

Modelos Box-Cox elípticos

Raúl Alejandro Morán Vásquez

TRABALHO APRESENTADO
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
DOUTOR EM CIÊNCIAS

Programa: Estatística

Orientadora: Silvia Lopes de Paula Ferrari

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da
CAPES e do CNPq

São Paulo, março de 2017

Modelos Box-Cox elípticos

Este exemplar corresponde à redação final da tese devidamente corrigida e defendida por Raúl Alejandro Morán Vásquez e aprovada pela Banca Examinadora.

Banca Examinadora:

- Profa. Silvia Lopes de Paula Ferrari (orientadora) - IME-USP.
- Prof. 2
- Prof. 3
- Prof. 4
- Prof. 5

Agradecimentos

À minha mãe Maria Patricia Vásquez Román e meu pai Ricardo Antonio Morán Bejarano, pelo amor e apoio incondicionais. Aos meus irmãos, Rubén Oswaldo Morán Vásquez e Ricardo Felipe Morán Vásquez, pelo apoio constante.

A Diana Carolina Londoño Moscoso, pelo carinho e paciência.

Ao professor Daya Krishna Nagar, pelo incentivo e preparação para o programa de doutorado.

A minha orientadora, Silvia Lopes de Paula Ferrari, pela confiança, paciência e excelente orientação que recebi durante o desenvolvimento deste trabalho.

A Giovana Fumes, pelos conselhos e contribuições a este trabalho.

Ao professor José Eduardo Corrente, por ceder o banco de dados para este trabalho.

Aos colegas do IME, pelos bons momentos.

Aos professores do IME, pelo conhecimento transmitido.

À CAPES e ao CNPq, pelo financiamento durante o programa de doutorado.

Dedico este trabalho a meus pais, Patricia Vásquez e Antonio Morán.

Resumo

Dados positivos multivariados aparecem com frequência em diversas áreas de estudo. A transformação de Box-Cox multivariada é uma metodologia habitualmente utilizada para modelar esse tipo de dados. Essa abordagem apresenta algumas desvantagens, como por exemplo a falta de interpretação dos parâmetros em termos de características do vetor de variáveis originais. Neste trabalho estudamos a classe de distribuições Box-Cox elípticas, que é uma alternativa para a modelagem de dados positivos multivariados através da transformação de Box-Cox multivariada. Definimos essa classe através de uma extensão da transformação de Box-Cox multivariada, e envolvendo uma nova classe de distribuições que denominamos de classe de distribuições elípticas truncadas, que também estudamos neste trabalho. A classe de distribuições Box-Cox elípticas tem como casos particulares as classes de distribuições log-elípticas e Box-Cox simétricas. Os parâmetros que conformam esta nova classe são interpretáveis em termos de características do vetor de variáveis originais, o que permite modelar dados positivos multivariados, marginalmente assimétricos e com presença de observações discrepantes. Além disso, alguns parâmetros estão relacionados a quantis das distribuições marginais, tornando esta classe atrativa para modelagem de regressão. Para abordar o problema de estimação dos parâmetros propomos o método de máxima verossimilhança. Estudamos aspectos teóricos e computacionais associados a essa metodologia, cuja adequação é verificada por meio de estudos de simulação. Posteriormente, desenvolvemos modelos de regressão lineares Box-Cox elípticos, que têm como casos particulares os modelos de regressão lineares log-elípticos e Box-Cox simétricos, que, por sua vez, também constituem uma nova contribuição à literatura estatística. Descrevemos o método de máxima verossimilhança aplicado a estes modelos e propomos métodos de diagnóstico para avaliar ajustes dos modelos de regressão lineares log-normal e log- t multivariados. Apresentamos aplicações das distribuições Box-Cox elípticas e dos modelos de regressão lineares Box-Cox elípticos a dados reais.

Palavras chave: distribuições Box-Cox simétricas; distribuições elípticas; distribuição truncada; modelos de regressão lineares multivariados; transformação de Box-Cox.

