 2.4 e 2.5, podemos notar que para os coeficientes de correlação  $\rho = 0.9$  e  $\rho = -0.9$  os níveis nominais de 5% e 10% são superestimados, principalmente para  $n = 25$  também observamos que os testes Wald e Razão de Verossimilhança convergem um pouco mais lento, mais o teste de Wald converge mais lentamente em comparação ao teste de Razão de Verossimilhança, o teste de Escore apresenta uma convergência mais rápida para os níveis nominais em comparação com os outros dois testes, além disso o teste de Razão de Verossimilhança converge mais rápido para o nível nominal em comparação ao teste de Wald. Também nota-se que para  $n$  suficientemente grande o comportamento dos três testes é otimizado consideravelmente.

Assim, no contexto de minimizar o erro do tipo I, podemos concluir que o teste de Escore é melhor considerando  $H_{01}$ ,  $H_{02}$ ,  $H_{03}$ , e  $H_{04}$ .

Tabela 2.5: Nível de significância empírica do teste baseado na estatística de Wald, Escore e Razão de Verossimilhança para  $\rho = -0.9$ 

$\alpha$	$n$	$H_{01}: \alpha_1 = \alpha_2$			$H_{02}: \beta_1 = \beta_2$			$H_{03}: \alpha_1 = \alpha_2, \beta_1 = \beta_2$		
		$W_{01}$	$S_{01}$	$Q_{01}$	$W_{02}$	$S_{02}$	$Q_{02}$	$W_{03}$	$S_{03}$	$Q_{03}$
5%	25	7.1	5.5	6.8	6.5	5.7	6.5	7.3	5.7	6.9
	50	6	5.3	5.9	6	5.2	5.5	6.2	5.4	5.8
	150	5.1	5	5.1	5.1	5	4.8	5	5.1	5.2
10%	25	13.2	10.6	13.1	13.6	10.7	13.5	12.4	10.7	12.2
	50	11.8	10.3	11.1	11.5	10.5	11	11.6	10.4	11.3
	150	10.4	10	10.2	10.3	10	10.2	10.1	10	10.1

## 2.9 Aplicação 1

Nesta seção, apresentamos a análise de dados reais para ilustrar as metodologias apresentadas. O conjunto de dados apresentado em Johnson & Wichern (1998), que representa a densidade mineral dos ossos (DMO) medido em  $g/cm^2$  para 24 indivíduos, que participaram de um estudo experimental, veja o conjunto de dados em I.

Os dados são bivariados e são apresentados na Tabela I.1. Estes dados foram analisados recentemente por Kundu *et al.* (2010), que ressaltou que os dados são mais adequados para a distribuição

Tabela 2.6: Conjunto de dados de ossos: Estatísticas descritivas para  $T_1$  e  $T_2$

	Mínimo	Média	Máximo	Percentil 25	Mediana	Percentil 75
$T_1$	0.493	0.8408	1.1030	0.7908	0.8490	0.9220
$T_2$	0.463	0.8410	1.0270	0.7745	0.8650	0.9170

$\mathcal{BS}$  bivariada.

Nas Tabelas 2.6 e 2.7 apresentamos algumas estatísticas descritivas e as estimativas, EP e IC dos parâmetros do modelo considerado. Na Tabela 2.7 notamos que os EMV dos parâmetros da distribuição  $\mathcal{BS}$  bivariada, junto com seus erros padrões correspondente, que são calculados a partir da matriz de informação esperada. Notamos que menores do que os EP obtidos usando a matriz de informação observada usado por Kundu *et al.* (2010).

Tabela 2.7: EMV para os parâmetros. Os valores dos EP, entre parênteses, e 95% intervalo de confiança (IC)

	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\rho}$
EMV	0.1491	0.1674	0.8313	0.8292	0.9343
EP	(2.579289e-06)	(3.251641e-06)	(2.094761e-05)	(2.613036e-05)	(1.592848e-05)
IC	(0.1459; 0.1526)	(0.1642, 0.1709)	(0.8223, 0.8413)	(0.8202, 0.8392)	(0.9265, 0.9422)

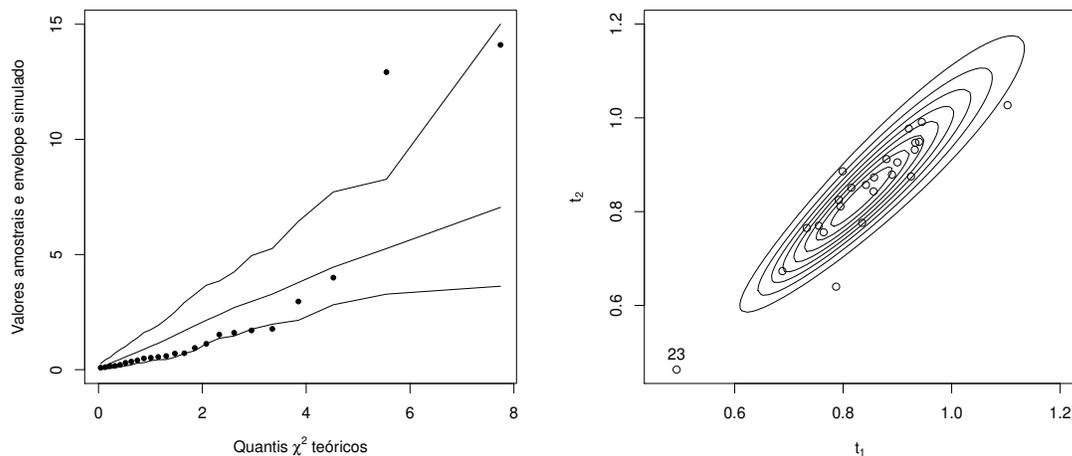


Figura 2.6: (Esquerda) Q-Q plots e envelope simulado e (Direita) Contorno do gráfico da função de densidade estimada para  $t_1$  e  $t_2$ .

A Figura 2.9 mostra os envelopes simulados (linhas que representam o percentual 5, a média, e do percentual 95 de 100 pontos simulados para cada observação). No lado direito da Figura 2.9 é apresentada no gráfico de dispersão os contornos da densidade estimada do modelo  $\mathcal{BS}$  bivariada.

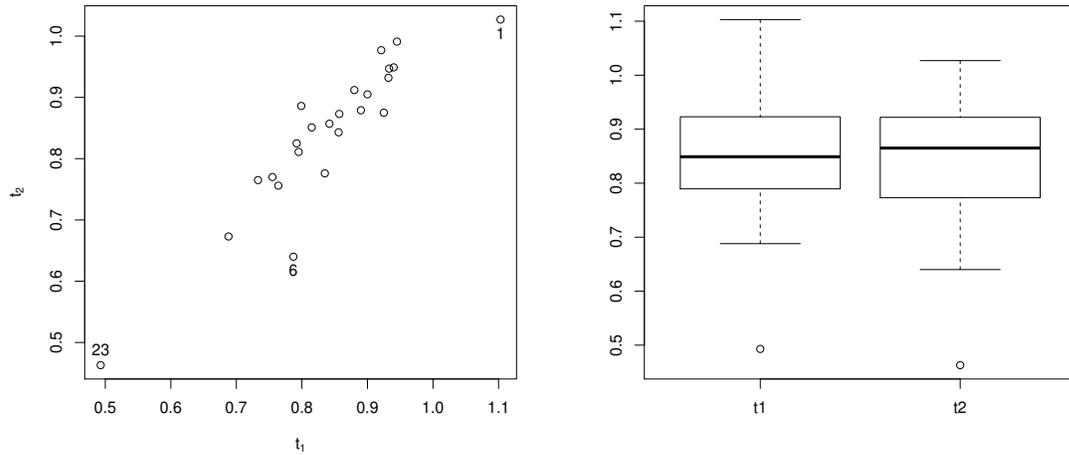


Figura 2.7: (Esquerda) Diagrama de dispersão entre  $T_1$  e  $T_2$  e (Direita) Diagrama de caixas entre  $T_1$  e  $T_2$ .

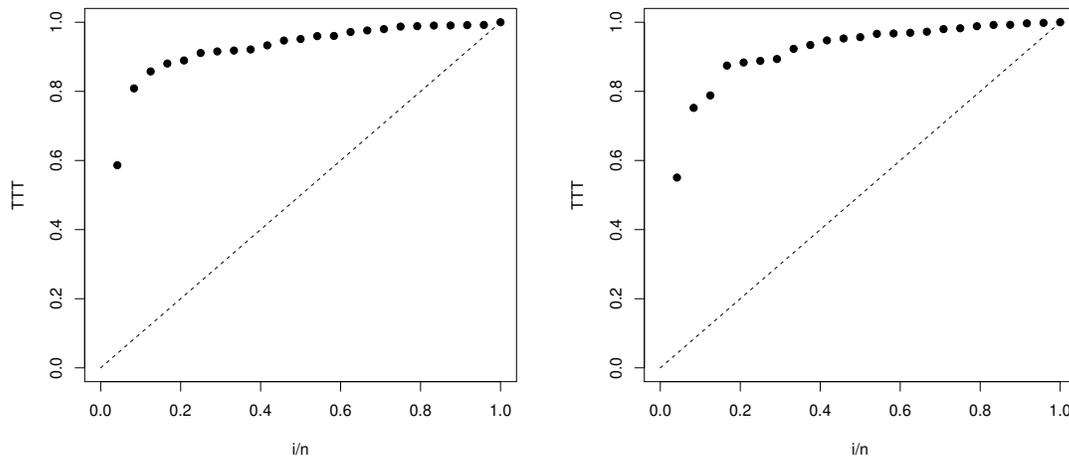


Figura 2.8: (Esquerda) TTT-Plot para o  $T_1$  e (Direita) TTT-Plot para o  $T_2$ .

A partir desta figura, vemos que a distribuição  $\mathcal{BS}$  bivariada fornece um melhor ajuste aos dados considerados. Além disso, a Figura 2.9 apresenta os ajustes marginais para as respostas  $T_1$  e  $T_2$ , a qual sugere um bom ajuste do modelo aos dados, de um ponto de vista marginal. Contudo, a partir da Figura 2.7, podemos ver que #23 aparece como um possível *outlier*.

Na Figura 2.8(a) e (b), percebe-se que, para ambas as variáveis, o risco parece ser crescente,

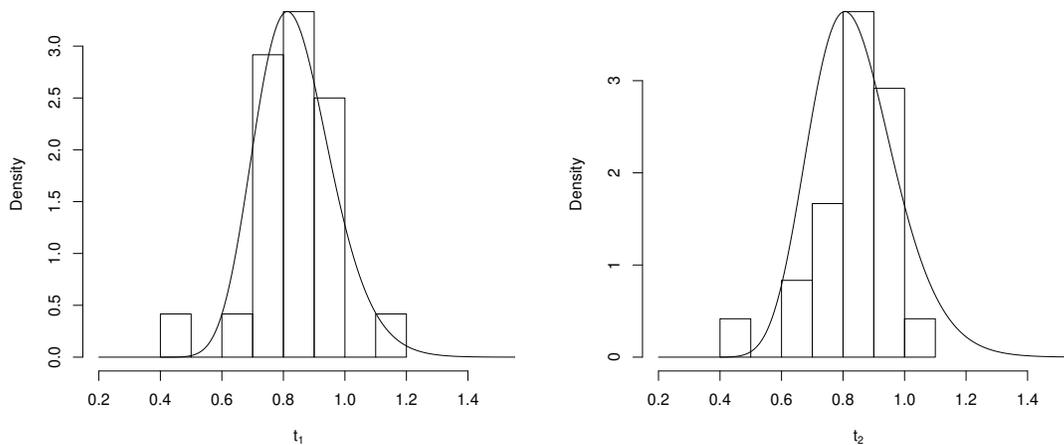


Figura 2.9: Histograma para os dados e as curvas estimados para as funções de densidade marginais de  $T_1$  e  $T_2$ .

devido às suas formas côncavas.

Tabela 2.8: Parâmetros EMV da distribuição BS bivariada para os dados apresentados na Tabela acima

Hipótese	$\alpha_1$	$\alpha_2$	$\beta_1$	$\beta_2$	$\rho$
$H_{01}$	0,1585	0,1585	0,8311	0,8291	0,9281
$H_{02}$	0,1491	0,1674	0,8319	0,8319	0,9342
$H_{03}$	0,1585	0,1585	0,8301	0,8301	0,9280
$H_{04}$	0,1491	0,1674	0,8316	0,8293	<0,000

A Tabela 2.8 apresenta as EMV dos parâmetros do modelo sob  $H_{0k}$ . Notamos que as estimativas de  $\beta_1$  e  $\beta_2$  não mostram muita diferença no âmbito das quatro hipóteses e as EMV sob hipóteses  $H_{01}$ ,  $H_{02}$  e  $H_{03}$  são menores.

Tabela 2.9: Estatísticas do teste Wald, Razão de Verossimilhança e Escore

Hipótese	W	p-valor	Q	p-valor	E	p-valor
$H_{01}$	2.520	0.284	2.420	0.298	2.301	0.316
$H_{02}$	0.044	0.978	0.041	0.979	0.040	0.980
$H_{03}$	2.564	0.633	2.460	0.652	2.336	0.674
$H_{04}$	20.952	0	49.526	0	20.938	0

A Tabela 2.9 apresenta os valores das estatísticas dos testes de Wald, Escore e Razão de Verossimilhança juntos com o p-valor. Analisando estes resultados, percebemos que as hipóteses

1, 2 e 3 não são rejeitadas, enquanto a hipótese 4 é rejeitada considerando um nível de significância de 5% y 10%.

## 2.10 Aplicação 2

Nesta seção, apresentamos a análise de dados reais para ilustrar as estimações do modelo proposto. Nos aplicamos alguns dos resultados do conjunto de dados apresentados em Johnson & Wichern (1998), pag. 43, representam os teores de minerais das quatro ossos maiores de 25 recém nascidos. Estes dados foram analisados recentemente por Kundu *et al.* (2013), que desenvolveram o caso  $\mathcal{BS}$  multivariado generalizado, veja na Tabela I.2 no Anexo o conjunto de dados.

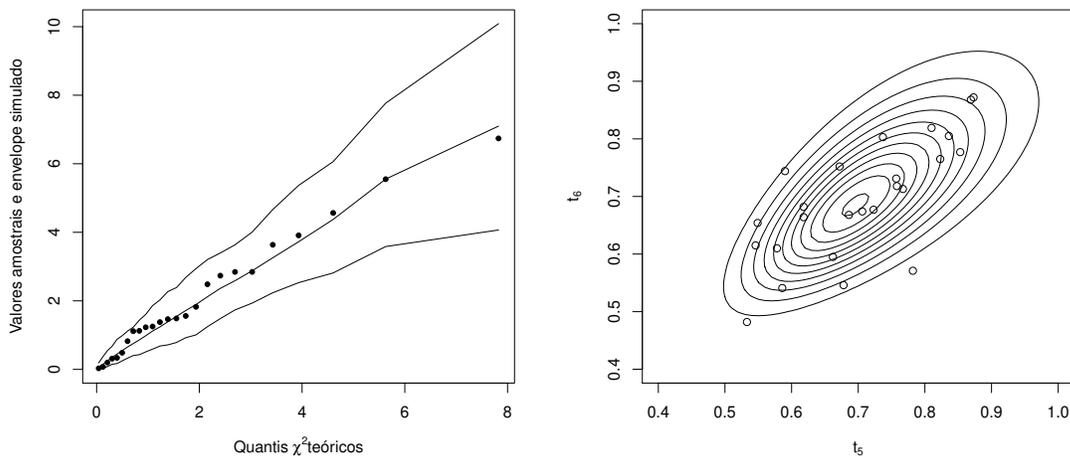


Figura 2.10: (Esquerda) Q-Q plots e envelope simulado e (Direita) Contorno do gráfico da função de densidade estimada para  $T_5$  e  $T_6$ .

A Figura 2.10 mostra os envelopes simulados (linhas que representam o percentual 5, a média, e do percentual 95 de 100 pontos simulados para cada observação) e nota-se que os pontos estão dentro do envelope e isto pode ser observado no gráfico de contorno, sendo assim os dados podem estar bem representados por uma distribuição  $\mathcal{BS}$  bivariada sendo assim vamos trabalhar com as variáveis  $T_5$  e  $T_6$  que tem por nomes, Dominant ulha e ulha respectivamente. Assim as EMV dos parâmetros da distribuição  $\mathcal{BS}$  bivariada, junto com seus erros padrão (EP) correspondentes, que são calculados a partir da matriz de informação esperada são apresentados na Tabela 2.10.

A partir da figura 2.10 (direita), vemos que o modelo BS bivariado fornece um bom ajuste aos dados considerados. Os resultados da figura acima indicam que o modelo ( $\mathcal{BS}$ ) bivariada é um modelo apropriado para os dados bivariados.

A Tabela 2.11 apresenta as EMV dos parâmetros do modelo sob  $H_{0k}$ . Notamos que as estimativas de  $\beta_1$  e  $\beta_2$  não mostram muita diferença no âmbito das quatro hipóteses e as EMV sob hipóteses  $H_{01}$ ,  $H_{02}$  e  $H_{03}$  são menores.

Tabela 2.10: EMV para os parâmetros. Os valores dos EP, entre parênteses, e 95% intervalo de confiança (IC)

	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\rho}$
EMV	0,1526	0,1503	0,6953	0,6852	0,7051
EP	(0,0032)	(0,0032)	(0,0038)	(0,0037)	(0,0180)
IC	(0,1462; 0,1588)	(0,1440; 0,1565)	(0,6878; 0,7027)	(0,6776; 0,6925)	(0,6697; 0,7405)

Tabela 2.11: Parâmetros EMV da distribuição BS bivariada para os dados apresentados na Tabela acima

Hipótese	$\alpha_1$	$\alpha_2$	$\beta_1$	$\beta_2$	$\rho$
H <sub>01</sub>	0,1514	0,1514	0,6961	0,6859	0,7050
H <sub>02</sub>	0,1528	0,1504	0,6907	0,6907	0,7011
H <sub>03</sub>	0,1516	0,1516	0,6910	0,6910	0,7010
H <sub>04</sub>	0,1525	0,1503	0,6962	0,6860	<0,000

A Tabela 2.12 apresenta os valores das estatísticas dos testes de Wald, Escore e Razão de Verossimilhança juntos com o p-valor. Analisando estes resultados, percebemos que as hipóteses 1, 2 e 3 não são rejeitados, enquanto a hipótese 4 é rejeitada a um nível de significância de 5%.

Tabela 2.12: Estatísticas do teste Wald, Razão de Verossimilhança e Escore

Hipótese	W	p-valor	Q	p-valor	E	p-valor
H <sub>01</sub>	0,011	0,915	0,009	0,923	0,011	0,917
H <sub>02</sub>	0,414	0,520	0,398	0,528	0,381	0,537
H <sub>03</sub>	0,425	0,808	0,409	0,815	0,392	0,822
H <sub>04</sub>	12,429	<0,000	17,184	<0,000	12,434	<0,000



# Capítulo 3

## Análise de influência local no modelo BS Bivariado

### 3.1 Introdução

Nesta seção, desenvolveremos um análise de diagnóstico para o modelo  $BS$  bivariado considerando eliminação de casos e análise de influência local, baseada no esquema de perturbação de ponderação de casos.

O análise de diagnóstico é feita assumindo o modelo como correto e então é investigada a robustez das conclusões às pequenas perturbações. A análise de diagnóstico pode ser vista como uma extensão do conceito de robustez para estudar e detectar subconjuntos de observações influentes que comprometem os resultados da análise estatística.

Os elementos de um conjunto de dados que efetivamente controlam aspectos da análise são ditos influentes. Em particular, uma observação é dita influente se ela produzir alterações relevantes no resultado da análise, quando a mesma for excluída ou submetida a uma pequena perturbação. Sendo assim, a presença de tais observações pode-se tornar inadequadas às conclusões com base num ajuste que não incorpore o caráter atípico destas observações por isso que uma identificação das observações responsáveis por essas discrepâncias pode ajudar na escolha de um modelo mais adequado aos dados. Uma abordagem para identificação de observações influentes é avaliar o efeito da exclusão de uma observação na estimativa dos parâmetros do modelo. Um trabalho pioneiro nessa direção foi proposto por Cook (1977), para o modelo de regressão, e este poder es adaptado para outros modelos, como pode ser vista em Cook (1986), que propõe um método geral para avaliar o efeito, sob o estimador de máxima verossimilhança, com pequenas perturbações no modelo estatístico e/ou nos dados, baseado no afastamento pela verossimilhança. O método de influência local proposto por Cook (1986) tem sido aplicado em diversos modelos estatísticos. A ideia principal de influência local é verificar, através de alguma medida apropriada de influência, o efeito das pequenas perturbações no modelo ou nos dados. Se essas perturbações causarem efeitos desproporcionais em determinados resultados do modelo, podem ser indícios de que o modelo proposto não é apropriado ou que existem afastamentos importantes das suposições feitas para o

mesmo.

Embora o método de influência local tenha sido aplicado com sucesso em muitos modelos estatísticos, alguns autores têm o criticado devido a falta de um critério objetivo para avaliar a grandeza da curvatura normal e tamanho relativo das componentes da direção correspondente à maior curvatura normal. Antes disso Poon & Poon (1999) propõe um método de influência local baseado na curvatura normal conformalizada, que é uma medida padronizada da curvatura normal utilizada por Cook (1986). Loynes (2001) propõe uma nova medida de influência local que chama de "propagação de curvatura" (tradução livre de "spread of curvature") baseado num estudo probabilístico da curvatura normal.

## 3.2 Método de eliminação de casos

A identificação de observações que influenciam demasiadamente nas estimativas dos parâmetros é uma etapa fundamental na análise estatística, já que a presença deste tipo de observações, pode gerar inferências inadequadas. Uma abordagem importante na identificação de observações atípicas é baseada na metodologia conhecida como Método de Eliminação de Casos, discutido em Cook (1977).

Para estudar o impacto da  $i$ -ésima observação na EMV de  $\boldsymbol{\theta}$ , é usual comparar as estimativas dos parâmetros obtidos com todas as observações, denotada por  $\hat{\boldsymbol{\theta}}$ , e as estimativas dos parâmetros obtidos quando  $i$ -ésima observação foi excluída, denotada por  $\hat{\boldsymbol{\theta}}_{[i]}$ . Esta metodologia corresponde ao método de eliminação de casos. Para o modelo  $\mathcal{BS}$  bivariado definido em (2.2.2) denotamos a log-verossimilhança para  $\boldsymbol{\theta}$  sem a  $i$ -ésima observação como

$$\ell_{[i]}(\boldsymbol{\theta}) = \sum_{j \neq i} \ell_j(\boldsymbol{\theta}).$$

Calcular  $\hat{\boldsymbol{\theta}}_{[i]} = (\boldsymbol{\alpha}_{[i]}^\top, \boldsymbol{\beta}_{[i]}^\top, \rho_{[i]})^\top$ ,  $\forall i$  e comparar  $\hat{\boldsymbol{\theta}}_{[i]}$  com  $\hat{\boldsymbol{\theta}}$ , pode envolver um extenso trabalho computacional, especialmente quando o tamanho da amostra  $n$  é muito grande. Felizmente, o seguinte resultado obtido por (Cook & Weisberg, 1982) proporciona uma fórmula aproximada sobre o método de eliminação de casos para estimar os parâmetros, em cada um dos  $n$  casos, isto é,

$$\hat{\boldsymbol{\theta}}_{[i]}^{-1} = \hat{\boldsymbol{\theta}} + \left\{ \mathbf{J}(\hat{\boldsymbol{\theta}}) \right\}^{-1} \dot{\ell}_{[i]}(\hat{\boldsymbol{\theta}}), \quad (3.2.1)$$

em que  $\mathbf{J}(\hat{\boldsymbol{\theta}}) = -\ddot{\mathbf{L}}$ , definido na Seção 2.3, conhecido como matriz Hessiana e  $\dot{\ell}_{[i]}(\hat{\boldsymbol{\theta}}) = \partial \ell_{(i)}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$ , com

$$\frac{\partial \ell_{[i]}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left( \sum_{j \neq i} \frac{\partial \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha}^\top}, \sum_{j \neq i} \frac{\partial \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}^\top}, \sum_{j \neq i} \frac{\partial \ell_j(\boldsymbol{\theta})}{\partial \rho} \right)^\top,$$

em que os elementos da derivada são dados por

$$\begin{aligned}\dot{\ell}_{[i]\alpha}(\hat{\boldsymbol{\theta}}) &= \frac{\partial \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha}} = D^{-1}(\boldsymbol{\alpha}) \left\{ -(n-1)\mathbf{1}_2 + \sum_{j \neq i}^n D(a(\mathbf{t}_j; \boldsymbol{\alpha}, \boldsymbol{\beta})) \boldsymbol{\Sigma}^{-1} a(\mathbf{t}_j; \boldsymbol{\alpha}, \boldsymbol{\beta}) \right\}, \\ \dot{\ell}_{[i]\beta}(\hat{\boldsymbol{\theta}}) &= \frac{\partial \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = -\frac{1}{2} D^{-1}(\boldsymbol{\beta}) \sum_{j \neq i}^n [C(\mathbf{t}_i; \boldsymbol{\beta}) - B(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}) \boldsymbol{\Sigma}^{-1} a(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta})], \\ \dot{\ell}_{[i]\rho}(\hat{\boldsymbol{\theta}}) &= \frac{\partial \ell_j(\boldsymbol{\theta})}{\partial \rho} = \frac{(n-1)\rho}{1-\rho^2} - \frac{1}{2} \sum_{i=1}^n a^\top(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}) \boldsymbol{\Sigma}_\rho^{-1} a(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}),\end{aligned}$$

em que

$$C(\mathbf{t}_i; \boldsymbol{\beta}) = D^{-1}(\mathbf{t}_i + \boldsymbol{\beta})(\mathbf{t}_i - \boldsymbol{\beta}),$$

e

$$B(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}) = D^{-1/2}(\boldsymbol{\beta}) D^{-1}(\boldsymbol{\alpha}) D^{-1/2}(\mathbf{t}_i) D(\mathbf{t}_i + \boldsymbol{\beta}),$$

com

$$\boldsymbol{\Sigma}_\rho^{-1} = \partial \boldsymbol{\Sigma}^{-1} / \partial \rho = -\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_\rho \boldsymbol{\Sigma}^{-1}, \quad \text{para } i = 1, \dots, n.$$

Deste resultado, podemos ver as diferenças entre as estimativas dos parâmetros, com todas as observações e sem a  $i$ -ésima observação. Consequentemente podemos obter medidas para avaliar o impacto das observações influentes no modelo  $\mathcal{BS}$  bivariado.