Abstract

Positive multivariate data are often encountered in several study areas. The multivariate Box-Cox transformation is a methodology usually employed to model this type of data. This approach has some drawbacks, one of them being the lack of interpretation of the parameters in terms of the vector of the original variables. In this work we study the Box-Cox elliptical class of distributions, which is an alternative strategy for modeling multivariate positive data through the multivariate Box-Cox transformation. We define this class of models through an extension of the multivariate Box-Cox transformation, involving a new class of distributions called truncated elliptical class of distributions, which we also study in this work. The Box-Cox elliptical class of distributions has as particular cases the log-elliptical and Box-Cox symmetric classes of distributions. The parameters that index this new class of distributions are interpretable in terms of characteristics of the vector of the original variables, which allows the modeling of multivariate positive data, marginally asymmetric in the presence of outliers. Furthermore, some parameters are related to quantiles of the marginal distributions, making this class attractive for regression modeling. To tackle the parameter estimation problem we propose the maximum likelihood method. We study theoretical and computational aspects associated with this methodology, whose adequacy is verified through simulation studies. Subsequently, we define a new class of regression models, the Box-Cox elliptical linear regression models, which have as particular cases the log-elliptical and Box-Cox symmetric linear regression models, that are not yet available in the literature. We describe the maximum likelihood method applied to these models and propose diagnostic methods to evaluate the goodness of fit in the multivariate log-normal and log- t linear regression models. We present applications of the Box-Cox elliptical distributions and Box-Cox elliptical linear regression models to real data.

Keywords: Box–Cox symmetric distributions; Box-Cox transformation; elliptical distributions; multivariate linear regression models; truncated distribution.

Índice

Lista de Figuras	1
Lista de Tabelas	3
Glossário de notações e abreviaturas	5
1 Introdução	9
1.1 Formulação do problema	9
1.2 Objetivos e estrutura da tese	10
1.3 Suporte computacional	11
2 A classe de distribuições elípticas truncadas	13
2.1 Introdução	13
2.2 A classe de distribuições elípticas	14
2.3 A classe de distribuições elípticas truncadas	19
2.4 Propriedades das distribuições elípticas truncadas	23
2.5 Geração de amostras aleatórias	27
2.6 Conclusões	29
3 A classe de distribuições Box-Cox elípticas	31
3.1 Introdução	31
3.2 A transformação de Box-Cox estendida	34
3.3 A classe de distribuições Box-Cox elípticas	35
3.4 Propriedades das distribuições Box-Cox elípticas	39
3.5 Interpretação dos parâmetros	44
3.6 Estimação de máxima verossimilhança	48
3.7 Aplicações a dados reais	51
3.8 Conclusões	62
4 Modelos de regressão linear Box-Cox elípticos	63
4.1 Introdução	63
4.2 Modelos de regressão lineares Box-Cox elípticos	64

4.3	Estimação de máxima verossimilhança	66
4.4	Diagnóstico para os modelos log-normal e log- t multivariados	68
4.5	Aplicações a dados reais	69
4.6	Conclusões	73
5	Considerações finais e propostas futuras	75
A	Provas dos teoremas e corolários	77
A.1	Prova Teorema 2.3.1	77
A.2	Prova Teorema 2.4.1	77
A.3	Prova Corolário 2.4.1.1	78
A.4	Prova Corolário 2.4.1.2	78
A.5	Prova Teorema 2.4.2	78
A.6	Prova Teorema 2.4.3	78
A.7	Prova Teorema 2.4.4	79
A.8	Prova Teorema 2.4.5	79
A.9	Prova Teorema 2.4.6	80
A.10	Prova Teorema 2.4.7	80
A.11	Prova Teorema 3.2.1	81
A.12	Prova Teorema 3.2.2	81
A.13	Prova Teorema 3.3.1	82
A.14	Prova Teorema 3.4.1	83
A.15	Prova Teorema 3.4.2	84
A.16	Prova Teorema 3.4.3	85
A.17	Prova Teorema 3.4.4	85
A.18	Prova Teorema 3.4.5	86
A.19	Prova Teorema 3.4.6	86
A.20	Prova Teorema 3.4.8	87
A.21	Prova Corolário 3.4.8.2	87
A.22	Prova Teorema 3.5.1	88
A.23	Prova Corolário 3.5.1.1	88
A.24	Prova Teorema 4.4.1	89
B	Programação em R	91
	Referências Bibliográficas	103