### Distância generalizada de Cook

A distância generalizada de Cook é definida como a distância padronizada entre  $\hat{\boldsymbol{\theta}}_{[i]}$  e  $\hat{\boldsymbol{\theta}}$ , isto é,

$$GD_i = (\hat{\boldsymbol{\theta}}_{[i]} - \hat{\boldsymbol{\theta}})^\top \mathbf{M} (\hat{\boldsymbol{\theta}}_{[i]} - \hat{\boldsymbol{\theta}}), \quad (3.2.2)$$

que avalia a combinação ponderada dos elementos para a diferença  $\hat{\boldsymbol{\theta}}_{[i]} - \hat{\boldsymbol{\theta}}$ . Cook & Weisberg (1982) consideram várias opções para  $\mathbf{M}$ . Uma escolha muito usada é a matriz de informação observada, isto é  $\mathbf{M} = -\ddot{\mathbf{L}}$ , veja em (??). Agora, substituindo (3.2.1) em (3.2.2), obtemos a seguinte aproximação:

$$GD_i^l = \dot{\ell}_{[i]}(\hat{\boldsymbol{\theta}})^\top \{-\ddot{\mathbf{L}}\}^{-1} \dot{\ell}_{[i]}(\hat{\boldsymbol{\theta}}), \quad i = 1, \dots, n. \quad (3.2.3)$$

### Afastamento pela log-verossimilhança

O afastamento pela verossimilhança ( $LD_i$ ) considerado para avaliar a exclusão da  $i$ -ésima observação como proposto por Cook & Weisberg (1982) é definido como:

$$LD_i = 2\{\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{[i]})\}. \quad (3.2.4)$$

Substituindo (3.2.1) em (3.2.4), obtemos a seguinte aproximação para  $LD_i$ .

$$LD_i = 2\{\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}} + \{-\ddot{\mathbf{L}}\}^{-1} \dot{\ell}_i(\widehat{\boldsymbol{\theta}}))\}, \quad i = 1, \dots, n,$$

em que  $\widehat{\boldsymbol{\theta}}_{[i]}$  é o EMV de  $\boldsymbol{\theta}$  obtido após eliminar a  $i$ -ésima observação,  $i = 1, \dots, n$ .

### 3.3 Influência local

Nesta seção vamos apresentar, resumidamente, o método de influência local proposto por Cook (1986).

Para um conjunto de dados observados, seja  $\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta})$  a log-verossimilhança do modelo postulado, onde  $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \rho)^\top$  é um vetor  $(p = 5) \times 1$  de parâmetros desconhecidos e  $\ell_i(\boldsymbol{\theta})$  é a contribuição da  $i$ -ésima observação à log-verossimilhança. A perturbação no modelo é introduzida através de um vetor  $\boldsymbol{\omega}$  de dimensão  $n \times 1$ , onde  $\boldsymbol{\omega} \in \boldsymbol{\Omega} \subseteq \mathbb{R}^n$ ,  $\boldsymbol{\Omega}$  um aberto. Seja  $\ell(\boldsymbol{\theta}, \boldsymbol{\omega})$  a função de log-verossimilhança do modelo perturbado, e suponha que existe um  $\boldsymbol{\omega}_0 \in \boldsymbol{\Omega}$  tal que  $\ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}, \boldsymbol{\omega}_0)$ , para todo  $\boldsymbol{\theta}$  e que  $\ell(\boldsymbol{\theta}, \boldsymbol{\omega})$  é duas vezes diferenciável em relação à  $(\boldsymbol{\theta}^\top, \boldsymbol{\omega}^\top)^\top$ .

Sejam  $\widehat{\boldsymbol{\theta}}$  e  $\widehat{\boldsymbol{\theta}}_\omega$  os EMV de  $\boldsymbol{\theta}$  sob o modelo postulado e perturbado, respectivamente. O objetivo é comparar  $\widehat{\boldsymbol{\theta}}$  e  $\widehat{\boldsymbol{\theta}}_\omega$  em torno de  $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ . A proposta de Cook (1986) é estudar o comportamento do afastamento pela verossimilhança, ou de alguma outra medida de influência, em torno do vetor de não-perturbação  $\boldsymbol{\omega}_0$ . Sugere-se que a comparação entre  $\boldsymbol{\theta}$  e  $\widehat{\boldsymbol{\theta}}_\omega$  seja feita através do afastamento pela verossimilhança  $LD(\boldsymbol{\omega})$  definida por

$$LD(\boldsymbol{\omega}) = 2 \left[ \ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_\omega) \right], \quad \boldsymbol{\omega} \in \boldsymbol{\Omega}. \quad (3.3.1)$$

O afastamento pela verossimilhança descrito acima é uma medida de influência que permite avaliar a perturbação  $\boldsymbol{\omega}$  na estimação de máxima verossimilhança do vetor paramétrico completo  $\boldsymbol{\theta}$ .

O sentido da distância entre  $\widehat{\boldsymbol{\theta}}$  e  $\widehat{\boldsymbol{\theta}}_\omega$ , baseado na função  $LD(\boldsymbol{\omega})$ , pode depender da concavidade da função de log-verossimilhança  $\ell(\boldsymbol{\theta})$ . Se  $\ell(\boldsymbol{\theta})$  é suficientemente achatada, pode-se dizer que  $\widehat{\boldsymbol{\theta}}$  e  $\widehat{\boldsymbol{\theta}}_\omega$  estão próximos entre si, enquanto que se  $\ell(\boldsymbol{\theta})$  for suficientemente concentrada em torno de  $\widehat{\boldsymbol{\theta}}$ , as estimativas  $\widehat{\boldsymbol{\theta}}$  e  $\widehat{\boldsymbol{\theta}}_\omega$  podem estar distantes entre si, veja Cook (1986).

A ideia de Cook (1986) é investigar o comportamento da função  $LD(\boldsymbol{\omega})$  numa vizinhança de  $\boldsymbol{\omega}_0$ , em que  $LD(\boldsymbol{\omega}_0) = 0$ , o qual é o ponto em que as duas verossimilhanças são iguais. Para isso considerou o gráfico de  $LD(\boldsymbol{\omega})$  versus  $\boldsymbol{\omega}$  que pode ser visto como a superfície geométrica formada pelos valores do vetor  $(q + 1)$  dimensional.

$$\boldsymbol{\alpha}(\boldsymbol{\omega}) = \begin{pmatrix} \boldsymbol{\omega} \\ LD(\boldsymbol{\omega}) \end{pmatrix},$$

em que  $\boldsymbol{\omega}$  varia em  $\boldsymbol{\Omega}$ . A função  $\alpha(\boldsymbol{\omega})$  é chamada de gráfico de influência local, o qual corresponde a uma superfície em  $\mathbb{R}^{q+1}$  e pode ser usada para avaliar a influência ao variar  $\boldsymbol{\omega}$  através de  $\boldsymbol{\Omega}$ . Uma ilustração gráfica é dada na Figura 3.1.

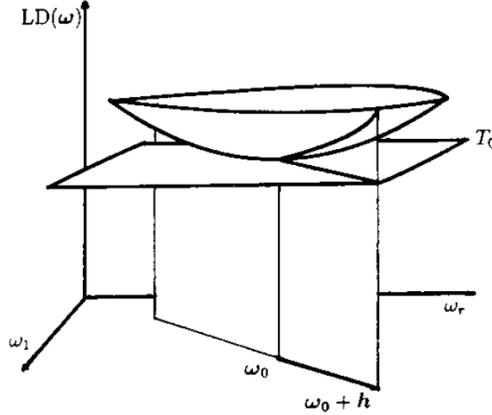


Figura 3.1: Representação gráfica do enfoque de influência local.

A ideia principal do método de Cook (1986) baseado na superfície  $\alpha(\boldsymbol{\omega})$  em torno de  $\boldsymbol{\omega}_0$  consiste em considerar o plano tangente ( $\mathbf{T}_0$ ) à superfície  $\alpha(\boldsymbol{\omega})$  em  $\boldsymbol{\omega}_0$ . Como  $LD(\boldsymbol{\omega})$  atinge o mínimo em  $\boldsymbol{\omega}_0$ , temos que  $\mathbf{T}_0$  é paralelo a  $\boldsymbol{\Omega} \subseteq \mathbb{R}^q$ . Cada vetor  $\mathbf{h}$  em  $\boldsymbol{\Omega}$ , de norma um, determina um plano que contém  $\mathbf{h}$  e que é ortogonal a  $\mathbf{T}_0$ . A intersecção, chamada de seção normal, deste plano com a superfície é chamada de linha levantada,  $C_{\mathbf{h}}$ , agora definida como a curvatura da curva plana  $(a, LD(\boldsymbol{\omega}_0 + a\mathbf{h}))$  em  $a = 0$  e pode ser visualizada como a inversa do raio do círculo que melhor se ajusta em  $\boldsymbol{\omega}_0$ . Valores grandes de  $C_{\mathbf{h}}$  indicam sensibilidade em relação à perturbação considerada na direção  $\mathbf{h}$ .

Cook (1986) mostra que a curvatura normal na direção  $\mathbf{h}$  pode ser expressa na seguinte forma:

$$C_{\mathbf{h}} = 2 \left| \mathbf{h}^{\top} \ddot{\mathbf{F}} \mathbf{h} \right|, \quad (3.3.2)$$

em que  $\|\mathbf{h}\| = 1$ ,  $\ddot{\mathbf{F}} = \boldsymbol{\Delta}^{\top} (-\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta}$ , com  $\boldsymbol{\Delta}$  sendo uma matriz  $p \times n$  com elementos  $\Delta_{ij} = \frac{\partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \theta_i \partial \omega_j}$ ,  $i = 1, \dots, p$ ,  $j = 1, \dots, n$  e  $-\ddot{\mathbf{L}}$  é a matriz de informação observada para o modelo postulado avaliada no ponto  $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$  e  $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ . A expressão em (3.3.2) depende da EMV de  $\boldsymbol{\theta}$  e do esquema de perturbação.

Quando o interesse é um subconjunto  $\boldsymbol{\theta}_1$  de  $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^{\top}, \boldsymbol{\theta}_2^{\top})^{\top}$  o equivalente a (3.3.1) é dado por

$$LD_s(\boldsymbol{\omega}) = 2 \left[ \ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{1w}, g(\hat{\boldsymbol{\theta}}_{1w})) \right], \quad (3.3.3)$$

em que é a função que maximiza  $\ell(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$  para cada  $\boldsymbol{\theta}_1$  fixo de  $\hat{\boldsymbol{\theta}}_{1w}$  e é determinada pela partição

de  $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}} = \left( \widehat{\boldsymbol{\theta}}_{1\boldsymbol{\omega}}^{\top}, \boldsymbol{\theta}_{2\boldsymbol{\omega}}^{\top} \right)^{\top}$ . Neste caso, a curvatura normal na direção  $\mathbf{h}$  é dado por

$$C_{\mathbf{h}}(\boldsymbol{\theta}_1) = 2 \left[ \mathbf{h}^{\top} \boldsymbol{\Delta}^{\top} \left( \ddot{\mathbf{L}}^{-1} - \mathbf{B}_{22} \right) \boldsymbol{\Delta} \mathbf{h} \right], \quad (3.3.4)$$

em que  $\mathbf{B}_{22} = \begin{pmatrix} 0 & 0 \\ 0 & \ddot{\mathbf{L}}_{22}^{-1} \end{pmatrix}$  e  $\ddot{\mathbf{L}}_{22}$  é determinada pela partição de  $\ddot{\mathbf{L}} = \begin{pmatrix} \ddot{\mathbf{L}}_{11} & \ddot{\mathbf{L}}_{12} \\ \ddot{\mathbf{L}}_{21} & \ddot{\mathbf{L}}_{22} \end{pmatrix}$ .

Uma outra escolha evidente corresponde à perturbação na direção do  $i$ -ésimo caso que é obtida tomando  $\mathbf{h}$  igual ao vetor  $\mathbf{h}_i$  de dimensão  $n \times 1$  de zeros com o valor de 1 na  $i$ -ésima posição, dando como resultado a seguinte medida de influência local

$$C_i = C_{h_i} = 2 \left| \boldsymbol{\Delta}_i^{\top} (-\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta}_i \right|, \quad i = 1, \dots, p, \quad (3.3.5)$$

em que  $\boldsymbol{\Delta}_i$  é a  $i$ -ésima coluna da matriz  $\boldsymbol{\Delta}$ . Verveke & Molenberghs (1997), propõe considerar a  $i$ -observação como influente se  $C_i$  é maior que o valor  $2 \sum_{i=1}^q C_i/n$ . Nota-se que  $C_i$  é de interesse particular uma vez que representa a influência local de cada observação separadamente na estimação de  $\boldsymbol{\theta}$ .

Finalmente, existem muitas formas possíveis de escolher o vetor  $\mathbf{h}$ . Por exemplo, a influência local na estimação de  $\boldsymbol{\theta}_1$  na direção do  $i$ -ésimo caso cuja curvatura normal será denotada por  $C_i(\boldsymbol{\theta}_1)$ .

Por Poon & Poon (1999), a curvatura normal conformalizada total em  $\boldsymbol{\omega}_0$  da superfície  $\boldsymbol{\alpha}(\boldsymbol{\omega})$  na direção  $\mathbf{h}_i$  pode ser expressa como

$$B_{\mathbf{h}_i} = - \frac{\boldsymbol{\Delta}_i^{\top} (-\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta}_i}{\sqrt{\text{tr} \left( \boldsymbol{\Delta}^{\top} (-\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta} \right)^2}}, \quad (3.3.6)$$

em que  $\mathbf{h}_i$  e  $\boldsymbol{\Delta}_i$  são como em (3.3.5) com  $i = 1, \dots, q$ . Se  $\lambda_1 \geq \dots \geq \lambda_k > 0$  denota os  $k$  autovalores diferentes de zero de

$$- \frac{\boldsymbol{\Delta}^{\top} (-\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta}}{\sqrt{\text{tr} \left( \boldsymbol{\Delta}^{\top} (-\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta} \right)^2}}, \quad (3.3.7)$$

e se  $\mathbf{a}_1, \dots, \mathbf{a}_k$  são os correspondentes autovetores, pode-se verificar que

$$B_{\mathbf{h}_i} = \sum_{j=1}^k \lambda_j a_{ji}^2, \quad (3.3.8)$$

em que  $a_{ji}$  é o  $i$ -ésimo componente do vetor  $a_j$ ,  $i = 1, \dots, n$ . Isto mostra que casos individuais podem ter uma grande curvatura local total  $B_{\mathbf{h}_i}$ , sem ter uma grande componente  $|a_{1i}|$  na direção

$\mathbf{h}_{max}$ . Segundo a ideia de Poon & Poon (1999) uma regra geral para avaliar a magnitude de um caso específico nos dados é baseado nas quantidades  $B_{\mathbf{h}_i}$ . Especificamente, a contribuição conjunta da  $j$ -ésima base de vetores de perturbações para todo autovetor  $r$ -influyente

$$m(r)_j = \sum_{i=1}^s \lambda_i a_{ji}^2, \quad (3.3.9)$$

em que  $s$  é o número de autovalores maiores a  $r/\sqrt{n}$ ,  $j = 1, \dots, n$  e uma adequada marca de nível para julgar a grandeza da influência local é  $2\bar{m}(r)$  pra  $r > 0$  e  $2b$  para  $r = 0$ , onde  $\bar{m}(r) = \sum_{i=1}^s \lambda_i/n$

e  $b = tr(\mathbf{A}) / (q\sqrt{tr(\mathbf{A}^2)})$ , com  $\mathbf{A} = -\mathbf{\Delta}^\top (-\mathbf{\ddot{L}})^{-1} \mathbf{\Delta}$ . Assim, a avaliação de casos influentes é baseada em  $\{m(r)_l, l = 1, \dots, n\}$ . Zhu & Lee (2001) propõe como ponto de corte o valor  $\bar{m}(r) + c^* DP(m(r))$ , onde  $DP(m(r))$  é o desvio padrão de  $\{m(r)_l, l = 1, \dots, n\}$  e  $c^*$  é uma constante arbitrária. Neste trabalho, consideramos  $c^* = 3.5$ .

### 3.4 Aplicação no Modelo Birnbaum-Saunders Bivariado

Nesta seção, vamos desenvolver o método de influência local para a distribuição  $\mathcal{BS}$  bivariada e assim vamos avaliar a matriz  $\mathbf{\Delta}$  para o esquema de perturbação de ponderação de casos. Para o esquema de perturbação a matriz  $\mathbf{\Delta}$  em (3.3.4) tem a forma  $\mathbf{\Delta} = (\mathbf{\Delta}_\alpha^\top, \mathbf{\Delta}_\beta^\top, \mathbf{\Delta}_\rho)^\top$ , em que cada um dos seus elementos é dado por

$$\begin{aligned} \mathbf{\Delta}_\alpha &= \frac{\partial^2 \ell(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\omega}^\top} \Big|_{\boldsymbol{\omega}_0 \in \mathbb{R}^{2 \times n}}, \\ \mathbf{\Delta}_\beta &= \frac{\partial^2 \ell(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\omega}^\top} \Big|_{\boldsymbol{\omega}_0 \in \mathbb{R}^{2 \times n}}, \\ \mathbf{\Delta}_\rho &= \frac{\partial^2 \ell(\boldsymbol{\theta}, \boldsymbol{\omega} | \hat{\boldsymbol{\theta}})}{\partial \rho \partial \boldsymbol{\omega}^\top} \Big|_{\boldsymbol{\omega}_0 \in \mathbb{R}^{1 \times n}}, \end{aligned}$$

e o vetor de perturbação é dado por  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top$

#### 3.4.1 Perturbação de Ponderação de Casos

Considere a distribuição  $\mathcal{BS}$  bivariada definida em (2.2.2) e o vetor de pesos  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top$  de dimensão  $n \times 1$ . A verossimilhança do modelo perturbado, considerando ponderação de casos, é dada por

$$\ell(\boldsymbol{\theta}, \boldsymbol{\omega}) = \sum_{j=1}^n \omega_j \ell_j(\boldsymbol{\theta}), \quad (3.4.1)$$

em que  $\ell_j(\boldsymbol{\theta})$ ,  $j = 1, \dots, n$ , foi definido em (2.3.2) e  $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \rho)^\top$ . Note que, neste caso o modelo não perturbado é obtido ao considerar  $\boldsymbol{\omega}_0 = \mathbf{1}_n$ .

Derivando (3.4.1) com respeito a  $\boldsymbol{\omega}$ , temos que

$$\frac{\partial \ell(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} = (\ell_1(\boldsymbol{\theta}), \dots, \ell_n(\boldsymbol{\theta})),$$

logo, derivando com respeito  $\boldsymbol{\theta}$  a expressão anterior, obtemos o seguinte resultado

$$\begin{aligned} \Delta &= \frac{\partial^2 \ell(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^\top} \\ &= \frac{\partial}{\partial \boldsymbol{\theta}} \left( \frac{\ell(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} \right) = \frac{\partial}{\partial \boldsymbol{\theta}} (\ell_1(\boldsymbol{\theta}), \dots, \ell_n(\boldsymbol{\theta})) \\ &= \left( \frac{\partial \ell_1(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \dots, \frac{\partial \ell_n(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \\ &= (\Delta_1, \dots, \Delta_n), \end{aligned}$$

em que  $\Delta_j = \frac{\partial \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ ,  $j = 1, \dots, n$ , com  $\frac{\partial \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$  obtida na Seção 2.3, que é dada por

$$\begin{aligned} \frac{\partial \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha}} &= D^{-1}(\boldsymbol{\alpha}) \left\{ -\mathbf{1}_2 + D(a(\mathbf{t}_j; \boldsymbol{\alpha}, \boldsymbol{\beta})) \boldsymbol{\Sigma}^{-1} a(\mathbf{t}_j; \boldsymbol{\alpha}, \boldsymbol{\beta}) \right\}, \\ \frac{\partial \ell_j(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} &= -\frac{1}{2} D^{-1}(\boldsymbol{\beta}) \left\{ C(\mathbf{t}_j; \boldsymbol{\beta}) - B(\mathbf{t}_j; \boldsymbol{\alpha}, \boldsymbol{\beta}) \boldsymbol{\Sigma}^{-1} a(\mathbf{t}_j; \boldsymbol{\alpha}, \boldsymbol{\beta}) \right\}, \\ \frac{\partial \ell_j(\boldsymbol{\theta})}{\partial \rho} &= \frac{\rho}{1 - \rho^2} - \frac{1}{2} a^\top(\mathbf{t}_j; \boldsymbol{\alpha}, \boldsymbol{\beta}) \boldsymbol{\Sigma}_\rho^{-1} a(\mathbf{t}_j; \boldsymbol{\alpha}, \boldsymbol{\beta}), \end{aligned}$$

avaliadas em  $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\alpha}}^\top, \hat{\boldsymbol{\beta}}^\top, \hat{\rho})^\top$ , com  $\boldsymbol{\Sigma}_\rho^{-1} = \partial \boldsymbol{\Sigma}^{-1} / \partial \rho = -\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}_\rho \boldsymbol{\Sigma}^{-1}$ . Para este esquema de perturbação, temos que

$$\begin{aligned} \Delta_\alpha &= D^{-1}(\boldsymbol{\alpha}) \left\{ -\mathbf{1}_2 \otimes \mathbf{1}_n^\top + \Delta_{\alpha D} \right\}, \\ \Delta_\beta &= -\frac{1}{2} D^{-1}(\boldsymbol{\beta}) \left\{ \Delta_{\beta C} - \Delta_{\beta B} \right\}, \\ \Delta_\rho &= \frac{\rho}{1 - \rho^2} \mathbf{1}_n^\top - \frac{1}{2} \Delta_{\rho A}, \end{aligned}$$

em que

$$\Delta_{\alpha D} = [D(a(\mathbf{t}_1; \boldsymbol{\alpha}, \boldsymbol{\beta})) \boldsymbol{\Sigma}^{-1} a(\mathbf{t}_1; \boldsymbol{\alpha}, \boldsymbol{\beta}), \dots, D(a(\mathbf{t}_n; \boldsymbol{\alpha}, \boldsymbol{\beta})) \boldsymbol{\Sigma}^{-1} a(\mathbf{t}_n; \boldsymbol{\alpha}, \boldsymbol{\beta})],$$

com

$$\Delta_{\beta C} = (C(\mathbf{t}_1; \boldsymbol{\alpha}, \boldsymbol{\beta}), \dots, C(\mathbf{t}_n; \boldsymbol{\alpha}, \boldsymbol{\beta})),$$

$$\Delta_{\beta B} = \left( B(\mathbf{t}_1; \boldsymbol{\alpha}, \boldsymbol{\beta}) \boldsymbol{\Sigma}^{-1} a(\mathbf{t}_1; \boldsymbol{\alpha}, \boldsymbol{\beta}), \dots, B(\mathbf{t}_n; \boldsymbol{\alpha}, \boldsymbol{\beta}) \boldsymbol{\Sigma}^{-1} a(\mathbf{t}_n; \boldsymbol{\alpha}, \boldsymbol{\beta}) \right),$$

$$\Delta_{\rho A} = \left( a^\top(\mathbf{t}_1; \boldsymbol{\alpha}, \boldsymbol{\beta}) \boldsymbol{\Sigma}_\rho^{-1} a(\mathbf{t}_1; \boldsymbol{\alpha}, \boldsymbol{\beta}), \dots, a^\top(\mathbf{t}_n; \boldsymbol{\alpha}, \boldsymbol{\beta}) \boldsymbol{\Sigma}_\rho^{-1} a(\mathbf{t}_n; \boldsymbol{\alpha}, \boldsymbol{\beta}) \right).$$

e

$$C(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}), B(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}) \text{ e } \boldsymbol{\Sigma}_\rho^{-1},$$

estão definidos na Seção 2.3. Observe que, para o esquema de perturbação de casos, a matriz  $\Delta$  não depende do vetor  $\boldsymbol{\omega}$ .

## 3.5 Ilustração numérica

Nesta seção, um exemplo será usado para ilustrar as técnicas introduzidas nas seções anteriores. Será desenvolvida a análise de diagnóstico com ênfase especial nos métodos de exclusão de casos proposto por Cook (1977) e o método de influência local proposto por Cook (1986) e Poon & Poon (1999), considerando a perturbação desenvolvida na seção prévia.

### 3.5.1 Exemplo 1

A aplicação que apresentamos consiste em analisar o conjunto de dados apresentados na Seção 2.9, tomados de Johnson & Wichern (1998), relativos à densidade mineral dos ossos (DMO) medido em  $g/cm^2$  para 24 indivíduos, que participaram de um estudo experimental.

#### Método de eliminação de casos

A Figura 3.2 apresenta uma análise de casos baseada na distância de Cook ( $GD_i^1$ ) e na Figura 3.3 apresenta análise de casos baseada no afastamento pela verossimilhança ( $LD_i^1$ ), definidas na Subseção 3.2. Baseado na distância de Cook, a Figura 3.2 indica que as observações #1, #6 e #23 são identificadas como influentes nas estimações do parâmetro  $\boldsymbol{\theta}$ . Con tudo, para a estimação do parâmetro  $\boldsymbol{\beta}$  as observações #1, #6 e #8 são identificadas como as mais influentes. Finalmente, a observação mais influente na estimação do parâmetro  $\rho$  é #6.

Um padrão similar pode ser observado quando usamos o afastamento pela verossimilhança, como pode ser visto na Figura 3.3, onde apenas as observações #1 e #6 são identificadas como as mais influentes na estimação dos parâmetros do modelo  $\mathcal{BS}$  bivariado.

A seguir, apresentamos um estudo de influência local baseado no esquema de perturbação de casos.

#### Perturbação de ponderação de casos

Da Figura 3.4(a) notamos que a observação #6 é influente na estimação do parâmetro completo  $\boldsymbol{\theta}$ . Na Figura 3.4(b) a observação #23 é influente na estimação do parâmetro  $\boldsymbol{\alpha}$ . Na Figura 3.4(c) a observação #23 é influente na estimação do parâmetro  $\boldsymbol{\beta}$ . Finalmente, na Figura 3.4(d), a observação #6 é influente na estimação do parâmetro  $\rho$ .

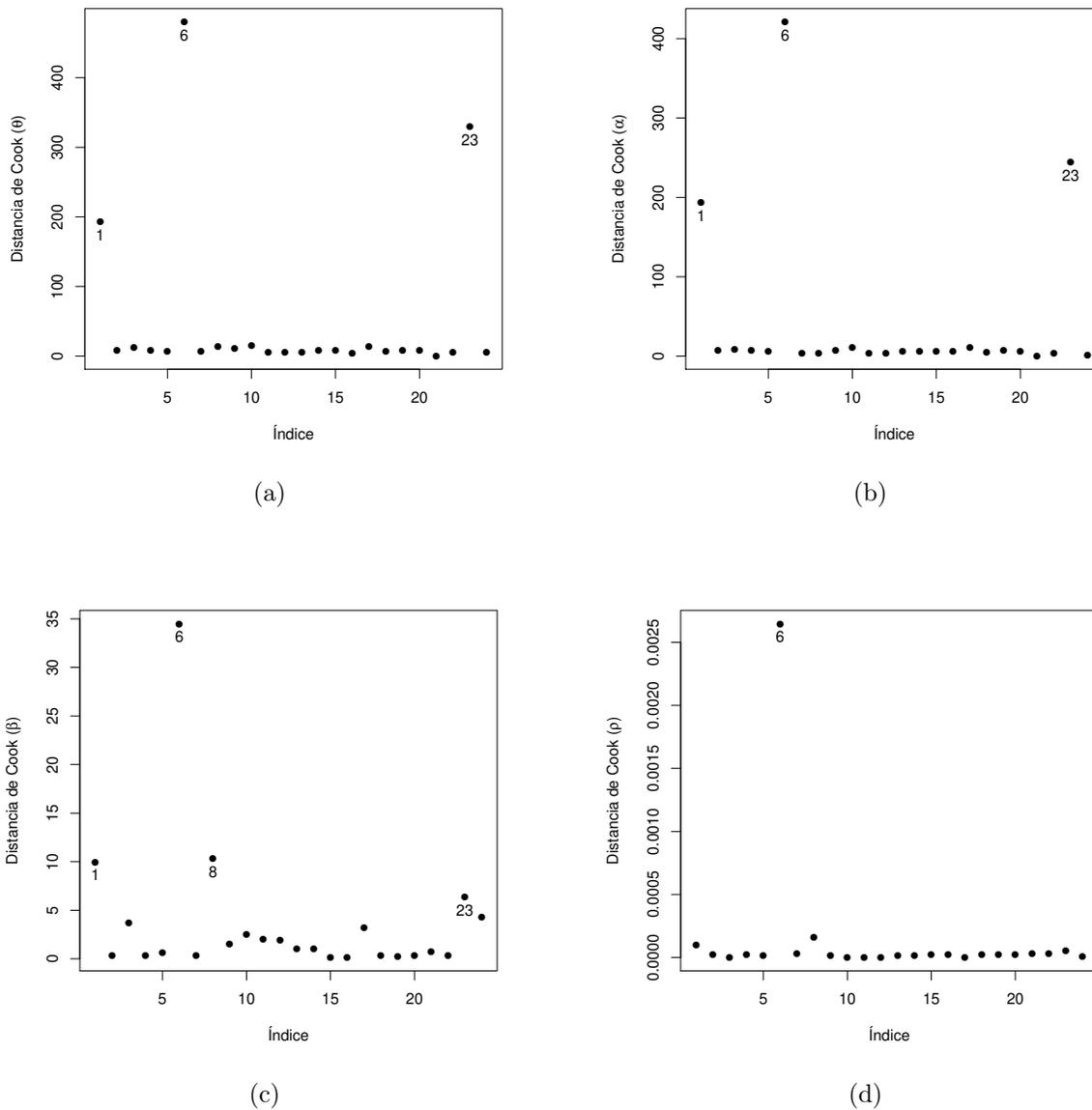


Figura 3.2: Distância Generalizada de Cook para os diferentes parâmetros

Eliminamos as observações #1, #6, #8 e #23 assim na Tabela 3.1, apresentamos a variação das estimativas de  $\theta$  obtidas a través de

$$RC = \frac{|\hat{\theta}_{j[i]} - \hat{\theta}_j|}{\hat{\theta}_j} \times 100, \quad j = 1, \dots, 5 \text{ e } i = 1, \dots, n,$$

Na Figura 3.5(a) mostramos um *index plot* da influência total  $C_i$  definida em (3.3.5), a marca de nível usada para avaliar as observações influentes dada por  $2 \sum_{i=1}^n C_i/n$  é igual a 0,1325 e é

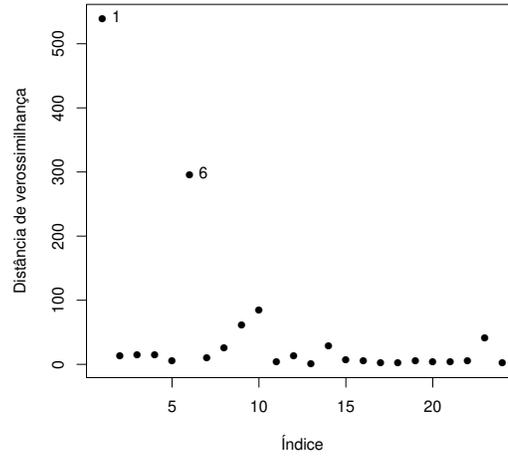


Figura 3.3: Gráfico de  $LD_i$  sob o modelo BS bivariado

Tabela 3.1: Mudança nas estimativas dos parâmetros (em %), excluindo as observações potencialmente influentes.

Obs. Eliminada	$\hat{\alpha}_1$	RC	$\hat{\alpha}_2$	RC	$\hat{\beta}_1$	RC	$\hat{\beta}_2$	RC	$\rho$	RC
1	0.1398	6.238	0.1648	1.5651	0.8211	1.2295	0.8215	0.9262	0.9365	0.2404
6	0.1518	1.827	0.1618	3.3651	0.8335	0.2685	0.8389	1.1725	0.9641	3.1919
8	0.1520	1.972	0.1704	1.7904	0.8327	0.1654	0.8268	0.2878	0.9436	0.9929
23	0.1023	31.38	0.1154	31.0464	0.8513	2.4085	0.8516	2.7011	0.8552	8.4698

indicada no gráfico por uma linha pontilhada. O gráfico indica que as observações #6 e #23 são influentes na estimação de todo o vetor  $\theta$ . Na Figura 3.6(b), mostramos um gráfico de dispersão de  $C_i(\alpha^\top)$  e  $C_i(\beta^\top)$ , as respectivas marcas de nível são dadas por 0,8303 e 1,1012, indicando que somente as observações #6 e #23 são localmente influentes na estimação dos parâmetros  $\alpha$  e  $\beta$ .

Na Figura 3.6(a) mostramos um *index plot* da influência total  $C_i$  definida em (3.3.5), a marca de nível usada para avaliar as observações influentes dada por  $2 \sum_{i=1}^n C_i/n$  é igual a 0,1325 e é indicada no gráfico por uma linha pontilhada. O gráfico indica que as observações #6 e #23 são influentes na estimação completa do  $\theta$ .

Na Figura 3.7(a) mostramos um gráfico de índices da influência total  $C_i(\alpha_1)$  definida em (3.3.5), a marca de nível usada para avaliar as observações influentes dada por  $2 \sum_{i=1}^n C_i(\alpha_1)/n$  é igual a 0.0370 e é indicada no gráfico por uma linha pontilhada. O gráfico indica que as observações #1, #6 e #23 são influentes na estimação completa do parâmetro  $\alpha_1$ . Na Figura 3.7(b) mostramos um *index plot* da influência total  $C_i(\alpha_2)$  definida em (3.3.5), a marca de nível usada para avaliar as observações influentes dada por  $2 \sum_{i=1}^n C_i(\alpha_2)/n$  é igual a 0.0676 e é indicada no gráfico por uma linha pontilhada. O gráfico indica que as observações #1, #6 e #23 são influentes na estimação completa do parâmetro  $\alpha_2$ .

Na Figura 3.8(c) mostramos um *index plot* da influência total  $C_i(\beta_1)$  definida em (3.3.5), a

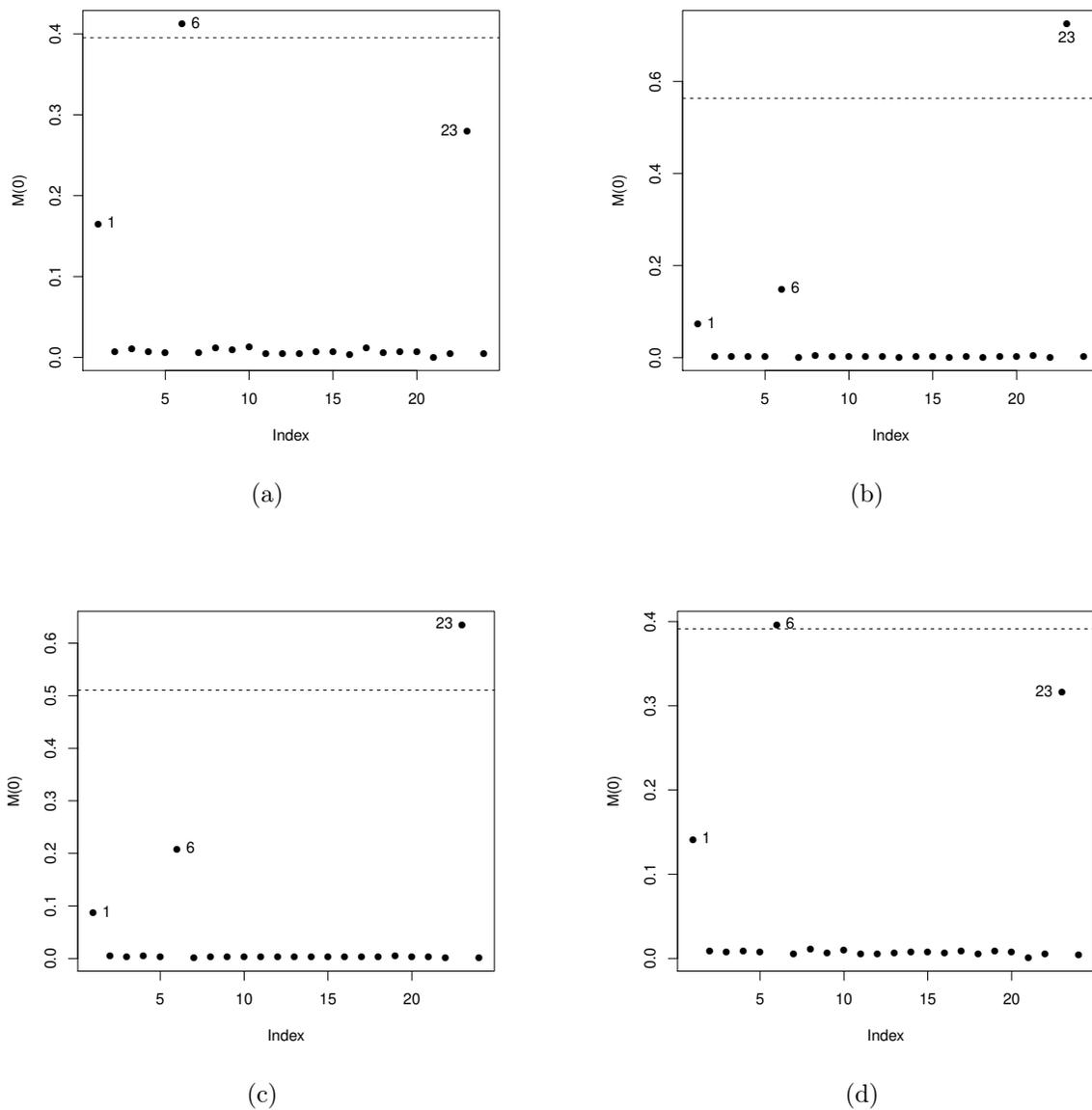


Figura 3.4: Gráfico de  $M(0)$  sob perturbação de ponderação de casos.

marca de nível usada para avaliar as observações influentes dada por  $2 \sum_{i=1}^n C_i(\beta_1)/n$  é igual a 0,0114 e é indicada no gráfico por uma linha pontilhada. O gráfico indica que as observações #1, #6 e #8 são influentes na estimação completa do parâmetro  $\beta_1$ . Na Figura 3.8(c) mostramos um *index plot* da influência total  $C_i(\beta_2)$  definida em (3.3.5), a marca de nível usada para avaliar as observações influentes dada por  $2 \sum_{i=1}^n C_i(\beta_2)/n$  é igual a 0,0130 e é indicada no gráfico por uma linha pontilhada. O gráfico indica que as observações #1, #6 e #8 são influentes na estimação completa do parâmetro  $\beta_2$ .

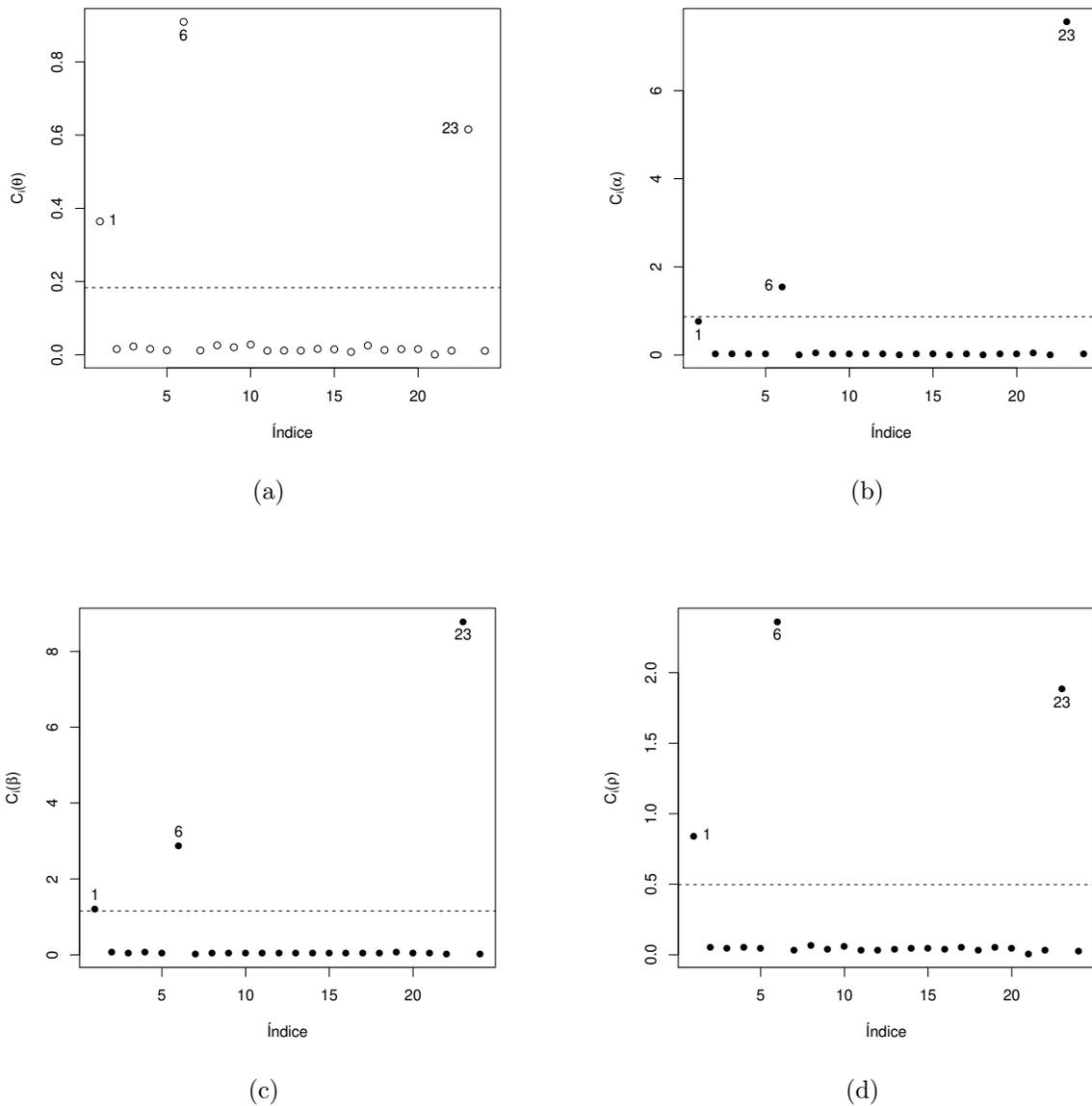


Figura 3.5: Gráfico das medidas de influência local total (a)  $C_i(\boldsymbol{\theta})$  vs observações (b)  $C_i(\boldsymbol{\alpha})$  vs observações (c)  $C_i(\boldsymbol{\beta})$  vs observações e (d)  $C_i(\rho)$  vs observações

Na Figura 3.9(a), mostramos um gráfico de dispersão de  $C_i(\alpha_1)$  vs  $C_i(\beta_1)$ , as respectivas marcas de nível são dadas por 0,0114 e 0,0370, indicando que as observações #1 e #6 são localmente influentes com  $\alpha_1$  e  $\beta_1$ . Na Figura 3.9(b), mostramos um gráfico de dispersão de  $C_i(\alpha_2)$  vs  $C_i(\beta_2)$ , as respectivas marcas de nível são dadas por 0,0130 e 0,0676, indicando que a observação #6 são localmente influente na estimação de  $\alpha_2$  e  $\beta_2$ . Na Figura 3.9(c), mostramos um gráfico de dispersão de  $C_i(\alpha_2)$  vs  $C_i(\beta_1)$ , as respectivas marcas de nível são dadas por 0,0113 e 0,0676, indicando que

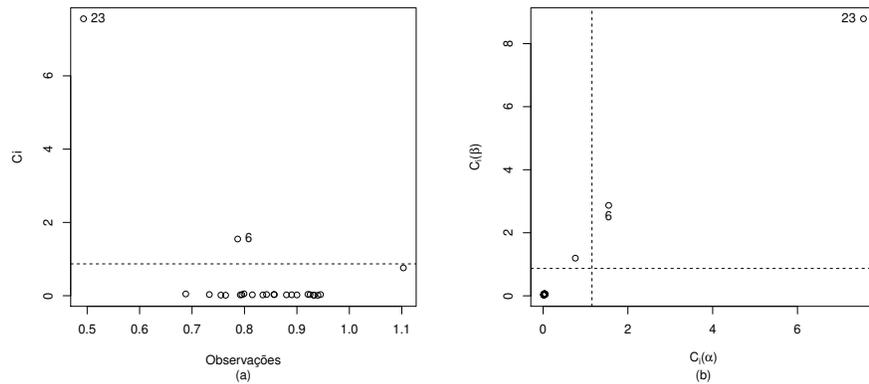


Figura 3.6: Ponderação de casos (a) Gráfico de influência total  $C_i$  vs observações. (b) Gráfico das medidas de influência local  $C_i(\alpha)$  vs  $C_i(\beta)$

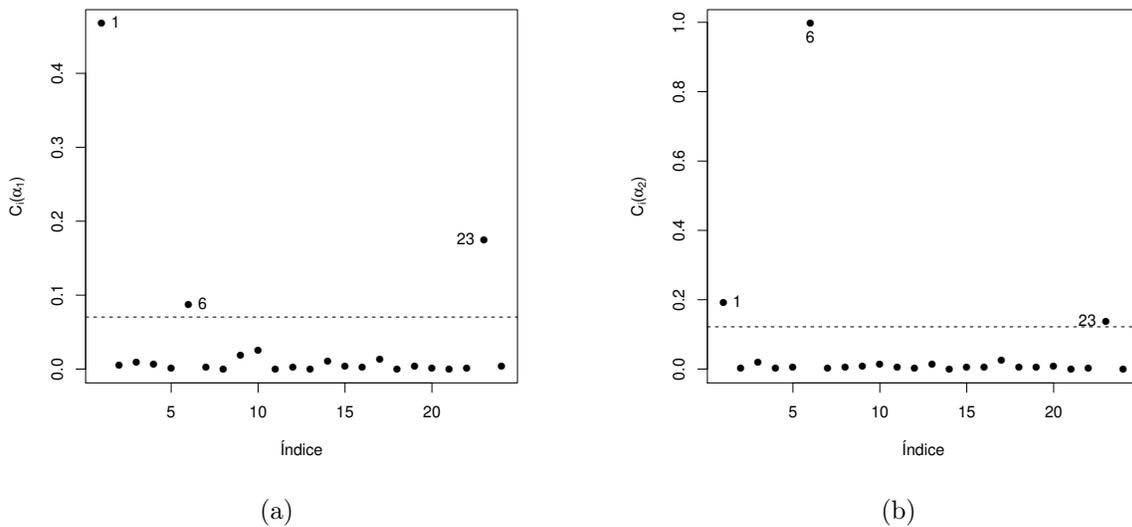


Figura 3.7: Gráfico das medidas de influência local (a)  $C_i(\alpha_1)$  vs observações (b)  $C_i(\alpha_2)$  vs observações

a observação #6 são localmente influentes na estimação de  $\alpha_2$  e  $\beta_1$ . Na Figura 3.9(d), mostramos um gráfico de dispersão de  $C_i(\alpha_1)$  vs  $C_i(\beta_2)$ , as respectivas marcas de nível são dadas por 0,0130 e 0,0370, indicando que a observação #6 são localmente influentes na estimação de  $\alpha_1$  e  $\beta_2$ .

Na Figura 3.10(a), mostramos um gráfico de dispersão de  $C_i(\alpha_1)$  vs  $C_i(\alpha_2)$ , as respectivas marcas de nível são dadas por 0,0370 e 0,0676, indicando que as observações #1, #6 e #23 são localmente influentes com  $\alpha_1$  e  $\alpha_2$ . Na Figura 3.10(b), mostramos um gráfico de dispersão de

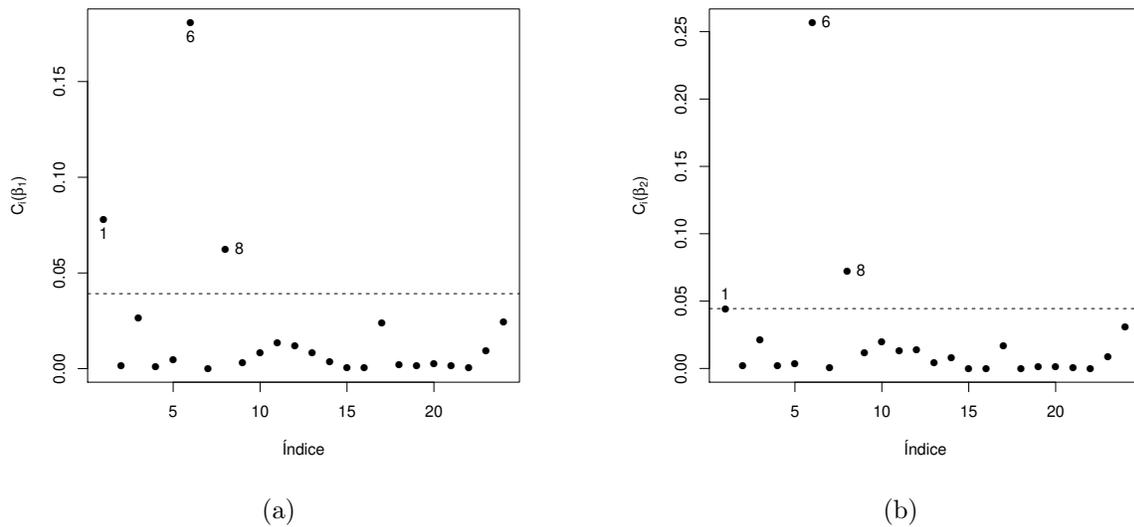


Figura 3.8: Gráfico das medidas de influência local (a)  $C_i(\beta_1)$  vs observações e (b)  $C_i(\beta_2)$  vs observações

$C_i(\beta_1)$  vs  $C_i(\beta_2)$ , com as respectivas marcas de nível são dadas por 0,0114 e 0,0130, indicando que as observações #1, #8 e #6 são localmente influentes na estimação de  $\beta_1$  e  $\beta_2$

Na Figura 3.11(a), mostramos um gráfico de dispersão de  $C_i(\rho)$  vs  $C_i(\alpha)$ , as respectivas marcas de nível são dadas por 0.8303 e 0.4585, indicando que as observações #6 e #23 são localmente influentes na estimação de  $\alpha$  e  $\rho$ . Na Figura 3.11(b), mostramos um gráfico de dispersão de  $C_i(\rho)$  vs  $C_i(\alpha)$ , as respectivas marcas de nível são dadas por 1.1012 e 0.4585, indicando que as observações #6 e #23 são localmente influentes na estimação de  $\alpha$  e  $\rho$ . Na Figura 3.11(a), mostramos um gráfico de dispersão de  $C_i(\rho)$  vs  $C_i(\alpha)$ , as respectivas marcas de nível são dadas por 0.8303 e 0.4585, indicando que as observações #6 e #23 são localmente influentes na estimação de  $\alpha$  e  $\rho$ . Na Figura 3.11(b), mostramos um gráfico de dispersão de  $C_i(\rho)$  vs  $C_i(\alpha)$ , as respectivas marcas de nível são dadas por 1.1012 e 0.4585, indicando que as observações #6 e #23 são localmente influentes na estimação de  $\alpha$  e  $\rho$

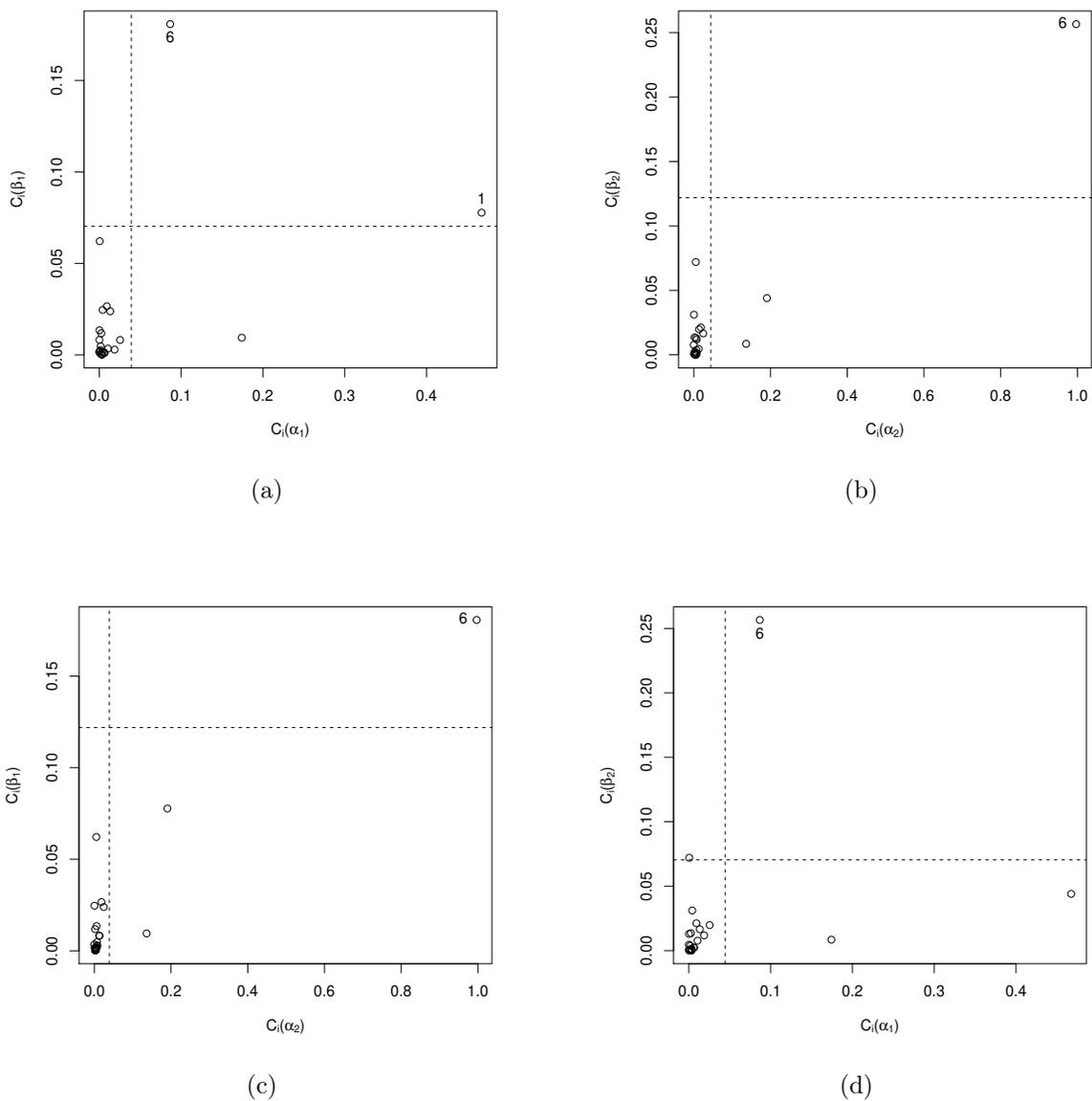


Figura 3.9: Gráfico das medidas de influência local (a)  $C_i(\alpha_1)$  vs  $C_i(\beta_1)$  (b)  $C_i(\alpha_2)$  vs  $C_i(\beta_2)$  (c)  $C_i(\alpha_2)$  vs  $C_i(\beta_1)$  e (d)  $C_i(\alpha_1)$  vs  $C_i(\beta_2)$