Lista de Figuras

2.1	Curvas de níveis e gráficos da f.d.p de $X \sim Tt_2(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \tau; B)$	22
3.1	Curvas de níveis e gráficos da f.d.p de $Y \sim BCt_2(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; \tau)$	38
3.2	Matriz de bagplots das variáveis nutricionais.	54
3.3	Curvas de níveis e gráficos das f.d.p das distribuições ajustadas a dados nutricionais.	57
3.4	Matriz de bagplots das variáveis antropométricas.	58
3.5	Curvas de níveis e gráficos das f.d.p das distribuições ajustadas a dados antropométricos.	61
4.1	<i>QQ-plots</i> das distâncias de Mahalanobis empíricas dos ajustes dos modelos de regressão lineares a dados nutricionais.	70
4.2	<i>QQ-plots</i> das distâncias de Mahalanobis empíricas dos ajustes dos modelos de regressão lineares a dados antropométricos.	72

Lista de Tabelas

3.1	Valores dos parâmetros nos estudos de simulação.	50
3.2	Valores da mediana, viés mediano, desvio absoluto mediano e amplitude interquartil associados ao estudo de simulação 1.	52
3.3	Valores da mediana, viés mediano, desvio absoluto mediano e amplitude interquartil associados ao estudo de simulação 2.	52
3.4	Valores da mediana, viés mediano, desvio absoluto mediano e amplitude interquartil associados ao estudo de simulação 3.	53
3.5	Valores da mediana, viés mediano, desvio absoluto mediano e amplitude interquartil associados ao estudo de simulação 4.	53
3.6	Estatísticas descritivas das variáveis nutricionais.	54
3.7	Correlações amostrais entre as variáveis nutricionais.	55
3.8	Valores do AIC das distribuições bivariadas ajustadas a duplas de variáveis nutricionais.	56
3.9	Estimativas (e erros padrão) dos parâmetros dos modelos ajustados a duplas de variáveis nutricionais.	56
3.10	Estatísticas descritivas das variáveis antropométricas.	59
3.11	Correlações amostrais entre as variáveis antropométricas.	59
3.12	Valores do AIC das distribuições bivariadas ajustadas a duplas de variáveis antropométricas.	60
3.13	Estimativas (e erros padrão) dos parâmetros dos modelos ajustados a duplas de variáveis antropométricas.	60
4.1	Estatísticas descritivas das variáveis nutricionais segundo o sexo.	69
4.2	Valores do AIC dos modelos de regressão lineares multivariados ajustados a dados nutricionais.	70
4.3	Estimativas (e erros padrão) dos parâmetros do modelo de regressão linear log- t multivariado ajustado a dados nutricionais.	71
4.4	Estatísticas descritivas das variáveis antropométricas segundo o sexo.	71
4.5	Valores do AIC dos modelos de regressão lineares multivariados ajustados a dados antropométricos.	72

4	LISTA DE TABELAS	0.0
4.6	Estimativas (e erros padrão) dos parâmetros do modelo de regressão linear log- t bivariado ajustado a dados antropométricos.	73