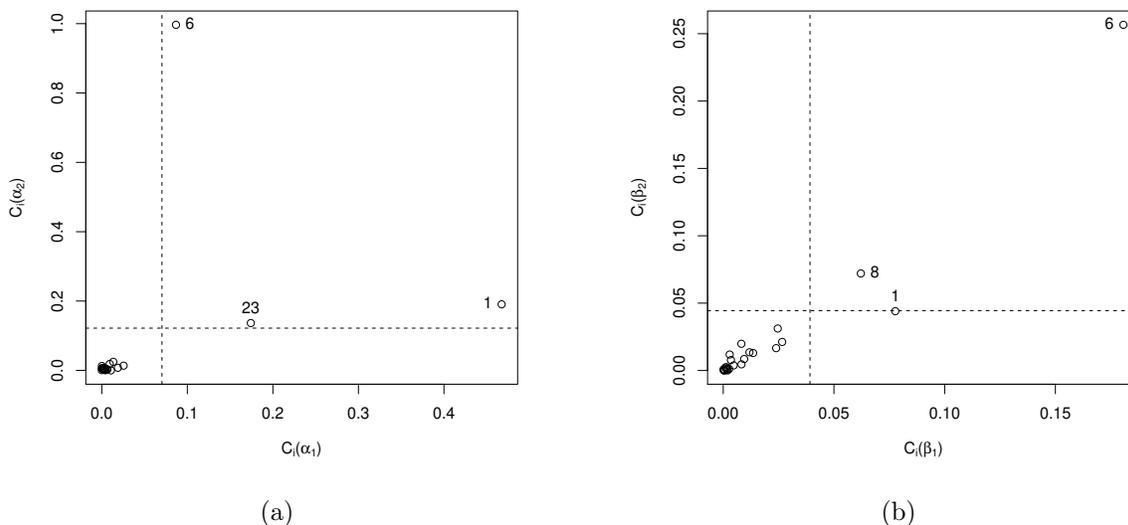


Figura 3.10: (a) Gráfico das medidas de influência local  $C_i(\alpha_1)$  vs  $C_i(\alpha_2)$  (b) Gráfico das medidas de influência local  $C_i(\beta_1)$  vs  $C_i(\beta_2)$

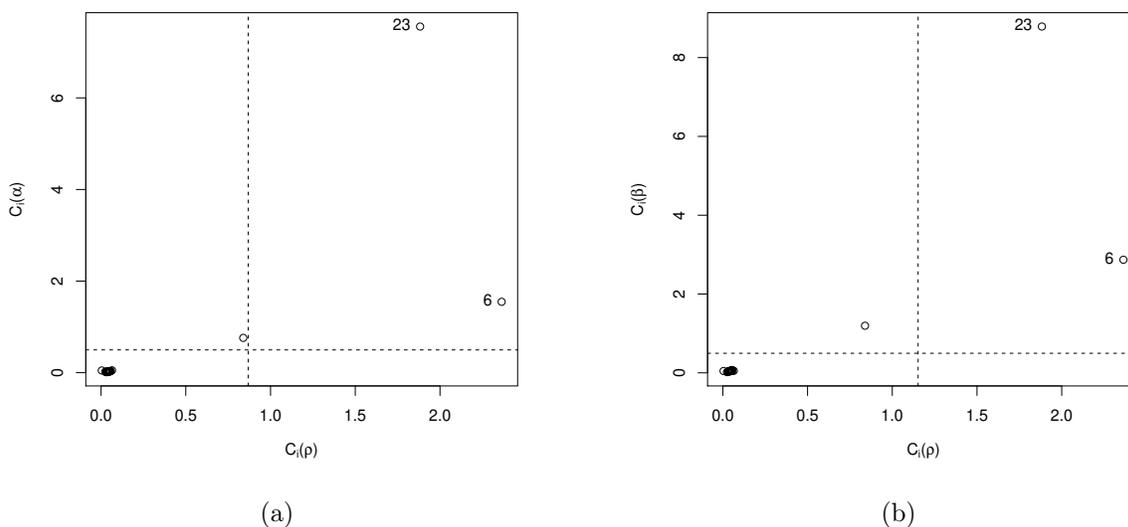


Figura 3.11: (a) Gráfico das medidas de influência local  $C_i(\rho)$  vs  $C_i(\alpha)$  (b) Gráfico das medidas de influência local  $C_i(\rho)$  vs  $C_i(\beta)$



# Capítulo 4

## Análise de sobrevivência para um modelo BS bivariado

### 4.1 Introdução

Dados de tempo de sobrevivência é um termo usado para se referir-se a dados de medição do tempo de ocorrências de determinados eventos, tal como uma falha, morte, o desenvolvimento de uma determinada doença. Esses dados são também referidos como dados da vida ou dados de tempo de falha. Os dados de sobrevivência são caracterizados pelos tempos de falha. Dados de tempo de sobrevivência frequentemente vêm da medicina Janurová & Briš (2013), mas podem vir de outras áreas aplicadas, como demografia Yashin & Manton (1997), engenharia Nelson (1972), economia e ciências sociais.

Precisamente, “a análise dos dados que correspondem ao tempo a partir de uma origem de tempo bem definido, até à ocorrência de um acontecimento específico ou de ponto final” Collett (1994). Pelo *tempo* queremos dizer anos, meses, semanas ou dias desde o início do acompanhamento de uma unidade experimental até que ocorra um evento. Alternativamente, o tempo pode referir-se a idade de um indivíduo quando ocorre um evento.

Dados multivariados de vida surgem quando cada unidade de estudo pode ter vários eventos de interesse associados ou quando existe algum agrupamento natural de indivíduos, o que induz a dependência entre os tempos de vida do mesmo grupo. Estes dados são comumente encontrados em investigações científicas, pois cada objeto de estudo pode experimentar vários eventos ou porque o estudo envolve vários membros de cada grupo.

A distribuição normal ocupa uma posição central na teoria estatística, de bem como a avaliação de integrais normais univariadas e multivariadas tem sido de interesse para os pesquisadores por um longo tempo. Muitas expressões diferentes estão disponíveis para a função de distribuição normal univariada. Estas são de quatro tipos: série convergente, frações contínuas, da série assintótica e aproximações empíricas. As fórmulas clássicas foram reunidas por Zelen e Severo (1964).

A função de sobrevivência e taxa de falha univariada depende do cálculo da fda de uma distribuição de interesse, assim da mesma maneira para o caso bivariado a função de sobrevivência e taxa de falha bivariada precisam de uma aproximação numérica, de suas integrais da fda bivariada. Por isso foram desenvolvidos diferentes aproximações.

## 4.2 A fda da distribuição normal bivariada

A fda, seja univariada ou multivariada, tem de ser avaliada numericamente. Existem inúmeros algoritmos disponíveis, muitos deles tendo sido aperfeiçoados, levando à avaliação mais rápida com maior precisão.

Cálculos que envolvem de distribuições de probabilidade bivariadas são necessários para muitas aplicações estatísticas. Embora algoritmos confiáveis, eficientes e precisos para probabilidades univariadas, os algoritmos de alta qualidade para os cálculos de distribuição de probabilidade bivariadas apenas mais recentemente tornarem-se disponível. Existe uma necessidade para estes algoritmos como componentes em bibliotecas de cálculo de estatística e pacotes no R, por exemplo o pacote `mvtnorm()`. É também uma necessidade crescente para esses algoritmos como um meio para a computação eficiente com bons limites para probabilidades multivariadas. Estes limites podem, então, ser usados para fornecer métodos de cálculo mais eficientes para o cálculo da probabilidade multivariada, veja Gassmann (2000) e Genz *et al.* (2000).

Vejamos a integral da fda da normal univariada,

$$F_T(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{u}{2}} du.$$

Como é bem conhecido, uma solução da forma fechada não existe para a integral como acima, de modo que uma aproximação numérica precisa ser implementada. O mais comum é uma aproximação que envolve uma exponencial e um quinto grau polinomial, dada em Abramowitz & Stegun (1974), e repetido por Hull (2002) e Haug (1998). Divgi (1979) trabalha com séries convergentes que requerem a integração da função de distribuição. Kerridge e Cook (1976) derivaram uma nova série convergente. Marsaglia (2004) propôs uma alternativa muito simples e intuitiva, mas poderosa, que se baseia-se na expansão de Taylor da razão de Mills ou funções semelhantes.

A integral da fda da normal bivariada é a função seguinte

$$F_{\mathbf{T}}(\mathbf{t}) = F_{\mathbf{T}}(t_1, t_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} e^{-\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)} dudv.$$

Assim, como no caso univariado, uma solução fechada para as integrais do caso bivariado não existe aproximação analítica. Por isso é necessário a avaliação numérica da função de distribuição normal bivariada para uma série de propósitos probabilísticos e estatísticos, assim uma aproximação numérica precisa ser implementada. Atualmente existem hoje muitos algoritmos disponíveis

para cálculos de probabilidades normais bivariadas, mas a qualidade desses algoritmos tem variações significativas.

Gupta (1963) apresentou um trabalho sobre integrais de probabilidade multivariadas e funções relacionadas e começou pelo caso bivariado. Um método mais adequado para os cálculos feitos com computador foi desenvolvido por Owen (1956). Ele introduziu pela primeira vez o  $T$ -função como um termo-chave para expressar a fda normalidade bivariada. Donnelly (1973a) codificou diretamente o método de Owen e formou o algoritmo Owen-Donnelly, que tem sido o mais amplamente implementado de todos os algoritmos normais bivariados, e é possível obter valores do fda da normal bivariado com uma precisão de 15 dígitos. Grande parte da pesquisa desde dos cálculos iniciais das integrais normais bivariados é baseada neste trabalho.

Terza & Welland (1991) apresentaram uma comparação de oito algoritmos (Owen, Young, Daley, Drezner, Divgi, Bouver e Bargmann (1979), Parrish e Bargmann (1981), Terza e Welland) para avaliar a fda da normal bivariada.

Recentemente Meyer (2013) estendeu a abordagem de Marsaglia (2004) ao caso bivariado, que consiste na redução da avaliação da fda bivariada para avaliação(s) de uma função univariada, ou seja, para a distribuição normal bivariada cumulativa, sobre a diagonal, e a expansão de Taylor dessa função.

### 4.2.1 Aproximação de Tsay and Ke (2009)

Seja  $\Phi_2(\cdot; \rho)$  a fda da distribuição normal bivariada padrão do vetor  $\mathbf{Z} = (Z_1, Z_2)^\top \sim N_2(\mathbf{0}, \Sigma)$  com  $\sigma_{11} = \sigma_{22} = 1$  e coeficiente de correlação  $\rho$ . A quantidade  $\Phi_2(\cdot; \rho)$  está estreitamente relacionada com

$$L(p, q; \rho) = P(Z_1 > p, Z_2 > q). \quad (4.2.1)$$

através da relação

$$\Phi_2(p, q; \rho) = 1 - L(p, -\infty; \rho) - L(-\infty, q; \rho) + L(p, q; \rho). \quad (4.2.2)$$

Para calcular  $\Phi_2(\cdot; \rho)$ , as tabelas de  $L(\cdot; \rho)$  são apresentados por Ederton et al. (1930), e Pearson (1931) entre outros. No entanto, não é conveniente utilizar as tabelas de  $L(\cdot; \rho)$  para calcular  $\Phi_2(\cdot; \rho)$  fazendo uso da relação (4.2.2), pois são três valores  $L(\cdot; \rho)$  os necessários para poder completar a tarefa. Consequentemente, vários métodos de aproximação para o cálculo de  $L(\cdot; \rho)$  são desenvolvidos, incluindo Owen (1956) e Cox & Wermuth (1991).

Métodos alternativos de quadratura têm sido propostos para avaliar  $\Phi_2(\cdot; \rho)$  baseados numa única integral de  $\Phi(\cdot)$ , a saber

$$\Phi_2(p, q; \rho) = \int_{-\infty}^q \Phi(av + b)\phi(\nu) d\nu, \quad (4.2.3)$$

em que  $a = -\rho/(1 - \rho^2)^{1/2}$ ,  $b = \rho/(1 - \rho^2)^{1/2}$  com  $a, b \in \mathbb{R}$  e  $\rho \in (-1, 1)$ . Da equação (7.1.1) de Abramowitz & Stegun (1974), denotam a função de erro como:

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt = 2 \int_0^{\sqrt{2}z} \phi(t) dt, \quad (4.2.4)$$

em que  $\operatorname{erf}(-z) = -\operatorname{erf}(z)$ . Tsay & Ke (2009) apresentaram o desenvolvimento de (4.2.3) que descrevemos a seguir:

$$\begin{aligned} \Phi_2(p, q; \rho) &= \int_{-\infty}^q \left[ \int_{-\infty}^{a\nu+b} \phi(\zeta) d(\zeta) \right] \phi(\nu) d\nu \\ &= \int_{-\infty}^q \left[ \int_{-\infty}^0 \phi(\zeta) d(\zeta) + \int_0^{a\nu+b} \phi(\zeta) d(\zeta) \right] \phi(\nu) d\nu \\ &= \int_{-\infty}^q \left[ \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left( \frac{a\nu+b}{\sqrt{2}} \right) \right] \phi(\nu) d\nu \\ &= \int_{-\infty}^q \frac{1}{2} \phi(\nu) d\nu + \int_0^q \frac{1}{2} \operatorname{erf} \left( \frac{a\nu+b}{\sqrt{2}} \right) \phi(\nu) d\nu \\ &= \int_{-\infty}^0 \frac{1}{2} \phi(\nu) d\nu + \int_0^q \frac{1}{2} \phi(\nu) d\nu + \int_{-\infty}^q \frac{1}{2} \operatorname{erf} \left( \frac{a\nu+b}{\sqrt{2}} \right) \phi(\nu) d\nu \\ &= \frac{1}{4} + \frac{1}{2} \int_0^Q \phi(\nu) d\nu + \int_{-\infty}^Q \frac{1}{2} \operatorname{erf} \left( \frac{a\nu+b}{\sqrt{2}} \right) \phi(\nu) d\nu \\ &= \frac{1}{4} + \frac{1}{4} \operatorname{erf} \left( \frac{q}{\sqrt{2}} \right) + \int_{-\infty}^q \frac{1}{2} \operatorname{erf} \left( \frac{a\nu+b}{\sqrt{2}} \right) \phi(\nu) d\nu. \end{aligned}$$

Os autores Tsay & Ke (2009) definem  $y = (a\nu + b)/\sqrt{2}$ , para obter:

$$\begin{aligned} \Phi_2(p, q; \rho) &= \frac{1}{4} + \frac{1}{4} \operatorname{erf} \left( \frac{q}{\sqrt{2}} \right) + \int_{-\infty}^0 \frac{\sqrt{2}}{2a} \operatorname{erf}(y) \phi \left( \frac{\sqrt{2}y - b}{a} \right) dy \\ &\quad + \int_0^{\frac{aQ+b}{\sqrt{2}}} \frac{\sqrt{2}}{2a} \operatorname{erf}(y) \phi \left( \frac{\sqrt{2}y - b}{a} \right) dy. \end{aligned}$$

Usando a mudança de variável,  $Z = -y$ , obtemos:

$$\Phi_2(p, q; \rho) = \frac{1}{4} + \frac{1}{4} \operatorname{erf} \left( \frac{q}{\sqrt{2}} \right) + A + B,$$

em que  $A = -\int_0^\infty \frac{\sqrt{2}}{2a} \operatorname{erf}(z) \phi \left( \frac{\sqrt{2}z - b}{a} \right) dz$ ,  $B = \int_0^{\frac{aQ+b}{\sqrt{2}}} \frac{\sqrt{2}}{2a} \operatorname{erf}(y) \phi \left( \frac{\sqrt{2}y - b}{a} \right) dy$ .

Agora, a análise se concentrará nas expressões A e B acima. Assim, Tsay & Ke (2009) apresentam pela primeira vez uma aproximação de  $\operatorname{erf}(x)$  por

$$g(x) = 1 - e^{c_1x + c_2x^2},$$

para  $x \geq 0$  onde  $c_1$  e  $c_2$  são escolhidos para fazer  $g(x)$  o mais próximo possível a  $erf(x)$ . As simulações realizadas em Tsay & Ke (2009) revela claramente a potência de  $g(x) = 1 - e^{c_1x + c_2x^2}$  na aproximação  $erf(x)$ . Veja Tsay & Ke (2009) para mais detalhes.

**Teorema 4.2.1.**  $\Phi_2(p, q; \rho)$  pode ser aproximado por  $F_{app}(p, q; \rho)$  tal como:

1. Se  $a > 0$  e  $aq + b \geq 0$ , então

$$F_{app}(p, q; \rho) = \frac{1}{2} \left[ erf \left( \frac{q}{\sqrt{2}} \right) + erf \left( \frac{b}{\sqrt{2a}} \right) \right] + \frac{1}{4\sqrt{1-a^2c_2}} e^{\frac{a^2c_1^2 - 2\sqrt{2}bc_1 + 2b^2c_2}{4(1-a^2c_2)}} \times \\ \left[ 1 - erf \left( \frac{\sqrt{2}b - a^2c_1}{2a\sqrt{1-a^2c_2}} \right) \right] - \frac{1}{4\sqrt{1-a^2c_2}} e^{\frac{a^2c_1^2 + 2\sqrt{2}bc_1 + 2b^2c_2}{4(1-a^2c_2)}} \times \\ \left[ erf \left( \frac{\sqrt{2}q - \sqrt{2}a^2c_2q - \sqrt{2}abc_2 - ac_1}{2\sqrt{1-a^2c_2}} \right) + erf \left( \frac{a^2c_1 + \sqrt{2}b}{2a\sqrt{1-a^2c_2}} \right) \right];$$

2. Se  $a = 0$ , então  $\rho = 0$ , pela propriedade de independência de duas variáveis, temos que:

$$F_{app}(p, q; \rho) = \Phi(p)\Phi(q);$$

3. Se  $a > 0$  e  $aq + b < 0$ , então

$$F_{app}(p, q; \rho) = \frac{1}{4\sqrt{1-a^2c_2}} e^{\frac{a^2c_1^2 - 2\sqrt{2}bc_1 + 2b^2c_2}{4(1-a^2c_2)}} \left[ 1 + erf \left( \frac{\sqrt{2}q - \sqrt{2}a^2c_2q - \sqrt{2}abc_2 + ac_1}{2\sqrt{1-a^2c_2}} \right) \right];$$

4. Se  $a < 0$  e  $aq + b \geq 0$ , então

$$F_{app}(p, q; \rho) = \frac{1}{2} + \frac{1}{2} erf \left( \frac{q}{\sqrt{2}} \right) - \frac{1}{4\sqrt{1-a^2c_2}} e^{\frac{a^2c_1^2 - 2\sqrt{2}bc_1 + 2b^2c_2}{4(1-a^2c_2)}} \times \\ \left[ 1 + erf \left( \frac{\sqrt{2}q - \sqrt{2}a^2c_2q - \sqrt{2}abc_2 - ac_1}{2\sqrt{1-a^2c_2}} \right) \right];$$

5. Se  $a < 0$  e  $aq + b < 0$ , então

$$F_{app}(p, q; \rho) = \frac{1}{2} - \frac{1}{2} erf \left( \frac{q}{\sqrt{2}} \right) - \frac{1}{4\sqrt{1-a^2c_2}} e^{\frac{a^2c_1^2 - 2\sqrt{2}bc_1 + 2b^2c_2}{4(1-a^2c_2)}} \times \\ \left[ 11 erf \left( \frac{\sqrt{2}b + a^2c_1}{2a\sqrt{1-a^2c_2}} \right) \right] + \frac{1}{4\sqrt{1-a^2c_2}} e^{\frac{a^2c_1^2 - 2\sqrt{2}bc_1 + 2b^2c_2}{4(1-a^2c_2)}} \times \\ \left[ erf \left( \frac{\sqrt{2}q - \sqrt{2}a^2c_2q - \sqrt{2}abc_2 + ac_1}{2\sqrt{1-a^2c_2}} \right) + erf \left( \frac{-a^2c_1 + \sqrt{2}b}{2a\sqrt{1-a^2c_2}} \right) \right],$$

em que  $a$  e  $b$  são definidos em (4.2.3),  $c_1 = -1.0950081470333$ , e  $c_2 = -0.75651138383854$ .

**Demonstração.** De Abramowitz & Stegun (1974) as equações (7.1.1) e (7.4.32) serão usadas para a demonstração para um uso posterior, assim como (4.2.4)

$$\int e^{-(kx^2+2mx+n)} = \frac{1}{2} \sqrt{\frac{\pi}{k}} e^{\frac{m^2-kn}{k}} \operatorname{erf} \left( \sqrt{k}x + \frac{m}{\sqrt{k}} \right) + C, \quad k \neq 0,$$

em que  $C$  é uma constante finita. Para fins expositivos, dividimos a análise em cinco partes, i.e., ( $a > 0$  e  $aq + b \geq 0$ ), ( $a = 0$  e  $aq + b \geq 0$ ), ( $a \geq 0$  e  $aq + b < 0$ ), ( $a < 0$  e  $aq + b \geq 0$ ), e ( $a < 0$  e  $aq + b < 0$ ).

Para o caso  $a > 0$  e  $aq + b \geq 0$ , temos que

$$\begin{aligned} A &= - \int_0^\infty \frac{\sqrt{2}}{2a} \operatorname{erf}(a) \phi \left( \frac{-\sqrt{2}z - b}{a} \right) da \\ &\approx - \int_0^\infty \frac{1}{2a} \left( 1 - e^{c_1z + c_2z^2} \right) \frac{e^{-\frac{1}{2} \left( \frac{-\sqrt{2}z - b}{a} \right)^2}}{\sqrt{2\pi}} dz \\ &= - \frac{1}{2a\sqrt{\pi}} \left[ \int_0^\infty e^{-\left( \frac{1}{a^2}z^2 + 2\frac{b}{\sqrt{2}a^2}z + \frac{b^2}{2a^2} \right)} dz - \int_0^\infty e^{-\left( \frac{1-a^2c_2}{a^2}z^2 + 2\frac{\sqrt{2}b-a^2c_1}{\sqrt{2}a^2}z + \frac{b^2}{2a^2} \right)} dz \right]. \end{aligned}$$

De Abramowitz & Stegun (1974), da equação (7.4.32) a expressão  $A$  pode ser aproximada por

$$A \approx -\frac{1}{4} + \frac{1}{4} \operatorname{erf} \left( \frac{b}{\sqrt{2}a} \right) + \frac{1}{4\sqrt{1-a^2c_2}} e^{\frac{a^2c_1^2 - 2\sqrt{2}bc_1 + 2b^2c_2}{4(1-a^2c_2)}} \left[ 1 - \operatorname{erf} \left( \frac{\sqrt{2}b - a^2c_1}{2a\sqrt{1-a^2c_2}} \right) \right].$$

Analogamente ao desenvolvimento descrito acima, podemos também aproximar o valor  $B$  para o caso  $a > 0$  e  $aq + b \geq 0$  como:

$$\begin{aligned} B &\approx \frac{1}{4} \left[ \operatorname{erf} \left( \frac{q}{\sqrt{2}} \right) + \operatorname{erf} \left( \frac{b}{\sqrt{2}a} \right) \right] - \frac{1}{4\sqrt{1-a^2c_2}} e^{\frac{a^2c_1^2 + 2\sqrt{2}bc_1 + 2b^2c_2}{4(1-a^2c_2)}} \times \\ &\quad \left[ \operatorname{erf} \left( \frac{\sqrt{2}aq - \sqrt{2}a^3c_2Q - \sqrt{2}a^2bc_2 - a^2c_1}{2a\sqrt{1-a^2c_2}} \right) + \operatorname{erf} \left( \frac{a^2c_1 + \sqrt{2}b}{2a\sqrt{1-a^2c_2}} \right) \right]. \end{aligned}$$

Combinando os resultados anteriores sobre  $A$  e  $B$ , nota-se que  $\phi_2(p, q; \rho)$  pode ser bem aproximada por  $F_{app}(p, q; \rho)$  para o caso ( $a > 0, aq + b \geq 0$ ) como:

$$\begin{aligned} F_{app}(p, q; \rho) &= \frac{1}{2} \left[ \operatorname{erf} \left( \frac{q}{\sqrt{2}} \right) + \operatorname{erf} \left( \frac{b}{\sqrt{2}a} \right) \right] + \frac{1}{4\sqrt{1-a^2c_2}} e^{\frac{a^2c_1^2 - 2\sqrt{2}bc_1 + 2b^2c_2}{4(1-a^2c_2)}} \times \\ &\quad \left[ 1 - \operatorname{erf} \left( \frac{\sqrt{2}b - a^2c_1}{2a\sqrt{1-a^2c_2}} \right) \right] - \frac{1}{4\sqrt{1-a^2c_2}} e^{\frac{a^2c_1^2 + 2\sqrt{2}bc_1 + 2b^2c_2}{4(1-a^2c_2)}} \times \\ &\quad \left[ \operatorname{erf} \left( \frac{\sqrt{2}q - \sqrt{2}a^2c_2q - \sqrt{2}abc_2 - ac_1}{2\sqrt{1-a^2c_2}} \right) + \operatorname{erf} \left( \frac{a^2c_1 + \sqrt{2}b}{2a\sqrt{1-a^2c_2}} \right) \right]. \end{aligned}$$

Seguindo os procedimentos anteriores, obtemos os quatro casos restantes, sem dificuldades. Combinando os resultados dos referidos cinco casos, temos demonstrado o Teorema (4.2.1). ■

Existem várias vantagens de usar  $F_{app}$  na aproximação  $\Phi_2(\cdot; \rho)$ . Em primeiro lugar, ao contrário de muitos algoritmos de integração numérica híbridas, as implementações de  $F_{app}$  não dependem do valor de  $\rho$  contanto que não seja igual a  $\pm 1$ .

### 4.2.2 Aproximação de Albers e Kallenberg (1994)

Nesta parte, vamos descrever o procedimento proposto por Albers & Kallenberg (1994) para uma aproximação da função de distribuição normal bivariada, com ênfase em situações em que o coeficiente de correlação seja grande, e, também apresenta uma aproximação para obter  $L(p, q; \rho)$ . Para  $\rho^2 \geq 1/2$  ou  $\rho > 1/\sqrt{2}$ , obtemos aproximação de primeiro, ordem e segundo, ordem, respectivamente, de  $\Phi_2(\cdot; \rho)$ ,

$$\begin{aligned}\Phi_2(p, q; \rho) &\approx \Phi(q) - g(p, q), \text{ e} \\ \Phi_2(p, q; \rho) &\approx \Phi(q) - g(p, q) - h(p, q).\end{aligned}$$

Para  $\rho < 2^{-1/2}$ , temos aproximações de primeira e segunda ordem, respetivamente de  $L(p, q; \rho)$ :

$$L(p, q; \rho) \approx P(Z_1 > p) - g(p, q), \text{ e} \tag{4.2.5}$$

$$L(p, q; \rho) \approx P(Z_1 > p) - g(p, q) - h(p, q), \tag{4.2.6}$$

em que

$$g(p, q) = \Phi(-c) \{ \Phi(p + \theta(c_1 - c)) - \Phi(p) \},$$

e

$$h(p, q) = \frac{1}{2} \theta^2 \Phi(-c) \phi'(p + \theta(c_1 - c)) \{ 1 + cc_1 - c_1^2 \},$$

com

$$c = (\rho a - b) / \sqrt{1 - \rho^2}, \quad \theta = (\sqrt{1 - \rho^2}) / \rho \text{ e } c_1 = \phi(c) / \Phi(-c).$$

### 4.2.3 Aproximação de Cox e Wermuth (1991)

Nesta parte, vamos descrever o procedimento para obter  $L(p, q; \rho)$  proposto por Cox & Wermuth (1991), veja (4.2.1). Seja  $\mathbf{Z} = (Z_1, Z_2)^\top \sim N_2(\mathbf{0}, \Sigma)$ , com  $\sigma_{11} = \sigma_{22} = 1$  e coeficiente de correlação  $\rho$ . Usando a propriedade de probabilidade condicional, temos que

$$\begin{aligned}L(p, q; \rho) &= P(Z_1 > p)P(Z_2 > q | Z_1 > p) \\ &= \Phi(-p)E \left\{ \Phi \left( \frac{\rho Z_1 - q}{\sqrt{1 - \rho^2}} \right) | Z_1 > p \right\},\end{aligned}$$

em que a segunda forma segue porque  $Z_1$  dado  $Z_1 = z_1$  é normalmente distribuída com média  $\rho x$  e variância  $1 - \rho^2$ . Se  $p, q, \rho > 0$ , podemos como uma primeira aproximação

$$L(p, q; \rho) \approx \Phi(-p)\Phi\{\xi(p, q; \rho)\}, \quad (4.2.7)$$

em que

$$\xi(p, q; \rho) = \frac{\rho\mu(p) - q}{\sqrt{1 - \rho^2}},$$

e

$$\mu(p) = E(Z_1|Z_1 > p) = \phi(p)/\Phi(-p). \quad (4.2.8)$$

em que  $\phi(\cdot)$  é a fdp da normal padrão. O refinamento de (4.2.7) é,

$$L(p, q; \rho) \approx \Phi(-p) \left[ \Phi\{\xi(p, q; \rho)\} - \frac{1}{2}\rho^2(1 - \rho^2)^{-1}\xi(p, q; \rho)\phi\{\xi(p, q; \rho)\}\sigma^2(p) \right], \quad (4.2.9)$$

em que  $\sigma^2(p) = \text{Var}(Z_1|Z_1 > p) = 1 + \mu(p) - \mu^2(p)$ .

O método de Cox & Wermuth (1991) apresenta problemas de aproximação quando  $p = q$  e  $\rho$  é grande.

## 4.3 Dados de tempo de vida bivariado

### 4.3.1 Função de sobrevivência bivariada

No caso bivariado seguindo as ideias de Basu (1971) temos que se  $\mathbf{T} = (T_1, T_2) \sim BS_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$  a função de sobrevivência bivariada é definida por

$$S_{\mathbf{T}}(\mathbf{t}) = \int_{a_{t_1}}^{\infty} \int_{a_{t_2}}^{\infty} f(u, v) du dv, \quad \mathbf{t} \in \mathbb{R}_+^2, \quad (4.3.1)$$

em que  $f(u, v)$  é a fdp da normal bivariada padrão. Porém, usando as propriedades estruturais de distribuição bivariada podemos expressar (4.3.1) como segue

$$S_{\mathbf{T}}(\mathbf{t}) = 1 - \Phi(a(t_1; \alpha_1, \beta_1)) - \Phi(a(t_2; \alpha_2, \beta_2)) + \Phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho), \quad (4.3.2)$$

em que  $\Phi(\cdot)$  e  $\Phi_2(\cdot; \rho)$  são as fdp de uma distribuição normal univariada padrão  $N(0, 1)$  e uma distribuição bivariada  $N_2(\mathbf{0}, \boldsymbol{\Sigma})$  respectivamente.

A função densidade de probabilidade conjunta do  $\mathbf{T}$  é dada por:

$$f_{\mathbf{T}}(\mathbf{t}) = \frac{\partial^2 S_{\mathbf{T}}(\mathbf{t})}{\partial t_1 \partial t_2}. \quad (4.3.3)$$

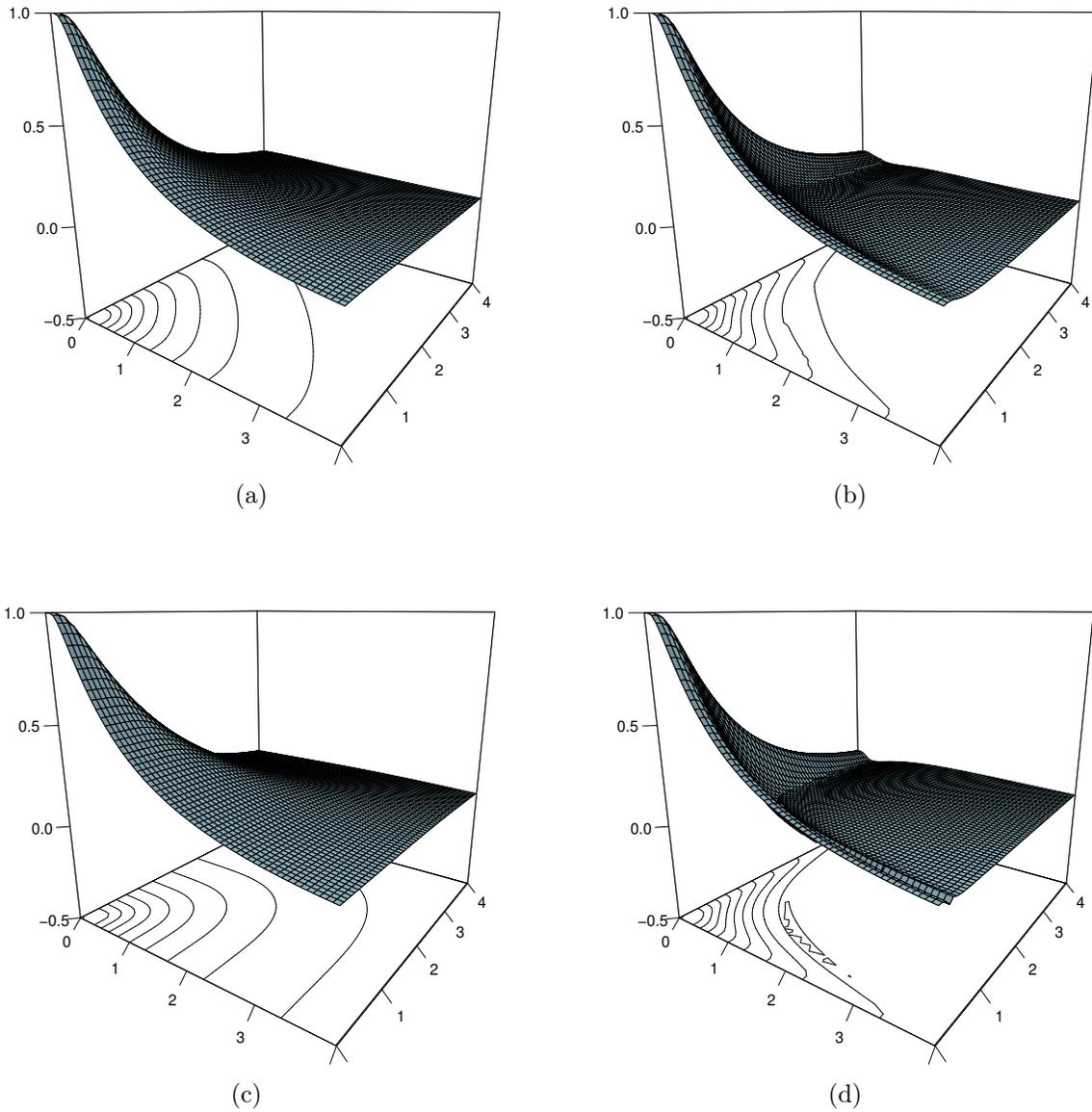


Figura 4.1: Função de sobrevivência bivariada conjunta de  $(T_1, T_2)$  quando  $\alpha_1 = \alpha_2 = 1$ ,  $\beta_1 = \beta_2 = 1$  para (a)  $\rho = 0.5$ , (b)  $\rho = -0.5$ , (c)  $\rho = 0.9$ , (d)  $\rho = -0.9$ .

**Observação 4.3.1.** Nas Figuras 4.1 e 4.2 foram obtidas usando a aproximação para  $\Phi_2(\cdot; \rho)$  proposta por Tsay & Ke (2009) e assim poder refinar os resultados de  $S_{\mathbf{T}}(\mathbf{t})$  e  $h_{\mathbf{T}}(\mathbf{t})$ , ou seja

$$S_{\mathbf{T}}(\mathbf{t}) = 1 - \Phi(a(t_1; \alpha_1, \beta_1)) - \Phi(a(t_2; \alpha_2, \beta_2)) + F_{app}(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho), \quad (4.3.4)$$

em que  $F_{app}$  é como no Teorema (4.2.1).

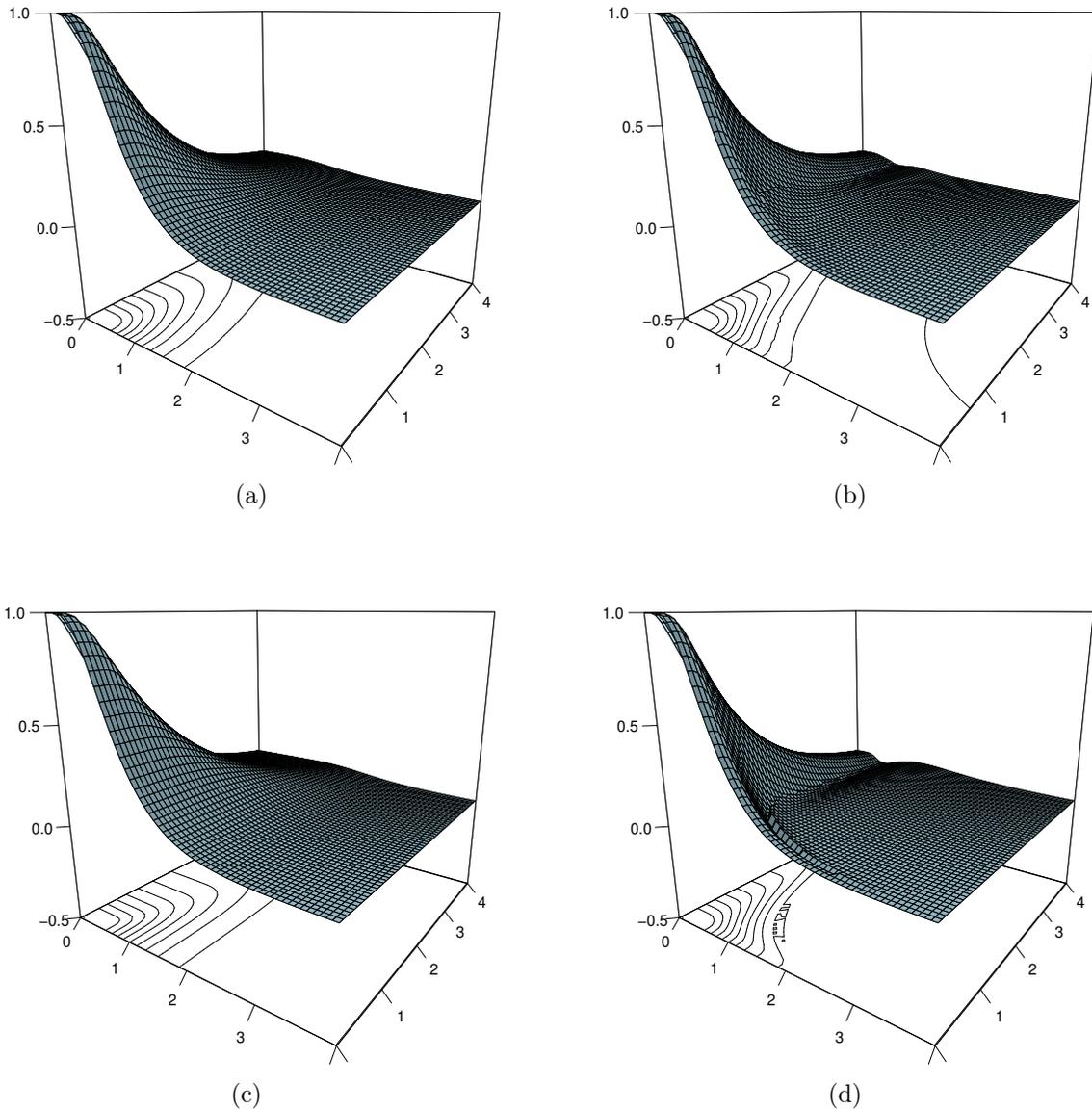


Figura 4.2: Função de sobrevivência bivariada conjunta de  $(T_1, T_2)$  quando  $\alpha_1 = 0.5$ ,  $\alpha_2 = 1$ ,  $\beta_2 = \beta_2 = 1$  para (a)  $\rho = 0.5$ , (b)  $\rho = -0.5$ , (c)  $\rho = 0.9$ , (d)  $\rho = -0.9$ .

### 4.3.2 Função de taxa de falha bivariada

Pode-se definir a função de taxa de falha, também definida como função de risco, de  $\mathbf{T}$  em mais de uma maneira no conjunto bivariado, por exemplo, seguindo as ideias de Basu (1971) consideramos a função de risco bivariada como uma quantidade escalar e é definida por:

**Definição 4.3.2.** Dada uma distribuição bivariada absolutamente contínua da função  $F(t_1, t_2)$

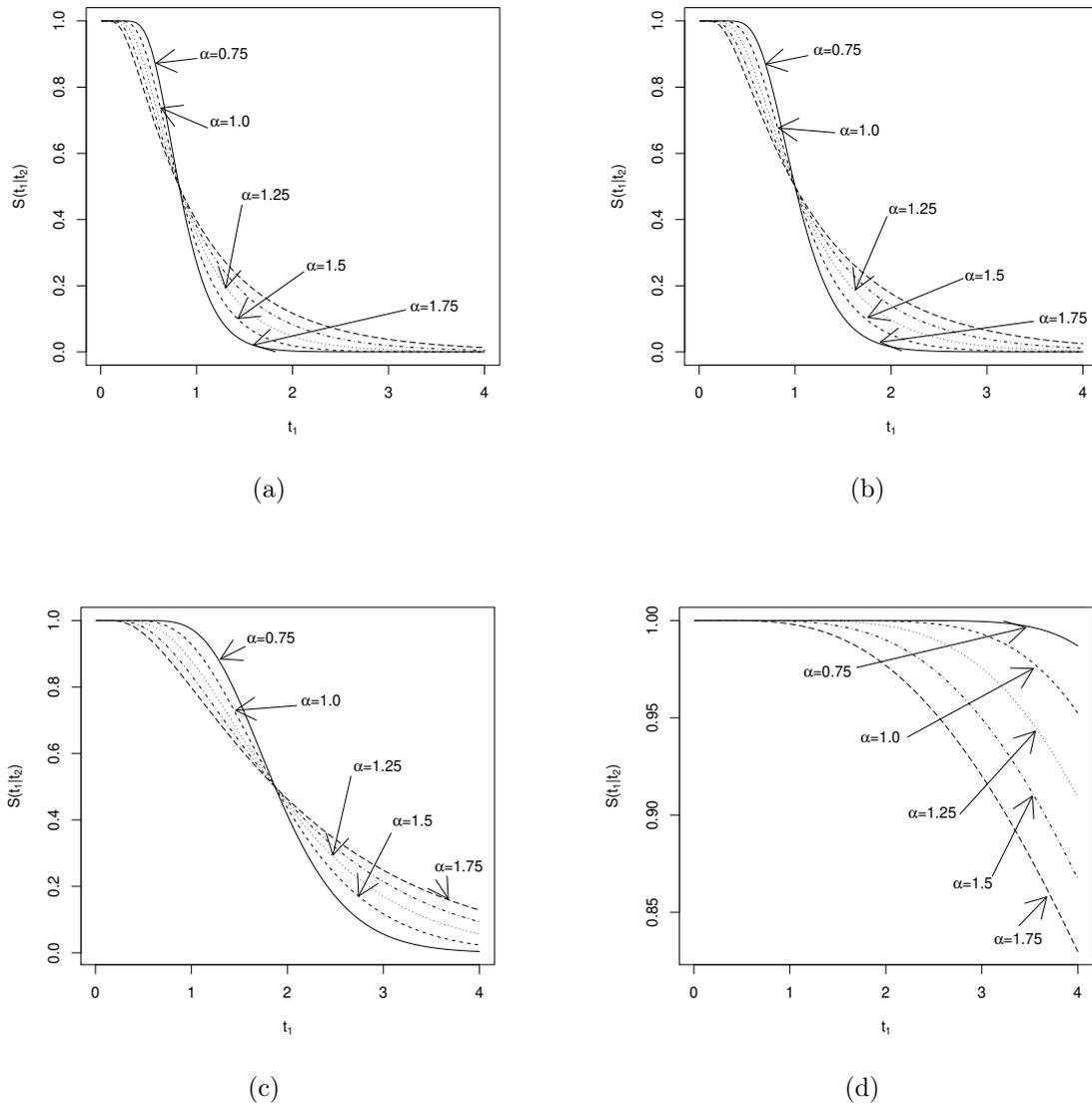


Figura 4.3: (a) Função de sobrevivência condicional  $T_1$ , dado  $T_2 = t_2 = 1$ , quando  $\alpha_1 = \alpha_2 = 1$ ,  $\beta_1 = \beta_2 = 2$  e (a)  $\rho = 0.9$ , (b)  $\rho = 0.5$ , (c)  $\rho = 0$ , (d)  $\rho = -0.9$ .

com função da densidade  $f(t_1, t_2)$ , a taxa de falha bivariada em  $(t_1, t_2)$  é dada por

$$h_{\mathbf{T}}(\mathbf{t}) = \frac{f_{\mathbf{T}}(t_1, t_2)}{P(T_1 > t_1, T_2 > t_2)} = \frac{f_{\mathbf{T}}(\mathbf{t})}{S_{\mathbf{T}}(\mathbf{t})} \quad \mathbf{t} \in \mathbb{R}_+^2. \quad (4.3.5)$$

**Observação 4.3.3.** Note-se que no caso de independência ( $\rho = 0$ ), temos

$$\begin{aligned}
 h_{\mathbf{T}}(\mathbf{t}) &= \frac{f(t_1, t_2)}{P(T_1 > t_1) \cdot P(T_2 > t_2)} \\
 &= \frac{f(t_1)}{1 - F(t_1)} \frac{f(t_2)}{1 - F(t_2)} \\
 &= h_{T_1}(t_1) \cdot h_{T_2}(t_2)
 \end{aligned}$$

em que  $h_{T_1}(t_1)$  e  $h_{T_2}(t_2)$  são as taxas de falhas univariadas correspondentes com  $T_1$  e  $T_2$  respectivamente.

De (4.3.8),  $h_{\mathbf{T}}(\mathbf{t})$ , temos que a função de taxa de falha para o modelo BS bivariado é

$$h_{\mathbf{T}}(\mathbf{t}) = \frac{\phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}))A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta})}{S_{\mathbf{T}}(\mathbf{t})}, \quad \mathbf{t} \in \mathbb{R}_+^2. \quad (4.3.6)$$

**Observação 4.3.4.** Note-se que de (4.3.6) pode ser escrito como

$$h_{\mathbf{T}}(\mathbf{t}) = \frac{\phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}))A(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta})}{1 - \Phi(a(t_1; \alpha_1, \beta_1)) - \Phi(a(t_2; \alpha_2, \beta_2)) + \Phi_2(a(\mathbf{t}; \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho)}, \quad \mathbf{t} \in \mathbb{R}_+^2.$$

A função taxa de falha condicional, veja as Figura 4.6 e 4.7, é dada por

$$h_{T_1|T_2}(t_1|t_2) = \frac{f_{T_1|T_2}(t_1|t_2)}{S_{T_1|T_2}(t_1|t_2)} = \frac{f_{T_1|T_2}(t_1|t_2)}{1 - F_{T_1|T_2}(t_1|t_2)}, \quad (4.3.7)$$

em que  $f_{T_1|T_2}(t_1|t_2) = \phi_1(a(t_1; \alpha_{1\rho}, \beta_1) - \mu_1(t_2))A(t_1; \alpha_{1\rho}, \beta_1)$ , com  $\alpha_{1\rho} = \alpha_1 \sqrt{1 - \rho^2}$  e  $\phi(\cdot)$  é a fdp univariada de uma normal padrão.

Mas, a principal desvantagem da função taxa de falha, definida em (4.3.8), é que não é possível determinar uma fdp em forma única. Mais tarde, Johnson & Kotz (1975) introduziu o conceito de uma taxa de risco vetorial multivariada e Marshall (1975) mostrou que esta função vetorial determina a distribuição de falha multivariada de forma única e é uma generalização do conceito univariado correspondente.

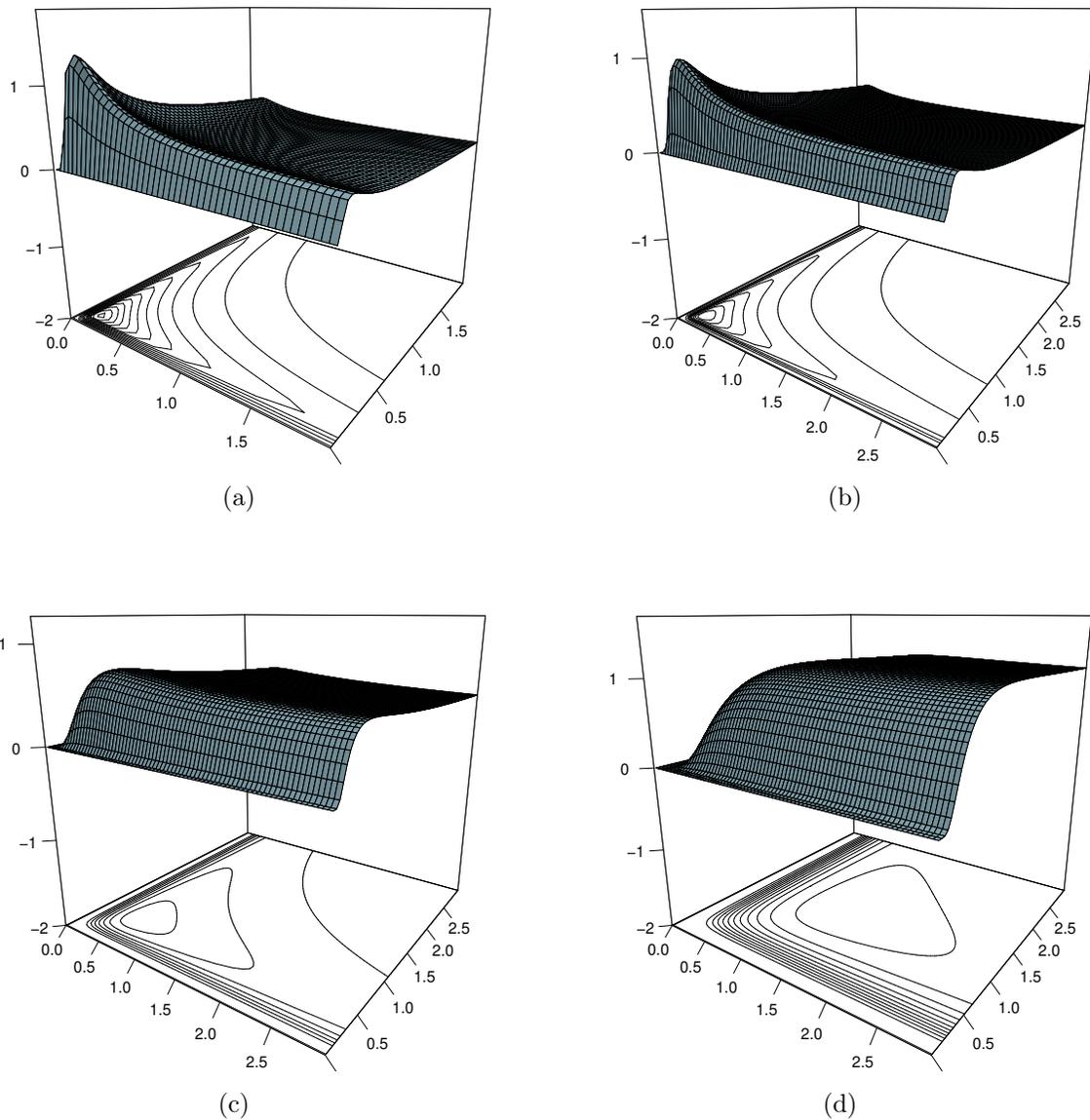


Figura 4.4: Função de taxa de falha bivariada quando  $\beta_1 = \beta_2 = 1$ ,  $\rho = 0$  e (a)  $\alpha_1 = \alpha_2 = 1.75$ , (b)  $\alpha_1 = \alpha_2 = 1.5$ , (c)  $\alpha_1 = \alpha_2 = 1$ , (d)  $\alpha_1 = \alpha_2 = 0.75$ .

Seguindo as ideias de Johnson & Kotz (1975) definimos a função de risco bivariada como um vetor (ou gradiente) e é dado por:

**Definição 4.3.5.** *Suponha que  $\mathbf{T} = (T_1, T_2)$  seja um vetor aleatório absolutamente contínuo com função de sobrevivência conjunta  $S_{\mathbf{T}}(\mathbf{t})$ . Sua função taxa de risco bivariada conjunta é definida como*

$$h_{\mathbf{T}}(\mathbf{t}) = (h_1(\mathbf{t}), h_2(\mathbf{t})), \tag{4.3.8}$$

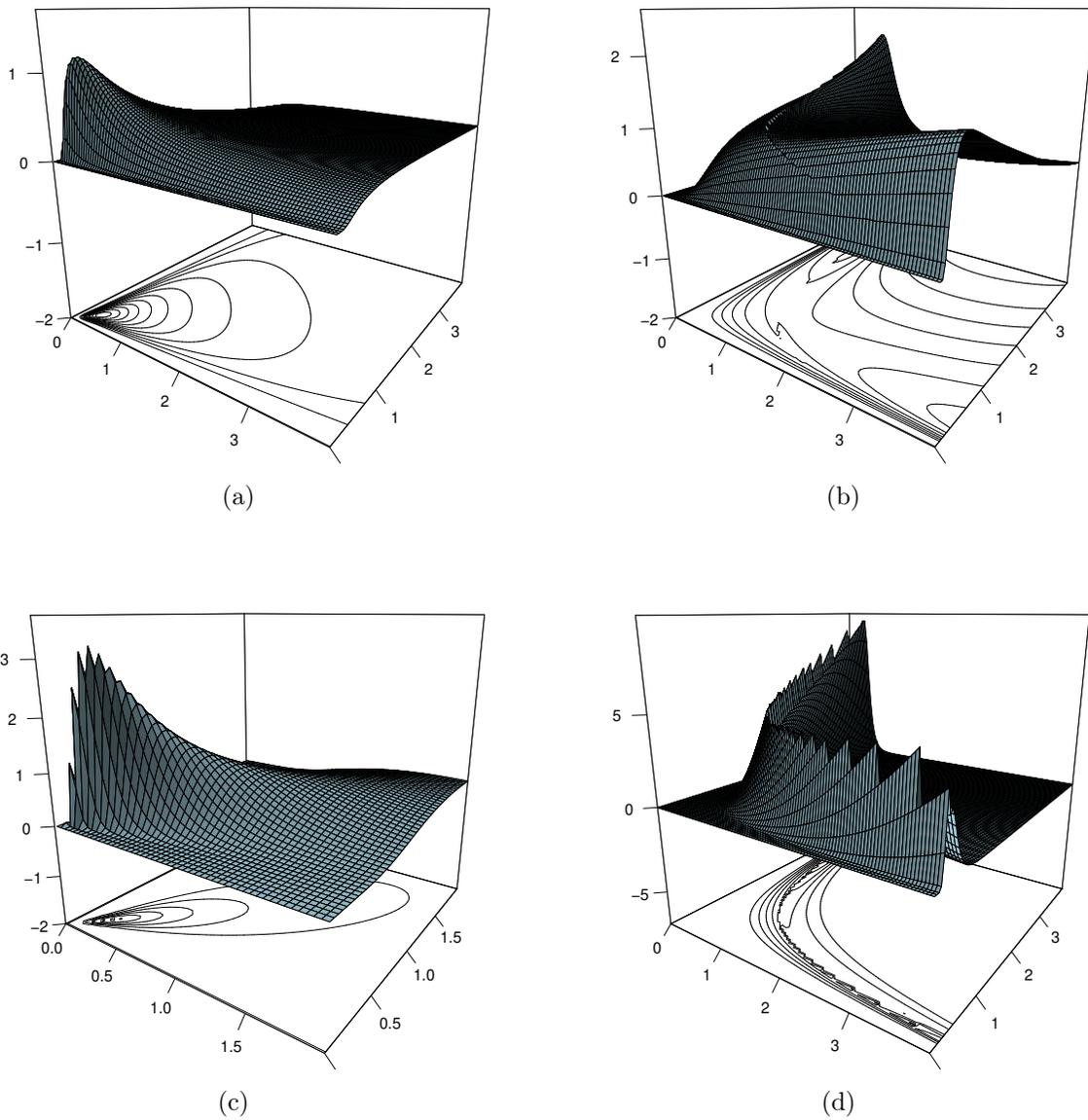


Figura 4.5: Função de taxa de falha bivariada conjunta de  $(T_1, T_2)$  quando  $\alpha_1 = \alpha_2 = 1, \beta_2 = \beta_2 = 1$  para (a)  $\rho = 0.5$ , (b)  $\rho = -0.5$ , (c)  $\rho = 0.9$ , (d)  $\rho = -0.9$ .

em que  $h_i(\mathbf{t}) = -\frac{\partial}{\partial t_i} \log S_{\mathbf{T}}(\mathbf{t}), \quad i = 1, 2.$

Deste modo, algumas caracterizações recentes utilizando o gradiente de risco foram dadas por Navarro e Ruiz (2004), Kotz et al. (2007) e Navarro et ai. (2007).