Glossário de notações e abreviaturas

- \mathbb{R}^p : Espaço Euclidiano p -dimensional.
- \mathbf{a} : Vetor coluna com elementos a_1, \dots, a_p ; $\mathbf{a} = (a_1, \dots, a_p)'$.
- $\mathbf{1}$: Vetor coluna de uns; $\mathbf{1} = (1, \dots, 1)'$.
- \mathbb{R}_+^p : Hiperoctante positivo; $\mathbb{R}_+^p = \{\mathbf{a} \in \mathbb{R}^p : a_k > 0, k = 1, \dots, p\}$.
- \mathbf{a}_{-k} : Subvetor obtido tirando o k -ésimo componente do vetor \mathbf{a} .
- $\mathbf{A}(p \times q)$: Matriz com p linhas e q colunas.
- $\mathbf{A} = (a_{jk})$: Matriz com elementos a_{jk} .
- \mathbf{A}' : Transposta da matriz \mathbf{A} .
- \mathbf{A}^{-1} : Inversa da matriz não-singular \mathbf{A} .
- \mathbf{I}_p : Matriz identidade de dimensão $p \times p$.
- $\mathbf{D}_\mathbf{a}$: Matriz diagonal conformada com os componentes de \mathbf{a} na diagonal principal; $\mathbf{D}_\mathbf{a} = \text{diag}\{a_1, \dots, a_p\}$.
- $\det(\mathbf{A})$: Determinante da matriz \mathbf{A} .
- $\mathbf{A} > 0$: Matriz \mathbf{A} definida positiva.
- $\mathbf{A} \geq 0$: Matriz \mathbf{A} semidefinida positiva.

- $\mathbf{A}_{k,-k}$: Subvetor linha obtido tirando o k -ésimo componente da k -ésima linha da matriz \mathbf{A} .
- $\mathbf{A}_{-k,k}$: Subvetor coluna obtido tirando o k -ésimo componente da k -ésima coluna da matriz \mathbf{A} .
- $\mathbf{A}_{-k,-k}$: Submatriz obtida ao excluir a k -ésima linha e a k -ésima coluna de \mathbf{A} .
- $\mathcal{O}(p)$: Grupo ortogonal; $\mathcal{O}(p) = \{\mathbf{H}(p \times p) : \mathbf{H}\mathbf{H}' = \mathbf{H}'\mathbf{H} = I_p\}$.
- $T : A \rightarrow B$: Transformação de domínio A e contradomínio B .
- $T(A)$: Imagem direta do conjunto A sob a transformação T .
- $T \circ U$: Transformação composta de T e U .
- $J(\mathbf{x} \rightarrow \mathbf{y})$: Jacobiano da transformação $\mathbf{y} = T(\mathbf{x})$.
- $\int_B f(\mathbf{x}) d\mathbf{x}$: Integral de $f(\mathbf{x})$ sobre o conjunto $B \subseteq \mathbb{R}^p$, em que $d\mathbf{x} = \prod_{i=1}^p dx_i$.
- i : Unidade imaginária; $i = \sqrt{-1}$
- $\text{Re}(z)$: Parte real do número complexo z .
- $|z|$: Módulo do número complexo z .
- $\text{arg}(z)$: Argumento do número complexo z .
- Γ : Função gama.
- γ : Função gama incompleta.
- K_q : Função de Bessel modificada de terceiro tipo.
- $P(\mathbf{X} \in A)$: Probabilidade de que os valores do vetor aleatório \mathbf{X} pertençam ao conjunto A .

$E(\mathbf{X})$:	Valor esperado do vetor aleatório \mathbf{X} .
$\text{Var}(\mathbf{X})$:	Matriz de covariância do vetor aleatório \mathbf{X} .
$\text{Cov}(\mathbf{X}, \mathbf{Y})$:	Matriz de covariância dos vetores aleatórios \mathbf{X} e \mathbf{Y} .
$\text{Corr}(\mathbf{X}, \mathbf{Y})$:	Matriz de correlação dos vetores aleatórios \mathbf{X} e \mathbf{Y} .
\sim	:	Distribuído como.
U	:	Distribuição uniforme.
IG	:	Distribuição gama inversa.
χ^2_ν	:	Distribuição qui-quadrado.
F_{ν_1, ν_2}	:	Distribuição F .
BCS	:	Distribuição Box-Cox simétrica.
LS	:	Distribuição log-simétrica.
S_p	:	Distribuição esférica.
El_p	:	Distribuição elíptica.
N_p	:	Distribuição normal multivariada.
t_p	:	Distribuição t multivariada.
SN_p	:	Distribuição skew-normal multivariada.
St_p	:	Distribuição skew- t multivariada.
TS_p	:	Distribuição esférica truncada.
TEl_p	:	Distribuição elíptica truncada.
TN_p	:	Distribuição normal multivariada truncada.