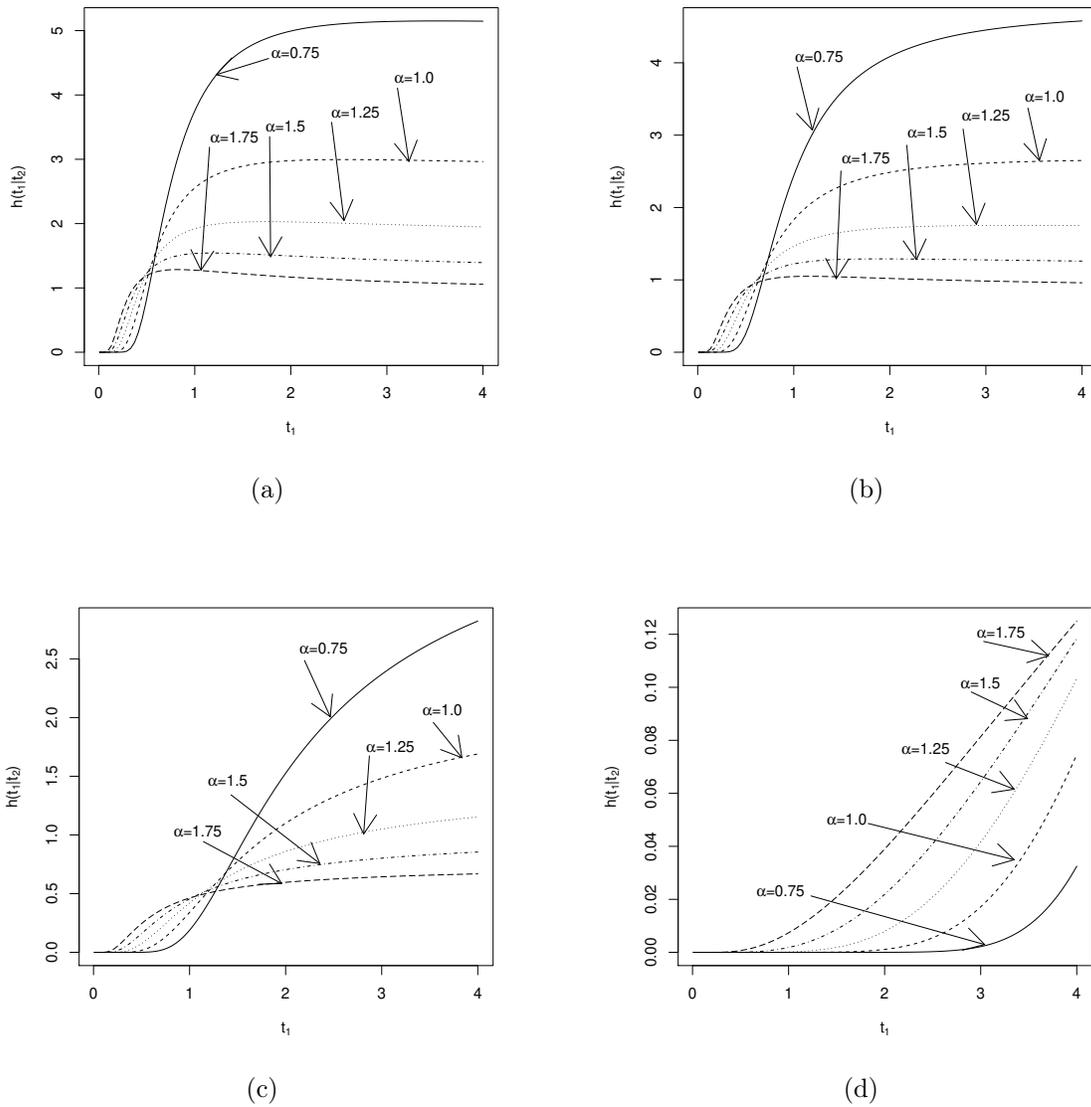


Figura 4.6: Função de taxa de falha condicional de  $T_1$ , dado  $T_2 = t_2 = 1$ , quando  $\rho = 0.9$  com (a)  $t_2 = 0.8$ , (b)  $t_2 = 1$ , (c)  $t_2 = 2$ , (d)  $t_2 = 8$ .

Da definição (4.3.5) temos que o seguinte Teorema:

**Teorema 4.3.6.** *Seja  $\mathbf{T} \sim \text{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ . Então,*

$$(i) \quad h_1(\mathbf{t}) = \frac{S_{T_2|T_1}(t_2|t_1)f_{T_1}(t_1)}{S_{\mathbf{T}}(\mathbf{t})},$$

$$(ii) \quad h_2(\mathbf{t}) = \frac{S_{T_1|T_2}(t_1|t_2)f_{T_2}(t_2)}{S_{\mathbf{T}}(\mathbf{t})}.$$

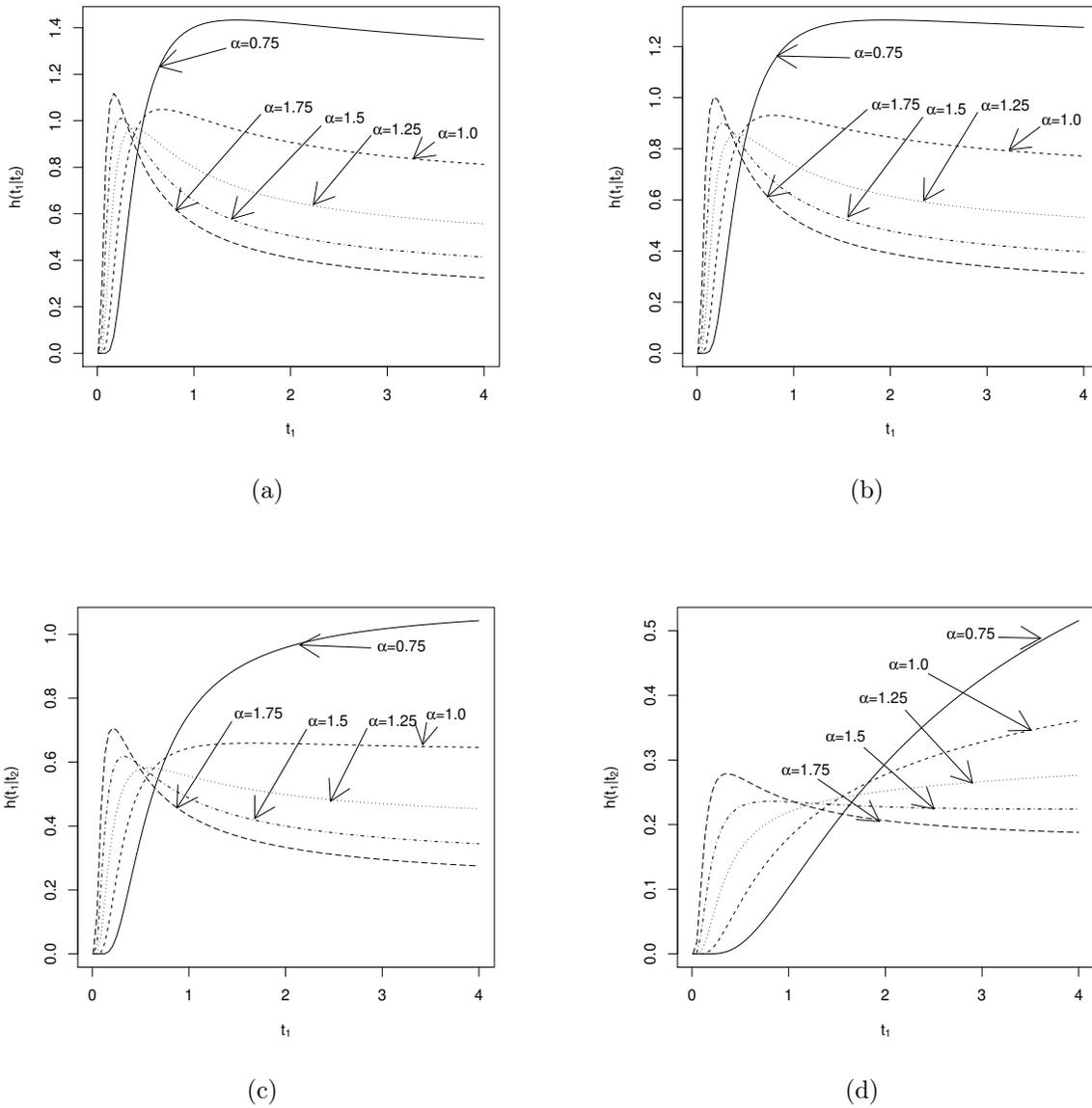


Figura 4.7: Função de taxa de falha condicional de  $T_1$ , dado  $T_2$ , quando  $\rho = 0.5$  com (a)  $t_2 = 0.8$ , (b)  $t_2 = 1$ , (c)  $t_2 = 2$ , (d)  $t_2 = 8$ .

em que

$$\begin{aligned}
 S_{T_1|T_2}(t_1 | T_2 = t_2) &= P(T_1 \geq t_1 | T_2 = t_2) = 1 - \Phi[a_{t_1}(\alpha_{1\rho}, \beta_1) - \mu_1(t_2)], \\
 S_{T_2|T_1}(t_2 | T_1 = t_1) &= P(T_2 \geq t_2 | T_1 = t_1) = 1 - \Phi[a_{t_2}(\alpha_{2\rho}, \beta_2) - \mu_2(t_1)],
 \end{aligned}$$

em que  $\mu_1(t_2)$  e  $\mu_2(t_1)$  são como em (2.2.4). Veja a Figura 4.3.

**Demonstração:**

(i) Seja  $\mathbf{T} \sim \text{BS}_2(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$ . Então

$$\begin{aligned} h_1(\mathbf{t}) &= -\frac{\partial}{\partial t_i} \log S_{\mathbf{T}}(\mathbf{t}) \\ &= -\frac{1}{S_{\mathbf{T}}(\mathbf{t})} \frac{\partial S_{\mathbf{T}}(\mathbf{t})}{\partial t_1}, \end{aligned} \quad (4.3.9)$$

em que

$$\begin{aligned} \frac{\partial S_{\mathbf{T}}(\mathbf{t})}{\partial t_1} &= \frac{\partial [1 - \Phi(a(t_1, \alpha_1, \beta_1)) - \Phi(a(t_2, \alpha_2, \beta_2)) + \Phi_2(a(\mathbf{t}, \boldsymbol{\alpha}, \boldsymbol{\beta}); \rho)]}{\partial t_1} \quad (4.3.10) \\ &= -\phi(a(t_1, \alpha_1, \beta_1)) \frac{\partial a(t_1, \alpha_1, \beta_1)}{\partial t_1} + \frac{\partial \Phi(a(t_2, \alpha_2, \beta_2))}{\partial t_1} \\ &= -\phi(a(t_1, \alpha_1, \beta_1)) A(t_1, \alpha_1, \beta_1) + \frac{\partial \Phi(a(t_2, \alpha_2, \beta_2))}{\partial t_1}, \end{aligned}$$

com

$$\begin{aligned} \frac{\partial \Phi(a(t_2, \alpha_2, \beta_2))}{t_1} &= \frac{\partial}{\partial t_1} \int_{-\infty}^{a(t_1, \alpha_1, \beta_1)} \int_{-\infty}^{a(t_2, \alpha_2, \beta_2)} \phi(z_1, z_2; \rho) dz_2 dz_1 \quad (4.3.11) \\ &= \frac{\partial}{\partial t_1} \int_{-\infty}^{a(t_1, \alpha_1, \beta_1)} \int_{-\infty}^{a(t_2, \alpha_2, \beta_2)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_1^2\right) \frac{1}{\sqrt{2\pi} \sqrt{1-\rho^2}} \\ &= \exp\left(-\frac{1}{2\sqrt{1-\rho^2}} (z_2^2 - \rho z_1)^2\right) dz_2 dz_1 \\ &= \frac{\partial}{\partial t_1} \int_{-\infty}^{a(t_1, \alpha_1, \beta_1)} \phi(z_1) \int_{-\infty}^{a(t_2, \alpha_2, \beta_2)} \phi(z_2, \rho z_1; 1-\rho^2) dz_2 dz_1 \\ &= \frac{\partial}{\partial t_1} \int_{-\infty}^{a(t_1, \alpha_1, \beta_1)} \phi(z_1) H(z_1, a(t_2, \alpha_2, \beta_2)) dz_1 \\ &= \phi(a(t_1, \alpha_1, \beta_1)) H(z_1, a(t_2, \alpha_2, \beta_2)) \frac{\partial a(t_1, \alpha_1, \beta_1)}{\partial t_1} \\ &= \phi(a(t_1, \alpha_1, \beta_1)) H(z_1, a(t_2, \alpha_2, \beta_2)) A(t_1, \alpha_1, \beta_1) \\ &= f_{T_1}(t_1) H(z_1, a(t_2, \alpha_2, \beta_2)), \end{aligned}$$

e

$$\begin{aligned}
 H(z_1, a(t_2, \alpha_2, \beta_2)) &= \int_{-\infty}^{a(t_2, \alpha_2, \beta_2)} \phi(z_2, \rho z_1; 1 - \rho^2) dz_2 & (4.3.12) \\
 &= P(z_2 \leq a(t_2, \alpha_2, \beta_2)) \\
 &= P\left(\frac{z_2 - \rho z_1}{\sqrt{1 - \rho^2}} \leq \frac{a(t_2, \alpha_2, \beta_2) - \rho a(t_1, \alpha_1, \beta_1)}{\sqrt{1 - \rho^2}}\right) \\
 &= \Phi\left(\frac{a(t_2, \alpha_2, \beta_2) - \rho a(t_1, \alpha_1, \beta_1)}{\sqrt{1 - \rho^2}}\right) \\
 &= \Phi(a(t_2, \alpha_{2\rho}, \beta_2) - \mu_2(t_1)),
 \end{aligned}$$

De (4.3.11), (4.3.12) e (4.3.13) em (4.3.9), temos que

$$h_1(\mathbf{t}) = \frac{S_{T_1|T_2}(t_1|t_2)f_{T_1}(t_1)}{S_{\mathbf{T}}(\mathbf{t})}. \quad \blacksquare$$

A demonstração de (ii) é em forma análoga.

## 4.4 Estudo de Simulação

A simulação foi baseada no trabalho de Johnson & Kotz (1973), o objetivo era poder ter uma ideia do valor a qual converge a primeira componente da função taxa de falha bivariada, para isto foi usado dois algoritmos diferentes, o proposto recentemente por Meyer (2013) e o trabalho não tão recente de Tsay & Ke (2009), para a implementação do algoritmo de Meyer (2013) foi desenvolvido em C++ e disponibilizado pelo autor em seu artigo, para o caso do algoritmo proposto por Tsay & Ke (2009) foi desenvolvido no R. Foram considerados os coeficientes de correlação de 0,2; 0,4; 0,6 e 0,8. Em relação aos parâmetros  $\alpha$  y  $\beta$ , foram fixados como por exemplo  $\alpha = (1, 1)$  e  $\beta = (1, 1)$ . Notamos que nas seguintes tabelas apresentadas, quando o  $\rho$  vai crescendo a primeira componente da função de taxa bivariada vai decrescendo mais rapidamente, isto também pode ser observado quando  $t_1$  ou  $t_2$  vai em aumento.

Tabela 4.1: Birnbaum Saunders: Primeiro componente da função taxa de falha, pela aproximação de Meyer (2013),

$$h_1(\mathbf{t}) = \frac{S_{T_2|T_1}(t_2|t_1)f_{T_1}(t_1)}{S_{\mathbf{T}}(\mathbf{t})} \quad \rho = 0.2$$

$$\boldsymbol{\alpha} = (1, 1) \quad , \quad \boldsymbol{\beta} = (1, 1)$$

$t_2 \backslash t_1$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5
0.5	0.2667	0.3211	0.3410	0.3539	0.3638	0.3720	0.3789	0.3847	0.3897	0.3940
1.0	0.2806	0.3462	0.3734	0.3926	0.4084	0.4223	0.4348	0.4462	0.4567	0.4664
1.5	0.3160	0.3931	0.4239	0.4448	0.4618	0.4765	0.4896	0.5016	0.5126	0.5227
2.0	0.3539	0.4492	0.4875	0.5127	0.5327	0.5496	0.5644	0.5776	0.5896	0.6005
2.5	0.3873	0.5037	0.5521	0.5840	0.6089	0.6297	0.6477	0.6636	0.6778	0.6906
3.0	0.4147	0.5526	0.6125	0.6524	0.6835	0.7095	0.7318	0.7515	0.7690	0.7848
3.5	0.4366	0.5952	0.6669	0.7154	0.7534	0.7852	0.8126	0.8368	0.8584	0.8777
4.0	0.4537	0.6315	0.7149	0.7720	0.8172	0.8551	0.8880	0.9171	0.9431	0.9666
4.5	0.4669	0.6622	0.7567	0.8224	0.8746	0.9187	0.9572	0.9913	1.022	1.050
5.0	0.4768	0.6880	0.7930	0.8668	0.9259	0.9761	1.020	1.059	1.095	1.127

Tabela 4.2:  $\rho = 0.4$

$t_2 \backslash t_1$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5
0.5	0.2824	0.3163	0.3133	0.3043	0.2935	0.2822	0.2707	0.2591	0.2477	0.2366
1.0	0.2785	0.3308	0.3449	0.3522	0.3569	0.3598	0.3614	0.3617	0.3610	0.3594
1.5	0.2939	0.3628	0.3793	0.3868	0.3918	0.3956	0.3987	0.4013	0.4033	0.4049
2.0	0.3043	0.4001	0.4256	0.4360	0.4414	0.4447	0.4469	0.4485	0.4497	0.4507
2.5	0.3066	0.4316	0.4715	0.4893	0.4984	0.5034	0.5061	0.5074	0.5079	0.5078
3.0	0.3029	0.4546	0.5113	0.5395	0.5554	0.5647	0.5701	0.5729	0.5740	0.5740
3.5	0.2950	0.4696	0.5433	0.5835	0.6080	0.6237	0.6337	0.6399	0.6433	0.6448
4.0	0.2846	0.4777	0.5675	0.6201	0.6544	0.6777	0.6938	0.7047	0.7119	0.7162
4.5	0.2728	0.4805	0.5849	0.6495	0.6937	0.7253	0.7483	0.7651	0.7771	0.7854
5.0	0.2604	0.4790	0.5962	0.6722	0.7262	0.7662	0.7966	0.8198	0.8373	0.8505

Tabela 4.3:  $\rho = 0.6$

$t_2 \backslash t_1$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5
0.5	0.3162	0.3228	0.2810	0.2375	0.1989	0.1661	0.1385	0.1155	0.0963	0.0803
1.0	0.2889	0.3393	0.3418	0.3358	0.3249	0.3105	0.2936	0.2750	0.2557	0.2361
1.5	0.2685	0.3586	0.3670	0.3647	0.3606	0.3555	0.3492	0.3414	0.3320	0.3211
2.0	0.2375	0.3752	0.4018	0.4030	0.3982	0.3925	0.3869	0.3813	0.3755	0.3692
2.5	0.2035	0.3777	0.4316	0.4450	0.4440	0.4380	0.4308	0.4235	0.4166	0.4101
3.0	0.1714	0.3676	0.4496	0.4812	0.4901	0.4884	0.4819	0.4734	0.4646	0.4561
3.5	0.1433	0.3487	0.4552	0.5072	0.5302	0.5372	0.5352	0.5285	0.5193	0.5093
4.0	0.1192	0.3250	0.4504	0.5220	0.5610	0.5798	0.5858	0.5839	0.5771	0.5676
4.5	0.0991	0.2990	0.4378	0.5265	0.5815	0.6137	0.6301	0.6357	0.6340	0.6274
5.0	0.0824	0.2726	0.4197	0.5225	0.5924	0.6381	0.6662	0.6812	0.6867	0.6853

Tabela 4.4:  $\rho = 0.8$

$t_2 \backslash t_1$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5
0.5	0.4061	0.3288	0.1864	0.0983	0.0514	0.0270	0.0144	0.0077	0.0042	0.0023
1.0	0.3056	0.4010	0.3812	0.3307	0.2653	0.2011	0.1466	0.1040	0.0724	0.0497
1.5	0.1830	0.3972	0.4101	0.3944	0.3701	0.3349	0.2913	0.2442	0.1985	0.1572
2.0	0.0983	0.3558	0.4324	0.4304	0.4142	0.3957	0.3728	0.3435	0.3085	0.2700
2.5	0.0516	0.2863	0.4276	0.4637	0.4566	0.4393	0.4214	0.4025	0.3804	0.3538
3.0	0.0271	0.2147	0.3913	0.4762	0.4961	0.4863	0.4679	0.4489	0.4306	0.4115
3.5	0.0144	0.1543	0.3362	0.4609	0.5181	0.5297	0.5183	0.4988	0.4785	0.4593
4.0	0.0077	0.1080	0.2759	0.4233	0.5162	0.5580	0.5645	0.5518	0.5314	0.5096
4.5	0.0042	0.0744	0.2191	0.3727	0.4921	0.5651	0.5971	0.6003	0.5866	0.5653
5.0	0.0023	0.0507	0.1698	0.3177	0.4520	0.5512	0.6107	0.6361	0.6369	0.6224

Tabela 4.5: Birnbaum Saunders: Primeiro componente da função taxa de falha, pela aproximação de Tsay & Ke (2009),

$$h_1(\mathbf{t}) = \frac{S_{T_2|T_1}(t_2|t_1)f_{T_1}(t_1)}{S_{\mathbf{T}}(\mathbf{t})} \quad \rho = 0.2$$

$$\boldsymbol{\alpha} = (1, 1) \quad , \quad \boldsymbol{\beta} = (1, 1)$$

$t_2 \backslash t_1$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5
0.5	0.3729	0.3395	0.3168	0.2994	0.2851	0.2729	0.2622	0.2527	0.2441	0.2364
1.0	0.7538	0.7082	0.6755	0.6494	0.6274	0.6082	0.5910	0.5754	0.5611	0.5478
1.5	1.0280	0.9776	0.9415	0.9131	0.8894	0.8693	0.8519	0.8368	0.8239	0.8131
2.0	1.2484	1.1938	1.1532	1.1201	1.0916	1.0660	1.0424	1.0201	0.9987	0.9776
2.5	1.4371	1.3802	1.3364	1.2997	1.2670	1.2366	1.2074	1.1785	1.1491	1.1185
3.0	1.6049	1.5477	1.5028	1.4647	1.4303	1.3979	1.3664	1.3349	1.3024	1.2683
3.5	1.7577	1.7023	1.6586	1.6216	1.5885	1.5578	1.5283	1.4993	1.4701	1.4401
4.0	1.8991	1.8476	1.8074	1.7743	1.7457	1.7205	1.6980	1.6778	1.6597	1.6436
4.5	2.0317	1.9858	1.9515	1.9250	1.9045	1.8894	1.8796	1.8756	1.8786	1.8902
5.0	2.1570	2.1186	2.0924	2.0755	2.0668	2.0667	2.0764	2.0983	2.1361	2.1959

Tabela 4.6:  $\rho = 0.4$

$t_2 \backslash t_1$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5
0.5	0.3392	0.2734	0.2312	0.2006	0.1768	0.1575	0.1414	0.1277	0.1158	0.1055
1.0	0.7229	0.6328	0.5675	0.5165	0.4746	0.4389	0.4078	0.3803	0.3557	0.3335
1.5	1.0040	0.9073	0.8343	0.7757	0.7262	0.6830	0.6445	0.6098	0.5782	0.5493
2.0	1.2313	1.1328	1.0550	0.9912	0.9367	0.8888	0.8459	0.8068	0.7707	0.7369
2.5	1.4260	1.3288	1.2487	1.1815	1.1231	1.0713	1.0243	0.9812	0.9410	0.9032
3.0	1.5985	1.5048	1.4247	1.3560	1.2956	1.2414	1.1919	1.1462	1.1034	1.0630
3.5	1.7548	1.6662	1.5879	1.5197	1.4591	1.4044	1.3544	1.3082	1.2652	1.2248
4.0	1.8987	1.8163	1.7414	1.6753	1.6163	1.5631	1.5146	1.4703	1.4297	1.3924
4.5	2.0326	1.9572	1.8870	1.8246	1.7689	1.7190	1.6742	1.6341	1.5986	1.5679
5.0	2.1582	2.0905	2.0262	1.9688	1.9179	1.8730	1.8339	1.8004	1.7732	1.7527

Tabela 4.7:  $\rho = 0.6$

$t_2 \backslash t_1$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5
0.5	0.3062	0.2027	0.1437	0.1061	0.0804	0.0619	0.0483	0.0381	0.0302	0.0241
1.0	0.7088	0.5662	0.4612	0.3824	0.3208	0.2714	0.2309	0.1974	0.1692	0.1455
1.5	1.0057	0.8629	0.7444	0.6482	0.5679	0.5000	0.4419	0.3916	0.3477	0.3092
2.0	1.2432	1.1099	0.9882	0.8845	0.7955	0.7178	0.6490	0.5876	0.5328	0.4836
2.5	1.4439	1.3237	1.2039	1.0972	1.0031	0.9195	0.8447	0.7771	0.7154	0.6588
3.0	1.6196	1.5135	1.3989	1.2924	1.1959	1.1086	1.0295	0.9574	0.8914	0.8307
3.5	1.7771	1.6848	1.5775	1.4736	1.3770	1.2881	1.2064	1.1312	1.0620	0.9981
4.0	1.9208	1.8415	1.7427	1.6433	1.5484	1.4596	1.3769	1.3002	1.2290	1.1632
4.5	2.0536	1.9862	1.8965	1.8028	1.7113	1.6241	1.5419	1.4650	1.3933	1.3267
5.0	2.1774	2.1208	2.0405	1.9535	1.8665	1.7822	1.7019	1.6261	1.5551	1.4890

Tabela 4.8:  $\rho = 0.8$

$t_2 \backslash t_1$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5
0.5	0.2705	0.1110	0.0481	0.0221	0.0105	0.0052	0.0026	0.0013	0.0007	0.0004
1.0	0.7262	0.5018	0.3280	0.2129	0.1380	0.0894	0.0579	0.0374	0.0242	0.0156
1.5	1.0417	0.8597	0.6612	0.4967	0.3689	0.2719	0.1991	0.1449	0.1048	0.0754
2.0	1.2807	1.1501	0.9648	0.7870	0.6323	0.5024	0.3959	0.3097	0.2406	0.1857
2.5	1.4770	1.3881	1.2300	1.0571	0.8936	0.7469	0.6183	0.5079	0.4143	0.3358
3.0	1.6469	1.5882	1.4604	1.3025	1.1408	0.9876	0.8475	0.7215	0.6098	0.5121
3.5	1.7990	1.7609	1.6615	1.5235	1.3709	1.2183	1.0730	0.9383	0.8152	0.7037
4.0	1.9379	1.9136	1.8385	1.7220	1.5831	1.4364	1.2911	1.1522	1.0221	0.9019
4.5	2.0667	2.0514	1.9960	1.9003	1.7775	1.6407	1.4996	1.3603	1.2267	1.1008
5.0	2.1874	2.1779	2.1378	2.0612	1.9553	1.8309	1.6972	1.5610	1.4270	1.2980

# Capítulo 5

## Considerações Finais

### 5.1 Conclusões

Neste trabalho, o objetivo principal foi considerar um estudo do modelo  $\mathcal{BS}$  bivariada, apresentado por Kundu *et al.* (2010).