Tt_p	: Distribuição t multivariada truncada.
$BCE\ell_p$: Distribuição Box-Cox elíptica.
BCN_p	: Distribuição Box-Cox normal multivariada.
BCt_p	: Distribuição Box-Cox t multivariada.
$LE\ell_p$: Distribuição log-elíptica.
LN_p	: Distribuição log-normal multivariada.
Lt_p	: Distribuição log- t multivariada.
f.d.p	: função densidade de probabilidade.
f.d.a	: função de distribuição acumulada.
f.g.d	: função geradora de densidades.
f.g.m	: função geradora de momentos.
f.c	: função característica.

Capítulo 1

Introdução

1.1 Formulação do problema

Dados positivos multivariados aparecem com frequência em diversas áreas de estudo. Consumos de nutrientes e medições antropométricas são exemplos desse tipo de dados que profissionais de saúde precisam modelar para avaliar diferentes características médicas (ver, por exemplo, Tooze et al. [1] e Pimentel et al. [2]). A distribuição normal multivariada é utilizada habitualmente para modelar dados multivariados, porém os modelos baseados nesta distribuição pressupõem simetria e apresentam problemas na presença de observações discrepantes, o que pode levar a conclusões erradas. Distribuições como a t multivariada, que tem caudas mais pesadas em relação à distribuição normal multivariada, são uma alternativa para modelar dados com presença de observações discrepantes (Lange et al. [3]), porém mantém o pressuposto de simetria.

Uma metodologia utilizada para reduzir a falta de normalidade em dados positivos multivariados é dada através da transformação de Box-Cox em cada componente do vetor de observações. Esta abordagem assume que o vetor de observações transformadas tem distribuição normal multivariada (elíptica), porém essa suposição apresenta uma desvantagem teórica, já que o domínio do vetor de variáveis transformadas não é necessariamente \mathbb{R}^p . Além disso, uma outra desvantagem deste modelo é que os parâmetros não são interpretáveis em termos de características do vetor de variáveis originais. No caso univariado, estes problemas são tratados por Ferrari e Fumes [4], que propuseram a classe de distribuições Box-Cox simétricas, que fornece distribuições úteis para modelar dados positivos, assimétricos e com presença de observações discrepantes. Esta classe tem como casos particulares a classe de distribuições log-simétricas (Vanegas e Paula [5]) e as distribuições Box-Cox Cole Green (Box-Cox normal, Cole e Green [6]), Box-Cox exponencial potência (Rigby e Stasinopoulos [7]) e Box-Cox t (Rigby e Stasinopoulos [8]). Essas autoras apresentam aplicações das distribuições Box-Cox simétricas por meio de ajustes ao consumo de micro e macronutrien-

tes, mostrando que essas distribuições apresentam melhor ajuste que outros modelos existentes na literatura. Caso o interesse seja analisar o consumo de vários nutrientes conjuntamente, as distribuições Box-Cox simétricas não são apropriadas, a menos que se suponha independência entre os consumos de nutrientes. Esta abordagem não é adequada uma vez que é comum a existência de associação entre os consumos de nutrientes. Desta maneira, é interessante considerar o problema de construir distribuições multivariadas que marginalmente tenham características similares a variáveis aleatórias com distribuições Box-Cox simétricas e que, além disso, a associação entre essas variáveis seja levada em consideração.