No Capítulo 2 foi determinado a matriz de informação de Fisher esperada, assim como a forma explícita de seu inversa, assim como um estudo de inferência baseados considerando quatro hipótese  $H_{0k}$ ,  $k = 1, 2, 3, 4$ , os testes são baseados nas estatísticas de Wald, Escore e Razão Verossimilhanças, respectivamente no modelo  $\mathcal{BS}$  bivariado. Mostramos através de simulação que o teste de escore tem melhor comportamento em relação aos outros dois teste, em termos de nível de significância, quando consideramos o modelo  $\mathcal{BS}$  bivariado. Os estimadores de máxima de verossimilhança dos parâmetros de escala do modelo  $\mathcal{BS}$  bivariado, sob as hipótese  $H_{0k}$ ,  $k = 1, 2, 3, 4$ . Nos casos em que os EMV não tem uma forma fechada, sua obtenção foi por métodos iterativos.

Foi determinado que estatística de Wald sob a hipótese  $H_{03}$  pode ser escrita como:  $W_{03} = W_{01} + W_{02}$ . Foi calculado a distribuição assintótica dos estimadores de momentos modificados. Foi demonstrado que os estimadores do parâmetros de forma e correlação é consistente e assintoticamente normal quando supunhamos que os estimadores do parâmetro de escala é consistente. Foram realizados estudos de simulação utilizando os programas estatísticos como o R e Matlab.

No Capítulo 3, conhecendo a importância de avaliar a presença das observações influentes, discutimos um procedimento de diagnóstico, baseado no método de influência local proposto por Cook (1986). Um esquema de perturbação de casos foi considerado na análise. Como observamos, a metodologia prevê um estudo unificado para avaliar a influência local e pode ser aplicado a modelos mais gerais. Os resultados obtidos foram comparados com outras medidas de diagnóstico usadas na literatura, com por exemplo, o afastamento pela verossimilhança  $LD_i$  e a distância de Cook generalizada  $L_i$ , notando que estas medidas prevê quase as mesmas observações influentes que as obtidas pelas metodologias consideradas ( $C_i; r = 0$ ) para a perturbações de ponderação de casos, são que estas possuem um critério mais objetivo para avaliar influências de observações, além de outras propriedades importantes apresentadas nesse capítulo.

Os resultados obtido foram aplicado em conjunto de dados reais, o conjunto de dados trata sobre a densidade dos ossos que foi considerado por Kundu *et al.* (2010) baseado no modelo  $\mathcal{BS}$  bivariado. Foram realizados estudos de simulação utilizando os programas estatísticos como o R e C++.

No capítulo 4, discutimos a função de sobrevivência e taxa de falha para o modelo  $\mathcal{BS}$  bivariado, em que foi apresentado a forma as componentes da função taxa de falha, proposto por Johnson & Kotz (1975). Foi apresentado diferentes algoritmos de aproximação para calcular a fda bivarida da normal padrão. Foram realizados estudos de simulação, utilizando o linguagem de programação, tais como R e C++.

# Apêndice A

## Código R da aproximação feita por Tsay Ke (2009)

```
Fapp=function(P,Q,rho){
  library(VGAM)
  #P=0
  #Q=-0.5
  #rho=0.8

  a=-rho/(1-rho^2)^0.5
  b=P/(1-rho^2)^0.5

  c1=-1.0950081470333
  c2=-0.75651138383854

  if(a>0 & (a*Q + b)>=0)
    Fapp=(1/2)*(erf(Q/sqrt(2)) + erf(b/(sqrt(2)*a))) + (1/(4*sqrt(1-
a^2*c2)))*exp((a^2*c1^2 - 2*sqrt(2)*b*c1 + 2*b^2*c2)/(4*(1-
a^2*c2)))*(1-erf((sqrt(2)*b-a^2*c1)/(2*a*(1-a^2*c2))))-(1/(4*sqrt(1-
a^2*c2)))*exp((a^2*c1^2 + 2*sqrt(2)*b*c1 + 2*b^2*c2)/(4*(1-
a^2*c2)))*(erf((sqrt(2)*Q-sqrt(2)*a^2*c2*Q-sqrt(2)*a*b*c2-
a*c1)/(2*sqrt(1-a^2*c2)))+erf((a^2*c1 + sqrt(2)*b)/(2*a*sqrt(1-a^2*c2))))

  if(a==0)
    Fapp=pnorm(P)*pnorm(Q)

  if(a>0 & (a*Q + b)<0)
    Fapp=(1/(4*sqrt(1-a^2*c2)))*exp((a^2*c1^2 - 2*sqrt(2)*b*c1 +
2*b^2*c2)/(4*(1-a^2*c2)))*(1+erf((sqrt(2)*Q-sqrt(2)*a^2*c2*Q-
sqrt(2)*a*b*c2+a*c1)/(2*sqrt(1-a^2*c2))))
```

```

if(a<0 & (a*Q + b)>=0)
  Fapp=1/2 + (1/2)*erf(Q/sqrt(2)) - (1/(4*sqrt(1-a^2*c2)))*exp((a^2*c1^2
+ 2*sqrt(2)*b*c1 + 2*b^2*c2)/(4*(1-a^2*c2)))*(1+erf((sqrt(2)*Q-
sqrt(2)*a^2*c2*Q-sqrt(2)*a*b*c2-a*c1)/(2*sqrt(1-a^2*c2))))

if(a<0 & (a*Q + b)<0)
  Fapp=1/2 - (1/2)*erf(b/(sqrt(2)*a)) - (1/(4*sqrt(1-
a^2*c2)))*exp((a^2*c1^2 + 2*sqrt(2)*b*c1 + 2*b^2*c2)/(4*(1-
a^2*c2)))*(1-erf((sqrt(2)*b+a^2*c1)/(2*a*sqrt(1-a^2*c2)))) +
(1/(4*sqrt(1-a^2*c2)))*exp((a^2*c1^2 - 2*sqrt(2)*b*c1 +
2*b^2*c2)/(4*(1-a^2*c2)))*(erf((sqrt(2)*Q-sqrt(2)*a^2*c2*Q-
sqrt(2)*a*b*c2+a*c1)/(2*sqrt(1-a^2*c2)))+erf((-a^2*c1 +
sqrt(2)*b)/(2*a*sqrt(1-a^2*c2))))

return(Fapp)
}

```

## Apêndice B

# Código C++, para obter as tabelas finais do Capítulo 4

Código da aproximação da fda da normal bivariada proposto por Tsay & Ke (2009).

```
#include <iostream>
#include <iomanip>
#include <cmath>

using namespace std;

double phi(double x)
{
    const double PI=3.141592653589793238463;

    return (1/sqrt(2*PI))*exp(-0.5*x*x);
}

double phi2(double x, double y, double rho)
{
    const double PI=3.141592653589793238463;

    return (1/(2*PI*sqrt(1-rho*rho)))*exp((-0.5/(1-rho*rho))*(x*x - 2*rho*x*y + y*y));
}

int main()
{
    int n=12;
    double h_t1t2_cond[n][n];
```

```

double s_t1t2_cond[n][n];
double ht1t2[n][n];
double hT1[n][n];
double St1t2[n][n];
double ft1t2[n][n];
double t1[12]={ 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6};
double t2[12]={ 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6};
double alpha1=1;
double alpha2=1;
double beta1=1;
double beta2=1;
double rho=0.8;
double at1rho[n];
double at2rho[n];
double At1rho[n];
double at1[n];
double at2[n];
double At1[n];
double At2[n];

for(int i=0; i<n; ++i)
{
    at1rho[i]=(rho/(alpha1*sqrt(1-rho*rho)))*(sqrt(t1[i]/beta1) -
sqrt(beta1/t1[i]));
    at2rho[i]=(1/(alpha2*sqrt(1-rho*rho)))*(sqrt(t2[i]/beta2) -
sqrt(beta2/t2[i]));
    At1rho[i]=pow(t1[i],(-3/2))*(t1[i] +
    beta1)/(2*sqrt(1-rho*rho)*alpha1*pow(beta1,(1/2)));

    at1[i]=(1/alpha1)*(sqrt(t1[i]/beta1) - sqrt(beta1/t1[i]));
    at2[i]=(1/alpha2)*(sqrt(t2[i]/beta2) - sqrt(beta2/t2[i]));
    At1[i]=pow(t1[i],(-3/2))*(t1[i] + beta1)/(2*alpha1*pow(beta1,(1/2)));
    At2[i]=pow(t2[i],(-3/2))*(t2[i] + beta2)/(2*alpha2*pow(beta2,(1/2)));

}

for(int i=0; i<n; ++i)
{
    for(int j=0; j<n; ++j)
        ft1t2[i][j]= phi2(at1[i],at2[j],rho)*At1[i]*At2[j];
}

for(int i=0; i<n; ++i)

```

```
{
    for(int j=0; j<n; ++j)
        St1t2[i][j]= 1 - Phi(at1[i])- Phi(at2[j]) + Phi2(at1[i],at2[j],rho);
}

for(int i=0; i<n; ++i)
{
    for(int j=0; j<n; ++j)
        hT1[i][j]=(1-phi(at2rho[i]-at1rho[j]))*ft1t2[i][j]/St1t2[i][j];
}

for(int i=0; i<n; ++i)
{
    for(int j=0; j<n; ++j)
        cout << setprecision(4) << showpoint << hT1[i][j] << ' ';
    cout<<endl;
}

return 0;
}
```



# Apêndice C

## Matriz de Informação de Fisher

Neste apêndice apresentamos a expressão desenvolvida da matriz de informação de Fisher observada assim como também o desenvolvimento parcial da prova do teorema (C.0.1) em que obtemos a matriz de informação de Fisher esperada e que é importante para a aplicação do estudo de diagnóstico. Alguns cálculos numéricos são apresentados como ilustração.

A função de verossimilhança foi definida em (C.0.1) e apresentamos novamente

$$\ell(\boldsymbol{\theta}) = -n \log(2\pi) - \frac{n}{2} \log(1 - \rho^2) + \sum_{i=1}^n \log(A(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta})) - \frac{1}{2} \sum_{i=1}^n d(\mathbf{t}_i).$$

A continuação temos as expressões da matriz, apresentanda em (??):

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \begin{pmatrix} \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^\top} & \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha} \partial \rho} \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\alpha}^\top} & \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \rho} \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \rho \partial \boldsymbol{\alpha}^\top} & \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \rho \partial \boldsymbol{\beta}^\top} & \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \rho \partial \rho} \end{pmatrix},$$

em que

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^\top} = \begin{pmatrix} \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_1 \partial \alpha_1} & \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_1 \partial \alpha_2} \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_1 \partial \alpha_2} & \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_2 \partial \alpha_2} \end{pmatrix},$$

com

$$\begin{aligned}\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_1 \partial \alpha_1} &= \frac{1}{\alpha_1^2} \left[ 1 - \frac{1}{1 - \rho^2} (3a^2(t_{1i}; \alpha_1, \beta_1) - 2\rho a(t_{1i}; \alpha_1, \beta_1)a(t_{2i}; \alpha_2, \beta_2)) \right], \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_1 \partial \alpha_2} &= \frac{\rho}{(1 - \rho^2)\alpha_1 \alpha_2} a(t_{1i}; \alpha_1, \beta_1)a(t_{2i}; \alpha_2, \beta_2), \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_2 \partial \alpha_2} &= \frac{1}{\alpha_2^2} \left[ 1 - \frac{1}{1 - \rho^2} (3a^2(t_{2i}; \alpha_2, \beta_2) - 2\rho a(t_{1i}; \alpha_1, \beta_1)a(t_{2i}; \alpha_2, \beta_2)) \right].\end{aligned}$$

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} = \begin{pmatrix} \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_1 \partial \beta_1} & \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_1 \partial \beta_2} \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_1 \partial \beta_2} & \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_2 \partial \beta_2} \end{pmatrix},$$

com

$$\begin{aligned}\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_1 \partial \beta_1} &= \frac{1}{\beta_1^2} \left\{ \frac{1}{2} - \frac{\beta_1^2}{(t_{1i} + \beta_1)^2} - \frac{1}{1 - \rho^2} \left[ \left( \frac{a^2(t_{1i}; \alpha_1, \beta_1)}{4} + \frac{1}{\alpha_1^2} \right) + \frac{1}{2\alpha_1^2} \left( \frac{t_{1i}}{\beta_1} - \frac{\beta_1}{t_{1i}} \right) \right. \right. \\ &\quad \left. \left. + \frac{\alpha_1^2}{2} a^2(t_{1i}; \alpha_1, \beta_1) \right) - \frac{\rho a(t_{2i}; \alpha_2, \beta_2)}{\alpha_1} \left( \sqrt{\frac{t_{1i}}{\beta_1}} + \sqrt{\frac{\beta_1}{t_{1i}}} + \frac{\alpha_1}{2} a(t_{1i}; \alpha_1, \beta_1) \right) \right] \right\}, \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_1 \partial \beta_2} &= \frac{\rho}{4(1 - \rho^2)\alpha_1 \alpha_2 \beta_1 \beta_2} \left( \sqrt{\frac{t_{1i} t_{2i}}{\beta_1 \beta_2}} + \sqrt{\frac{\beta_1 t_{2i}}{t_{1i} \beta_2}} + \sqrt{\frac{t_{1i} \beta_2}{\beta_1 t_{2i}}} + \sqrt{\frac{\beta_1 \beta_2}{t_{1i} t_{2i}}} \right), \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_2 \partial \beta_2} &= \frac{1}{\beta_2^2} \left\{ \frac{1}{2} - \frac{\beta_2^2}{(t_{2i} + \beta_2)^2} - \frac{1}{1 - \rho^2} \left[ \left( \frac{a^2(t_{2i}; \alpha_2, \beta_2)}{4} + \frac{1}{\alpha_2^2} \right) + \frac{1}{2\alpha_2^2} \left( \frac{t_{2i}}{\beta_2} - \frac{\beta_2}{t_{2i}} \right) \right. \right. \\ &\quad \left. \left. + \frac{\alpha_2^2}{2} a^2(t_{2i}; \alpha_2, \beta_2) \right) - \frac{\rho a(t_{1i}; \alpha_1, \beta_1)}{\alpha_2} \left( \sqrt{\frac{t_{2i}}{\beta_2}} + \sqrt{\frac{\beta_2}{t_{2i}}} + \frac{\alpha_2}{2} a(t_{2i}; \alpha_2, \beta_2) \right) \right] \right\}.\end{aligned}$$

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\beta}^\top} = \begin{pmatrix} \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_1 \partial \beta_1} & \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_1 \partial \beta_2} \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_2 \partial \beta_1} & \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_2 \partial \beta_2} \end{pmatrix},$$

com

$$\begin{aligned}\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_1 \partial \beta_1} &= -\frac{1}{2(1-\rho^2)\alpha_1^2\beta_1} \left[ \left( \sqrt{\frac{t_{1i}}{\beta_1}} + \sqrt{\frac{\beta_1}{t_{1i}}} \right) (2a(t_{1i}; \alpha_1, \beta_1) + \rho a(t_{2i}; \alpha_2, \beta_2)) \right], \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_1 \partial \beta_2} &= \frac{\rho}{2(1-\rho^2)\alpha_1\alpha_2\beta_2} a(t_{1i}; \alpha_1, \beta_1) \left( \sqrt{\frac{t_{2i}}{\beta_2}} + \sqrt{\frac{\beta_2}{t_{2i}}} \right), \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_2 \partial \beta_1} &= \frac{\rho}{2(1-\rho^2)\alpha_1\alpha_2\beta_1} a(t_{2i}; \alpha_2, \beta_2) \left( \sqrt{\frac{t_{1i}}{\beta_1}} + \sqrt{\frac{\beta_1}{t_{1i}}} \right), \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_2 \partial \beta_2} &= -\frac{1}{2(1-\rho^2)\alpha_2^2\beta_2} \left[ \left( \sqrt{\frac{t_{2i}}{\beta_2}} + \sqrt{\frac{\beta_2}{t_{2i}}} \right) (2a(t_{2i}; \alpha_2, \beta_2) + \rho a(t_{1i}; \alpha_1, \beta_1)) \right].\end{aligned}$$

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\alpha} \partial \rho} = \begin{pmatrix} \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_1 \partial \rho} \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_2 \partial \rho} \end{pmatrix},$$

com

$$\begin{aligned}\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_1 \partial \rho} &= \frac{2\rho}{(1-\rho^2)^2\alpha_1} (a^2(t_{1i}; \alpha_1, \beta_1) - \rho a(t_{1i}; \alpha_1, \beta_1) a(t_{2i}; \alpha_2, \beta_2)) \\ &\quad - \frac{1}{(1-\rho^2)\alpha_1} a(t_{1i}; \alpha_1, \beta_1) a(t_{2i}; \alpha_2, \beta_2), \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \alpha_2 \partial \rho} &= \frac{2\rho}{(1-\rho^2)^2\alpha_2} (a^2(t_{2i}; \alpha_2, \beta_2) - \rho a(t_{1i}; \alpha_1, \beta_1) a(t_{2i}; \alpha_2, \beta_2)) \\ &\quad - \frac{1}{(1-\rho^2)\alpha_2} a(t_{1i}; \alpha_1, \beta_1) a(t_{2i}; \alpha_2, \beta_2).\end{aligned}$$

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \rho} = \begin{pmatrix} \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_1 \partial \rho} \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_2 \partial \rho} \end{pmatrix},$$

com

$$\begin{aligned}\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_1 \partial \rho} &= \frac{\rho}{(1-\rho^2)^2\alpha_1} \left\{ \frac{1}{\alpha_1} \left( \frac{t_{1i}}{\beta_1^2} - \frac{1}{t_{1i}} \right) + \frac{\rho}{\alpha_2\beta_1} \left( \sqrt{\frac{t_{1i}t_{2i}}{\beta_1\beta_2}} - \sqrt{\frac{\beta_1 t_{2i}}{t_{1i}\beta_2}} + \sqrt{\frac{t_{1i}\beta_2}{\beta_1 t_{2i}}} - \sqrt{\frac{\beta_1\beta_2}{t_{1i}t_{2i}}} \right) \right\}, \\ \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_2 \partial \rho} &= \frac{\rho}{(1-\rho^2)^2\alpha_2} \left\{ \frac{1}{\alpha_2} \left( \frac{t_{2i}}{\beta_2^2} - \frac{1}{t_{2i}} \right) + \frac{\rho}{\alpha_1\beta_2} \left( \sqrt{\frac{t_{1i}t_{2i}}{\beta_1\beta_2}} - \sqrt{\frac{\beta_1 t_{2i}}{t_{1i}\beta_2}} + \sqrt{\frac{t_{1i}\beta_2}{\beta_1 t_{2i}}} - \sqrt{\frac{\beta_1\beta_2}{t_{1i}t_{2i}}} \right) \right\}.\end{aligned}$$

Finalmente,

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \rho \partial \rho} = \frac{1 + \rho^2}{(1 - \rho^2)^2} - \frac{2}{(1 - \rho^2)^2} a(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}) \Sigma_{\rho^*}^{-1} a(\mathbf{t}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}),$$

em que  $\Sigma_{\rho^*}^{-1} = (1 - \rho^2)\Sigma_{\rho}^{-1} + (1 - \rho^2)\Sigma^{-1}$ .

Por simplicidade, apresentamos a matriz de informação de Fisher. Lembremos que como as variáveis aleatórias são identicamente distribuídas temos que

$$\mathbf{I}(\boldsymbol{\theta}) = n\mathbf{I}_F,$$

em que

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{I}_{\alpha\alpha} & \mathbf{I}_{\alpha\beta} & \mathbf{I}_{\alpha\rho} \\ \mathbf{I}_{\beta\alpha} & \mathbf{I}_{\beta\beta} & \mathbf{I}_{\beta\rho} \\ \mathbf{I}_{\rho\alpha} & \mathbf{I}_{\rho\beta} & \mathbf{I}_{\rho\rho} \end{pmatrix},$$

em que

$$\begin{aligned} \mathbf{I}_{\alpha\alpha} &= -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^\top} \right) = \begin{pmatrix} -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \alpha_1 \partial \alpha_1} \right) & -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \alpha_1 \partial \alpha_2} \right) \\ -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \alpha_2 \partial \alpha_1} \right) & -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \alpha_2 \partial \alpha_2} \right) \end{pmatrix} \\ &= \begin{pmatrix} \frac{2 - \rho^2}{\alpha_1^2 (1 - \rho^2)} & -\frac{\rho^2}{\alpha_1 \alpha_2 (1 - \rho^2)} \\ -\frac{\rho^2}{\alpha_1 \alpha_2 (1 - \rho^2)} & \frac{2 - \rho^2}{\alpha_2^2 (1 - \rho^2)} \end{pmatrix}, \\ \mathbf{I}_{\alpha\beta} &= -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\beta}^\top} \right) = \begin{pmatrix} -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \alpha_1 \partial \beta_1} \right) & -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \alpha_1 \partial \beta_2} \right) \\ -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \alpha_2 \partial \beta_1} \right) & -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \alpha_2 \partial \beta_2} \right) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \\ \mathbf{I}_{\alpha\rho} &= -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \boldsymbol{\alpha} \partial \rho} \right) = \begin{pmatrix} -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \alpha_1 \partial \rho} \right) \\ -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \alpha_2 \partial \rho} \right) \end{pmatrix} = \begin{pmatrix} -\frac{\rho}{(1 - \rho^2)\alpha_1} \\ -\frac{\rho}{(1 - \rho^2)\alpha_2} \end{pmatrix}, \\ \mathbf{I}_{\beta\rho} &= -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \boldsymbol{\beta} \partial \rho} \right) = \begin{pmatrix} -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \beta_1 \partial \rho} \right) \\ -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \beta_2 \partial \rho} \right) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ \mathbf{I}_{\rho\rho} &= -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \rho \partial \rho} \right) = \frac{1 + \rho^2}{(1 - \rho^2)^2}. \end{aligned}$$

Vamos obter, por exemplo,  $\mathbf{I}_{\beta\beta}$  como ilustração, os outros elementos de  $\mathbf{I}_F(\boldsymbol{\theta})$ , seguem facilmente de (??).

$$\begin{aligned}
 \mathbf{I}_{\beta\beta} &= -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right) = \begin{pmatrix} -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \beta_1 \partial \beta_1} \right) & -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \beta_1 \partial \beta_2} \right) \\ -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \beta_2 \partial \beta_1} \right) & -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \beta_2 \partial \beta_2} \right) \end{pmatrix} \\
 &= \begin{pmatrix} \frac{1}{\beta_1^2} \left[ J(\alpha_1) + \frac{1}{1-\rho^2} \left( \frac{\rho^2}{4} + \frac{1}{\alpha_1^2} \right) \right] & -\frac{\rho(\psi_1(\boldsymbol{\alpha}, \rho) + \psi_1(\boldsymbol{\alpha}, -\rho))}{2\alpha_1\alpha_2\beta_1\beta_2(1-\rho^2)} \\ -\frac{\rho(\psi_1(\boldsymbol{\alpha}, \rho) + \psi_1(\boldsymbol{\alpha}, -\rho))}{2\alpha_1\alpha_2\beta_1\beta_2(1-\rho^2)} & \frac{1}{\beta_2^2} \left[ J(\alpha_2) + \frac{1}{1-\rho^2} \left( \frac{\rho^2}{4} + \frac{1}{\alpha_2^2} \right) \right] \end{pmatrix}.
 \end{aligned}$$

Apresentamos a expressão correta de  $I_{22} = -\mathbb{E} \left( \frac{\partial^2 \ell_i}{\partial \beta_1 \partial \beta_1} \right)$  que foi apresentado por Kundu *et al.* (2010).

$$\begin{aligned}
 \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_1} &= -\frac{1}{2\beta_1} + \frac{1}{(t_1 + \beta_1)} - \frac{1}{(1-\rho^2)} \left\{ \frac{\xi(t_1, \beta_1)}{\alpha_1^2} \frac{\partial \xi(t_1, \beta_1)}{\partial \beta_1} - \frac{\rho}{\alpha_1\alpha_2} \xi(t_2, \beta_2) \frac{\partial \xi(t_1, \beta_1)}{\partial \beta_1} \right\} \\
 \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_1 \partial \beta_1} &= \frac{1}{2\beta_1^2} - \frac{1}{(t_1 + \beta_1)^2} - \frac{1}{(1-\rho^2)} \left\{ \frac{1}{\alpha_1^2} \left( \frac{\partial \xi(t_1, \beta_1)}{\partial \beta_1} \right)^2 + \frac{\xi(t_1, \beta_1)}{\alpha_1^2} \frac{\partial^2 \xi(t_1, \beta_1)}{\partial \beta_1 \partial \beta_1} \right. \\
 &\quad \left. - \frac{\rho}{\alpha_1\alpha_2} \xi(t_2, \beta_2) \frac{\partial^2 \xi(t_1, \beta_1)}{\partial \beta_1 \partial \beta_1} \right\},
 \end{aligned}$$

em que

$$\begin{aligned}
 \left( \frac{\partial \xi(t_1, \beta_1)}{\beta_1} \right)^2 &= \frac{1}{4\beta_1^2} (\alpha_1^2 a_{t_1}^2 + 4), \\
 \xi(t_1, \beta_1) \frac{\partial^2 \xi(t_1, \beta_1)}{\partial \beta_1 \partial \beta_1} &= \alpha_1 a_{t_1} \frac{1}{2\beta_1} \left[ \left( \sqrt{\frac{t_1}{\beta_1}} + \sqrt{\frac{\beta_1}{t_1}} \right) + \frac{\alpha_1}{2} a_{t_1} \right], \\
 \xi(t_2, \beta_2) \frac{\partial^2 \xi(t_1, \beta_1)}{\partial \beta_1 \partial \beta_1} &= \alpha_2 a_{t_2} \frac{1}{2\beta_1} \left[ \left( \sqrt{\frac{t_1}{\beta_1}} + \sqrt{\frac{\beta_1}{t_1}} \right) + \frac{\alpha_1}{2} a_{t_1} \right].
 \end{aligned}$$