Neste trabalho abordamos o problema descrito acima construindo uma classe de distribuições que denominaremos de classe de distribuições Box-Cox elípticas. Essa construção é feita através de uma extensão da transformação de Box-Cox, e envolvendo uma nova classe de distribuições que chamaremos de classe de distribuições elípticas truncadas. A classe de distribuições Box-Cox elípticas tem como casos particulares as classes de distribuições log-elípticas (Fang et al. [9, Seção 2.8.1]) e Box-Cox simétricas. As desvantagens da abordagem clássica da transformação de Box-Cox multivariada são tratadas através das distribuições Box-Cox elípticas, sendo que o vetor de variáveis transformadas tem distribuição elíptica truncada com suporte num subconjunto adequado de \mathbb{R}^p . Também, alguns parâmetros que determinam as distribuições Box-Cox elípticas são interpretados como características das distribuições marginais do vetor de variáveis de interesse e outros parâmetros induzem associação entre essas distribuições. Essa interpretação dos parâmetros como características do vetor de variáveis originais permite modelar dados positivos multivariados, marginalmente assimétricos e com presença de observações discrepantes. Além disso, alguns parâmetros da classe de distribuições Box-Cox elípticas estão relacionados a quantis das distribuições marginais, tornando esta classe atrativa para modelagem de regressão.

1.2 Objetivos e estrutura da tese

O principal objetivo deste trabalho é definir, estudar e mostrar aplicações da classe de distribuições Box-Cox elípticas, que é uma alternativa para a modelagem de dados positivos multivariados através da transformação de Box-Cox multivariada. Para alcançar esse objetivo é essencial definir e estudar, previamente, a classe de distribuições elípticas truncadas. Neste trabalho também pretendemos definir e estudar alguns aspectos dos modelos de regressão lineares Box-Cox elípticos.

Este trabalho está organizado em cinco capítulos. No Capítulo 2 apresentamos uma abordagem teórica da classe das distribuições elípticas truncadas, necessária para definir e estudar a classe de distribuições Box-Cox elípticas apresentada no Capítulo 3. Também propomos um algoritmo para gerar amostras aleatórias de um vetor aleatório

com distribuição elíptica truncada sobre retângulos em \mathbb{R}^p , que é útil para gerar amostras aleatórias de um vetor aleatório com distribuição Box-Cox elíptica. No Capítulo 3 definimos a classe de distribuições Box-Cox elípticas, estudamos suas propriedades e a interpretação dos parâmetros que determinam essa classe. Também, apresentamos aspectos teóricos e computacionais associados ao método de máxima verossimilhança, que propomos para abordar o problema de estimação dos parâmetros. Entre os aspectos computacionais desenvolvemos estudos de simulação para verificar a adequação da metodologia de estimação proposta. Finalizamos este capítulo mostrando aplicações das distribuições Box-Cox elípticas a dados nutricionais e antropométricos. No Capítulo 4 definimos os modelos de regressão lineares Box-Cox elípticos, estudamos alguns aspectos teóricos e descrevemos o método de máxima verossimilhança aplicado a estes modelos. Propomos métodos de diagnóstico para avaliar ajustes dos modelos de regressão lineares log-normal e log- t multivariados. Finalizamos este capítulo mostrando aplicações dos modelos de regressão lineares Box-Cox elípticos a dados nutricionais e antropométricos. No Capítulo 5 apresentamos conclusões e propostas de estudos futuros derivados deste trabalho. Finalmente, incluímos dois apêndices, no Apêndice A apresentamos as provas dos teoremas e corolários estabelecidos neste trabalho. No Apêndice B disponibilizamos a programação em \mathbb{R} relacionada a alguns cálculos computacionais das famílias de distribuições Box-Cox normal e Box-Cox t multivariadas.

1.3 Suporte computacional

Os cálculos e gráficos deste trabalho foram feitos na linguagem de programação \mathbb{R} [10], versão 3.3.3, disponível em <http://www.R-project.org/>. Para os estudos de simulação desenvolvidos na Seção 3.6 utilizamos o servidor `BrucutuIV`, administrado pelo Instituto de Matemática e Estatística da Universidade de São Paulo. A escrita desta tese foi feita por meio do sistema tipográfico \LaTeX ; detalhes podem ser encontrados em <http://www.latex-project.org/>. Os programas computacionais foram executados sob o sistema operacional `Ubuntu`, versão 16.04; detalhes podem ser encontrados em <http://www.ubuntu.com/>.