Finalmente, aplicamos a esperança na expressão  $-\frac{\partial^2 \ell_i}{\partial \beta_1 \partial \beta_1}$ , com  $a_{t_1}^2 = \frac{\xi^2(t_1, \beta_1)}{\alpha_1^2} \sim \chi_1^2$ .

$$\begin{aligned}
 -\mathbb{E}\left(\frac{\partial^2 \ell_i}{\partial \beta_1 \partial \beta_1}\right) &= -\frac{1}{2\beta_1^2} + \mathbb{E}\left(\frac{1}{(t_1 + \beta_1)^2}\right) + \frac{1}{1 - \rho^2} \left\{ \frac{1}{\alpha_1^2} \mathbb{E}\left(\frac{\partial \xi(t_1, \beta_1)}{\partial \beta_1}\right)^2 \right. \\
 &\quad \left. + \frac{1}{\alpha_1^2} \mathbb{E}\left(\xi(t_1, \beta_1) \frac{\partial^2 \xi(t_1, \beta_1)}{\partial \beta_1 \partial \beta_1}\right) - \frac{\rho}{\alpha_1 \alpha_2} \mathbb{E}\left(\xi(t_2, \beta_2) \frac{\partial^2 \xi(t_1, \beta_1)}{\partial \beta_1 \partial \beta_1}\right) \right\} \\
 &= -\frac{1}{2\beta_1^2} + \mathbb{E}\left(\frac{1}{(t_1 + \beta_1)^2}\right) + \frac{1}{1 - \rho^2} \left( \frac{1}{\alpha_1^2} \frac{1}{4\beta_1^2} (\alpha_1^2 + 4) + \frac{1}{\alpha_1^2} \frac{\alpha_1^2}{4\beta_1^2} - \frac{\rho}{\alpha_1 \alpha_2} \frac{\alpha_1 \alpha_2 \rho}{4\beta_1^2} \right) \\
 &= \frac{1}{\beta_1^2} \left\{ -\frac{1}{4} + \mathbb{E}\left(\frac{\beta_1^2}{(t_1 + \beta_1)^2}\right) + \frac{1}{1 - \rho^2} \left( \frac{1}{4} + \frac{1}{\alpha_1^2} \right) \right\} \\
 &= \frac{1}{\beta_1^2} \left\{ -\frac{1}{4} + J(\alpha_1) + \frac{1}{1 - \rho^2} \left( \frac{1}{4} + \frac{1}{\alpha_1^2} \right) \right\}.
 \end{aligned}$$

Apresentamos também a expressão correta de  $I_{24} = -\mathbb{E}\left(\frac{\partial^2 \ell_i}{\partial \beta_2 \partial \beta_1}\right)$  que foi apresentado por Kundu *et al.* (2010)

$$\begin{aligned}
 \frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial \beta_2 \partial \beta_1} &= -\frac{1}{1 - \rho^2} \left[ -\frac{\rho}{\alpha_1 \alpha_2} \left( \frac{\partial \xi(t_2, \beta_2)}{\partial \beta_2} \right) \left( \frac{\partial \xi(t_1, \beta_1)}{\partial \beta_1} \right) \right] \\
 &= \frac{\rho}{(1 - \rho^2) \alpha_1 \alpha_2} \left[ -\frac{1}{2\beta_2} \left( \sqrt{\frac{t_2}{\beta_2}} + \sqrt{\frac{t_2}{\beta_2}} \right) \right] \left[ -\frac{1}{2\beta_1} \left( \sqrt{\frac{t_1}{\beta_1}} + \sqrt{\frac{t_1}{\beta_1}} \right) \right] \\
 &= \frac{\rho}{4(1 - \rho^2) \alpha_1 \alpha_2 \beta_1 \beta_2} \left( \sqrt{\frac{t_2 t_1}{\beta_2 \beta_1}} + \sqrt{\frac{\beta_2 t_1}{t_2 \beta_1}} + \sqrt{\frac{t_2 \beta_1}{\beta_2 t_1}} + \sqrt{\frac{\beta_2 \beta_1}{t_1 t_2}} \right).
 \end{aligned}$$

Finalmente, aplicamos a esperança na expressão  $-\frac{\partial^2 \ell_i}{\partial \beta_2 \partial \beta_1}$ .

$$\begin{aligned}
 -\mathbb{E}\left(\frac{\partial^2 \ell_i}{\partial \beta_2 \partial \beta_1}\right) &= -\frac{\rho}{4(1 - \rho^2) \alpha_1 \alpha_2 \beta_1 \beta_2} \left[ \mathbb{E}\left(\frac{t_2 t_1}{\beta_2 \beta_1}\right) + \mathbb{E}\left(\sqrt{\frac{\beta_2 t_1}{t_2 \beta_1}}\right) + \mathbb{E}\left(\sqrt{\frac{t_2 \beta_1}{\beta_2 t_1}}\right) \right. \\
 &\quad \left. - \mathbb{E}\left(\sqrt{\frac{\beta_2 \beta_1}{t_1 t_2}}\right) \right] \\
 &= -\frac{\rho}{4(1 - \rho^2) \alpha_1 \alpha_2 \beta_1 \beta_2} (\psi_1(\boldsymbol{\alpha}, \rho) + \psi_1(\boldsymbol{\alpha}, -\rho) + \psi_1(\boldsymbol{\alpha}, -\rho) + \psi_1(\boldsymbol{\alpha}, \rho)) \\
 &= -\frac{\rho}{2(1 - \rho^2) \alpha_1 \alpha_2 \beta_1 \beta_2} (\psi_1(\boldsymbol{\alpha}, \rho) + \psi_1(\boldsymbol{\alpha}, -\rho)).
 \end{aligned}$$

# Anexo I

## Conjuntos de dados

Tabela I.1: Densidade mineral dos ossos

Numero	Dominant radius	Radius
1	1,103	1,052
2	0,842	0,859
3	0,925	0,873
4	0,857	0,744
5	0,795	0,809
6	0,787	0,779
7	0,933	0,880
8	0,799	0,851
9	0,945	0,876
10	0,921	0,906
11	0,792	0,825
12	0,815	0,751
13	0,755	0,724
14	0,880	0,866
15	0,900	0,838
16	0,764	0,757
17	0,733	0,748
18	0,932	0,898
19	0,856	0,786
20	0,890	0,950
21	0,688	0,532
22	0,940	0,850
23	0,493	0,616
24	0,835	0,752

Tabela I.2: Contenido de mineral em ossos

Numero	Dominant radius	Radius	Dominant humerus	Humerus	Dominant ulha	Ulha
1	1.103	1.052	2.139	2.238	.873	.872
2	.842	.859	1.873	1.741	.590	.744
3	.925	.873	1.887	1.809	.767	.713
4	.857	.744	1.739	1.547	.706	.674
5	.795	.809	1.734	1.715	.549	.654
6	.787	.779	1.509	1.474	.782	.571
7	.933	.880	1.695	1.656	.737	.803
8	.799	.851	1.740	1.777	.618	.682
9	.945	.876	1.811	1.759	.853	.777
10	.921	.906	1.954	2.009	.823	.765
11	.792	.825	1.624	1.657	.686	.668
12	.815	.751	2.204	1.846	.678	.546
13	.755	.724	1.508	1.458	.662	.595
14	.880	.866	1.786	1.811	.810	.819
15	.900	.838	1.902	1.606	.723	.677
16	.764	.757	1.743	1.794	.586	.541
17	.733	.748	1.863	1.869	.672	.752
18	.932	.898	2.028	2.032	.836	.805
19	.856	.786	1.390	1.324	.578	.610
20	.890	.950	2.187	2.087	.758	.718
21	.688	.532	1.650	1.378	.533	.482
22	.940	.850	2.334	2.225	.757	.731
23	.493	.616	1.037	1.268	.546	.615
24	.835	.752	1.509	1.422	.618	.664
25	.915	.936	1.971	1.869	.869	.868

## Anexo II

# Código C++ da aproximação feita por Meyer (2013)

```
using namespace std;

double phi(double x)
{
    const double PI=3.141592653589793238463;

    return (1/sqrt(2*PI))*exp(-0.5*x*x);
}

double phi2(double x, double y, double rho)
{
    const double PI=3.141592653589793238463;

    return (1/(2*PI*sqrt(1-rho*rho)))*exp((-0.5/(1-rho*rho))*(x*x - 2*rho*x*y + y*y));
}

double Phi(double x)
{
    long double s=x;
    long double t=0;
    long double b=x;
    long double q=x*x;
    long double i=1;

    while(s!=t)
```

```
s=(t=s)+(b*=q/(i+=2));

return .5+s*exp(-.5*q-.91893853320467274178L);
}

double sqr( double x )
{
    return x * x;
}

double Phi2diag( double x, double a, double px, double pxs )
{
    if(a <= 0.0)
        return px;

    if(a >= 1.0)
        return px * px;

    double b = 2.0 - a;
    double sqrt_ab = sqrt(a*b);
    double c1 = 6.36619772367581343e-001;
    double c2 = 1.25331413731550025;
    double c3 = 1.57079632679489662;
    double c4 = 1.591549430918953358e-001;
    double asr = ( a > 0.1 ? asin( 1.0 - a ) : acos( sqrt_ab ) );
    double comp = px * pxs;

    if( comp*( 1.0 - a - c1*asr ) < 5e-17 )
        return b * comp;

    double tmp = c2 * x;
    double alpha = a * x * x / b;
    double a_even = -tmp * a;
    double a_odd = -sqrt_ab * alpha;
    double beta = x * x;
    double b_even = tmp * sqrt_ab;
    double b_odd = sqrt_ab * beta;
    double delta = 2.0 * x * x / b;
    double d_even = ( 1.0 - a ) * c3 - asr;
    double d_odd = tmp * ( sqrt_ab - a );
    double res = 0.0, res_new = d_even + d_odd;
    int k = 2;
```

```

while( res != res_new )
{
    d_even = ( a_odd + b_odd + delta * d_even ) / k;
    a_even *= alpha / k;
    b_even *= beta / k;
    k++;
    a_odd *= alpha / k;
    b_odd *= beta / k;
    d_odd = ( a_even + b_even + delta * d_odd ) / k;
    k++;
    res = res_new;
    res_new += d_even + d_odd;
}

res *= exp( -x * x / b ) * c4;
return fmax( ( 1.0 + c1 * asr ) * comp, b * comp - fmax( 0.0, res ) );
}

```

```

double Phi2help( double x, double y, double rho )
{
    double s = sqrt( ( 1.0 - rho ) * ( 1.0 + rho ) );
    double a = 0.0, b1 = -fabs( x ), b2 = 0.0;

    if( rho > 0.99 )
    {
        double tmp = sqrt( ( 1.0 - rho ) / ( 1.0 + rho ) );
        b2 = -fabs( ( x - y ) / s - x * tmp );
        a = sqr( ( x - y ) / x / s - tmp );
    }
    else if( rho < -0.99 )
    {
        double tmp = sqrt( ( 1.0 + rho ) / ( 1.0 - rho ) );
        b2 = -fabs( ( x + y ) / s - x * tmp );
        a = sqr( ( x + y ) / x / s - tmp );
    }
    else
    {
        b2 = -fabs( rho * x - y ) / s;
        a = sqr( b2 / x );
    }

    double p1 = Phi( b1 );
    double p2 = Phi( b2 );
}

```

```

double q = 0.0;

if( a <= 1.0 )
    q = 0.5 * Phi2diag( b1, 2.0 * a / ( 1.0 + a ), p1, p2 );
else
    q = p1 * p2 - 0.5 * Phi2diag( b2, 2.0 / ( 1.0 + a ), p2, p1 );

int c1 = ( y / x >= rho ), c2 = ( x < 0.0 ), c3 = c2 && ( y >= 0.0 );

return ( c1 && c3 ? q - 0.5
: c1 && c2 ? q
: c1 ? 0.5 - p1 + q
: c3 ? p1 - q - 0.5
: c2 ? p1 - q
: 0.5 - q );
}

double Phi2( double x, double y, double rho )
{
    if( ( 1.0 - rho ) * ( 1.0 + rho ) <= 0.0 )
        if( rho > 0.0 )
            return Phi( fmin(x, y) );
    else
        return fmax( 0.0, fmin( 1.0, Phi( x ) + Phi( y ) - 1.0 ) );

    if( x == 0.0 && y == 0.0 )
        if( rho > 0.0 )
            return Phi2diag( 0.0, 1.0 - rho, 0.5, 0.5 );
    else
        return 0.5 - Phi2diag( 0.0, 1.0 + rho, 0.5, 0.5 );

    return max( 0.0, min( 1.0, Phi2help( x, y, rho ) + Phi2help( y, x, rho ) ) );
}

```

# Anexo III

## Licença

Copyright (c) 2014 de Luis Enrique Benites Sánchez.

Exceto quando indicado o contrário, esta obra está licenciada sob a licença Creative Commons Atribuição 3.0 Não Adaptada. Para ver uma cópia desta licença, visite <http://creativecommons.org/licenses/by/3.0/>.



A marca e o logotipo da UNICAMP são propriedade da Universidade Estadual de Campinas. Maiores informações sobre encontram-se disponíveis em <http://www.unicamp.br/unicamp/a-unicamp/logotipo/normas%20oficiais-para-uso-do-logotipo>.

### III.1 Sobre a licença dessa obra

A licença Creative Commons Atribuição 3.0 Não Adaptada utilizada nessa obra diz que:

1. Você tem a liberdade de:

- Compartilhar — copiar, distribuir e transmitir a obra;
- Remixar — criar obras derivadas;
- fazer uso comercial da obra.

2. Sob as seguintes condições:

- Atribuição — Você deve creditar a obra da forma especificada pelo autor ou licenciante (mas não de maneira que sugira que estes concedem qualquer aval a você ou ao seu uso da obra).



# Referências Bibliográficas

- Aarset, M. (1987). How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, **36(1)**, 106–108.
- Abramowitz, M. & Stegun, I. A. (1974). *Handbook of Mathematical Functions*. Dover, New York.
- Albers, W. & Kallenberg, C. (1994). A simple approximation to the bivariate normal distribution with large correlation coefficient. *Journal of Multivariate Analysis*, **49**, 87–96.
- Arnold, B. & Zahedi, H. (1988). On multivariate mean remaining life functions. *Journal of Multivariate Analysis*, **25**, 1–9.
- Balakrishnan, N. & Lai, C.-D. (2009). *Continuous Bivariate Distributions*. Second Edition. Springer, New York.
- Barlow, R. & Campo, R. (1974). Total time on test processes and applications to failure data analysis. *Reliability and fault tree analysis*, pages 451–481.
- Barros, M., Paula, G. & Leiva, V. (2008). A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. *Lifetime Data Analysis*, **14**, 316–332.
- Basu, A. P. (1971). Bivariate failure rate. *Journal of Applied Probability*, **66**, 103–104.
- Bebbington, M., C-D., L. & Zitikis, R. (2008). A proof of the shape of the birnbaum-saunders hazard rate function. *Mathematical Scientist*, **33**, 49–56.
- Bergman, B. (1979). On age replacement and the total time on test concept. *Scand. J. Statist*, **6**, 161–168.
- Birnbaum, Z. W. & Saunders, S. C. (1969a). A new family of life distributions. *Journal of Applied Probability*, **6**, 319–327.
- Birnbaum, Z. W. & Saunders, S. C. (1969b). Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability*, **6**, 328–347.
- Birnbaum, Z. W. & Saunders, S. C. (1958). A statistical model for life-length of materials. *Journal of the American Statistical Association*, **53**, 151–160.

- Birnbaum, Z. W. & Saunders, S. C. (1968). A probabilistic interpretation of miner's rule. *Journal of Applied Mechanics*, **16**, 637–652.
- Bo-Cheng, W. (1998). *Exponential Family Nonlinear Models*. Springer, Singapore.
- Brown, B. W. & Flood, M. M. (1958). Tumbler mortality. *Journal of the American Statistical Association*, **42**, 562–574.
- Buchanan, W. & Singpurwalla, N. (1977). Some stochastic characterization of multivariate survival. *Theory and Applications of Reliability*, **1**, 329–348.
- Cecilia, A., Víctor, L., A., E. & Balakrishnan, N. (2012). Shape and change point analyses of the birnbaum–saunders-t hazard rate and associated estimation. *Computational Statistics Data Analysis*, **56**, 3887–3897.
- Charlier, C. V. (1931). Application de la théorie des probabilités à l'astronomie. traité du calcul des probabilités et de ses applications. *Paris: Gauthier-Villais et Cie.*, **2**.
- Collett, D. (1994). *Modelling Survival Data in Medical Research*. Chapman & Hall, Boca Raton, FL.
- Colosimo, E. A. & Giolo, S. (2006). *Análise de sobrevivência aplicada*. Edgard Blücher, Brasil.
- Cook, R. (1977). Detection of influential observation in linear regression. *Technometrics*, **19**, 15–18.
- Cook, R. (1986). Assessment of local influence. *Journal of the Royal Statistical Society. Series B (Methodological)*, **48**, 133–169.
- Cook, R. & Weisberg, S. (1982). *Residuals and Influence in Regression*. Chapman & Hall, UK.
- Cook, R. & Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression. *Biometrika*, **70**, 1–10.
- Cook, R. D. (1987). Influence assessment. *Journal of Applied Statistics*, **14**, 117–131.
- Cooray, K. (2006). A generalization of the weibull distribution: the odd weibull family. *Statistical Modelling*, **6(3)**, 265–277.
- Cox, D. & Hinkley, D. (1974). *Theoretical Statistics*. Chapman & Hall, UK.
- Cox, D. & Wermuth, N. (1991). A simple approximation for bivariate and trivariate normal integrals. *International Statistical Review*, **59**, 263–269.
- Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, Princeton, USA.
- Cysneiro, A., Cribari-Neto, F. & Araujo, C. (2008). On birnbaum-saunders inference. *Computational Statistics and Data Analysis*, **52**, 4939–4950.

- Dabrowska, D. (1998). Kaplan-meier estimate on the plane. *The Annals of Statistics*, **16**, 1475–1489.
- Desmond, A. (1985). Stochastic models of failure in random environments. *Canadian Journal of Statistics*, **13**, 171–183.
- Desmonde, A. F. (1986). On the relationship between two fatigue-life models. *IEEE Transactions on Reliability*, **35**, 167–169.
- Divgi, D. (1979). Calculation of univariate and bivariate normal probability functions. *The Annals of Statistics*, **7**, 903–910.
- Donnelly, T. (1973a). Algorithm 462: Bivariate normal distribution. *Communications of the Association for Computing Machinery*, **16**, 636.
- Donnelly, T. (1973b). Algorithm 462: Bivariate normal distribution. *Communications of the Association for Computing Machinery*, **16**, 636.
- Drezner, Z. & Wesolowsky, G. O. (1990). The computation of the bivariate normal integral. *Journal of Statistical Computation and Simulation*, **35**, 101–107.
- Engelhardt, M., Bain, L. J. & Wright, F. T. (1981). Inferences on the parameters of the birnbaum-saunders fatigue life distribution based on maximum likelihood estimation. *Technometrics*, **23**, 251–255.
- Gassmann, H. (2000). Rectangle probabilities of trivariate normal distributions. *Technical report available at <http://www.mgmt.dal.ca/sba/profs/hgassmann/>*.
- Genz, A., Bretz, F. & Hochberg, Y. (2000). Approximations to multivariate  $t$  probabilities with application to multiple comparison procedures. *Technical report available at <http://www.math.wsu.edu/faculty/genz>*.
- Guiraud, P., Leiva, V. & Fierro, R. (2009). A non-central version of the birnbaum-saunders distribution for reliability analysis. *IEEE Transactions on Reliability*, **58**, 152–160.
- Gupta, S. (1963). Probability integral of multivariate normal and multivariate  $t$ . *The Annals of Statistics*, **34**, 792–828.
- Haug, E. G. (1998). *The complete guide to option pricing formulas*. McGraw - Hill.
- Hull, J. (2002). *Options, Futures, and Other Derivatives*. fifth edn, Prentice Hall.
- James, B. R. (2006). *Probabilidade: Um curso em nível intermediário*. Projeto Euclides. IMPA.
- Janurová, K. & Briš, R. (2013). A nonparametric approach to medical survival data: Uncertainty in the context of risk in mortality analysis. *Reliability Engineering and System Safety*.
- Johnson, N. & Kotz, S. (1973). A vector valued multivariate hazard rate. *Institute of Statistics, University of North Carolina, Mimeo Series No. 873*.

- Johnson, N. & Kotz, S. (1975). A vector multivariate hazard rate. *Journal of Multivariate Analysis*, **5**, 53–66.
- Johnson, N. L., Kotz, S. & Balakrishnan, N. (1994). *Continuous Univariate Distributions, Vol. 1, Second Edition*. John Wiley and Sons, New York.
- Johnson, N. L., Kotz, S. & Balakrishnan, N. (1995). *Continuous Univariate Distributions, Vol. 2, Second Edition*. John Wiley and Sons, New York.
- Johnson, N. L., Kotz, S. & Balakrishnan, N. (1997). *Discrete Multivariate Distributions*. John Wiley and Sons, New York.
- Johnson, R. & Wichern, D. (1998). *Applied Multivariate Statistical Analysis*. Pearson, US.
- Kalbfleisch, J. D. & Prentice, R. (2002). *The Statistical Analysis of Failure Time Data*. John Wiley and Sons, New York.
- Kleinbaum, D. G. & Klein, M. (2005). *Survival Analysis, Second Edition*. Springer, New York.
- Kleinbaum, D. G. & Klein, M. (2011). *Survival Analysis, 3rd Edition*. Springer, New York.
- Kotz, S., Balakrishnan, N. & Johnson, N. L. (2000). *Continuous Multivariate Distributions - Vol. 1, Second Edition*. John Wiley and Sons, New York.
- Kundu, D., K., N. & Balakrishnan, N. (2008). On the hazard function of birnbaum–saunders distribution and associated inference. *Computational Statistics Data Analysis*, **52**, 2692–2702.
- Kundu, D., Balakrishnan, N. & Jamalizadeh, A. (2010). Bivariate birnbaum saunders distribution and associated inference. *Journal of Multivariate Analysis*, **101**, 113–125.
- Kundu, D., Balakrishnan, N. & Jamalizadeh, A. (2013). Generalized multivariate birnbaum–saunders distributions and related inferential issues. *Computational Statistics and Data Analysis*, **116**, 230–244.
- Leiva, V., Barros, M., Paula, G. & Galea, M. (2007). Influence diagnostics in log-birnbaum-saunders regression models with censored data. *Computational Statistics and Data Analysis*, **51**, 5694–5707.
- Leiva, V., Barros, M., Paula, G. & Sanhueza, A. (2008). Generalized birnbaum-saunders distributions applied to air pollutant concentration. *Environmetrics*, **19**, 235–249.
- Lemonte, A. J. (2013). Multivariate birnbaum–saunders regression model. *Journal of Statistical Computation and Simulation*, **83**, 2244–2257.
- Lemonte, A. J. & Patriota, A. G. (2011). Influence diagnostics in birnbaum–saunders nonlinear regression models. *Journal of Applied Statistics*, **38**, 871–884.
- Li, L. & From, S. (2006). Estimation of the parameters of the birnbaum-saunders distribution. *Communications in Statistics Theory and Methods*, **35**, 2157–2169.

- Loynes, R. M. (2001). A new measure in local influence. *Journal of Statistical Planning and Inference*, **92**, 47–53.
- Mann, N. R., Schafer, R. E. & Singpurwalla, N. (1974). *Methods for Statistical Analysis of Reliability and Life Data*. John Wiley Sons, New York.
- Marsaglia, G. (2004). Evaluating the normal distribution. *Journal of Statistical Software*, **11(4)**, 1–11.
- Marshall, A. (1975). Some comments on hazard gradient". *Stochastic Process and Applications*, **3**, 293–300.
- Marshall, A. W. & Olkin, I. (2007). *Life Distributions*. Springer, New York.
- Meyer, C. (2013). The bivariate normal copula. *Communications in Statistics - Theory and Methods*.
- Miner, M. A. (1945). Cumulative damage in fatigue. *Journal of Applied Mechanics*, **12**, 159–164.
- Mittelbach, F., Goossens, M., Braams, J., Carlisl, D. & Rowley, C. (2004). *The LATEX Companion. Tools and Techniques for Computer Typesetting*. Addison Wesley, Boston.
- Mudholkar, G., Srivastava, D. & Kollia, G. (1996). A generalization of the weibull distribution with application to the analysis of survival data. *Reliability and fault tree analysis*, **91**, 1575–1583.
- Navarro, J. & J., M. S. (2011). Reliability properties of bivariate conditional proportional hazard rate models. *Journal of Multivariate Analysis*, **113**, 116–127.
- Nelson, W. (1972). Graphical analysis of accelerated life test data with the inverse power law model. *IEEE Transactions on Reliability*, **21**, 2–11.
- Ng, H., Kundu, D. & Balakrishnan, N. (2003). Modified moment estimation for the two-parameter birnbaumsaunders distribution. *Computational Statistics and Data Analysis*, **43**, 283–298.
- Ng, H., a D., K. & Balakrishnan, N. (2006). Point and interval estimation for the two-parameter birnbaum-saunders distribution based on type-ii censored samples. *Computational Statistics and Data Analysis*, **50**, 3222–3242.
- Owen, D. (1956). Tables for computing bivariate normal probabilities. *Annals of Mathematical Statistics*, **27**, 1075–1090.
- Owen, W. & Padgett, W. (1999). Accelerated test models for system strength based on birnbaum-saunders distribution. *Lifetime Data Anal.*, **5**, 133–147.
- Paula, G., Leiva, V.; Barros, M. & Liu, S. (2012). Robust statistical modeling using the birnbaum-saunders-t distribution applied to insurance. *Applied Stochastic Models in Business and Industry*, **28**, 16–34.

- Paula, G. A., Leiva, V., M., B. & Liu, S. (2011). Robust statistical modeling using the birnbaum–saunders-t distribution applied to insurance. *Applied Stochastic Models in Business and Industry*, **28**, 16–34.
- Plackett, R. L. (1954). A reduction for normal multivariate integrals. *Biometrika*, **41**, 351–360.
- Poon, W. Y. & Poon, Y. S. (1999). Conformal normal curvature and assessment of local influence. *Journal of the Royal Statistical Society B*, **61**, 51–61.
- Santana, L. R. (2009). *Estimação e Diagnóstico em Modelos Birnbaum-Saunders Skew-Normal*. Ph.D. thesis, Instituto De Matemática, Estatística e Computação Científica, Campinas, Brasil.
- Shanbag, D. & Kotz, S. (1987). Some new approaches to multivariate probability distributions. *Journal of Multivariate Analysis*, **22**, 189–211.
- Terza, J. & Welland, U. (1991). A comparison of bivariate normal algorithms. *Journal of Statistical Computation and Simulation*, **39**, 115–127.
- Tsay, W.-J. & Ke, P.-H. (2009). A simple approximation for bivariate normal integral based on error function and its application on probit model with binary endogenous regressor. *The Institute of Economics, Academia Sinica, Taiwan*.
- Verveke, G. & Molenberghs, G. (1997). *Linear Mixed Models in Practice*. Springer-Verlag. New York.
- Yashin, A. & Manton, K. (1997). Effects of unobserved and partially observed covariate processes on system failure: a review of models and estimation strategies. *Statistical Science*, **12**, 20–34.
- Zhu, H. & Lee, S. (2001). Local influence for incomplete data models. *Journal of the Royal Statistical Society B*, **63**, 111–126.