Capítulo 2

A classe de distribuições elípticas truncadas

2.1 Introdução

Na estatística teórica e aplicada a distribuição normal multivariada é de grande importância devido a suas boas propriedades, em particular na análise multivariada e suas aplicações, incluindo modelos de regressão. Uma generalização da família de distribuições normais multivariadas é a classe de distribuições elípticas, proposta por Kelker [11]. Vários trabalhos sobre esta classe de distribuições têm sido desenvolvidos por diferentes autores nas últimas décadas. Fang et al. [9] apresentam um estudo unificado desses trabalhos. Além da família de distribuições normais multivariadas, outros membros na classe de distribuições elípticas são, por exemplo, as famílias de distribuições t , exponencial potência, slash, Bessel, logística, entre outras. Muitas propriedades satisfeitas pela distribuição normal multivariada também são satisfeitas pelas distribuições elípticas. Nos últimos anos, esta classe de distribuições tem sido de grande importância na análise multivariada e suas aplicações, ampliando as possibilidades na modelagem estatística sem ser limitada apenas à distribuição normal multivariada como ocorre tipicamente.

Neste capítulo definimos e estudamos uma nova classe de distribuições multivariadas que denominaremos de classe de distribuições elípticas truncadas. Esta classe será necessária para definir e estudar a classe de distribuições Box-Cox elípticas apresentada no Capítulo 3. A classe de distribuições elípticas truncadas generaliza a classe das distribuições elípticas através do truncamento num subconjunto de seu domínio, de modo que, quando este subconjunto seja o domínio completo, a distribuição elíptica será um caso particular. Esta classe tem como membros as distribuições normal e t multivariadas truncadas, as quais aparecem na literatura, mas estudos sobre estas distribuições raramente são discutidos. Birnbaum e Meyer [12] encontraram os momen-

tos centrais da distribuição normal padrão multivariada truncada inferiormente. Em uma série de artigos Tallis [13, 14, 15] obteve a função geradora de momentos (f.g.m) da distribuição normal truncada multivariada e apresenta algumas aplicações. Nesses artigos o autor considera truncamento inferior, elíptico, radial e através de planos. Recentemente, Horrace [16] estudou várias propriedades da distribuição normal multivariada truncada unilateralmente, mostrando resultados em relação a transformações lineares, distribuições marginais e condicionais, e independência. Manjunath e Wilhelm [17] tratam o problema de computar os primeiro e segundo momentos da distribuição normal multivariada considerando truncamento duplo, que tem como casos particulares os truncamentos inferiores e superiores, estendendo algumas das abordagens tratadas por Tallis [13, 14, 15]; estes autores também apresentam algoritmos para calcular os dois primeiros momentos e probabilidades das distribuições marginais bivariadas. Em relação à distribuição t multivariada truncada, Ho et. al [18] apresentaram fórmulas para calcular os dois primeiros momentos considerando truncamento duplo. Estes autores derivaram esses momentos com base numa representação estocástica de misturas de escala de distribuições normais multivariadas. Também desenvolveram um algoritmo eficiente baseado no método de *slice sampling* para gerar amostras aleatórias da distribuição t multivariada truncada. Geweke [19] propõe um algoritmo eficiente baseado no amostrador de Gibbs para a geração de amostras das distribuições normal e t multivariadas truncadas sob restrições lineares. Wilhelm e Manjunath [20] desenvolveram o pacote *tmvtnorm* do *software* livre \mathbb{R} [10] em que consideram vários aspectos de cálculo computacional em relação às distribuições normal e t multivariadas truncadas, entre os quais estão a geração de vetores aleatórios usando amostragem de Gibbs e *slice sampling*, cálculo de probabilidades, entre outros. Várias funções deste pacote estão baseadas no pacote *mvtnorm* (Genz et al. [21]) do *software* \mathbb{R} , onde são implementados vários algoritmos propostos por Genz e Bretz [22] para o cálculo computacional relacionado com as distribuições normal e t multivariadas.

2.2 A classe de distribuições elípticas

Nesta seção apresentamos alguns conceitos básicos sobre as distribuições elípticas que serão necessários para definir e estudar a classe de distribuições elípticas truncadas.

Denotamos por $\mathcal{O}(p)$ ao grupo ortogonal de matrizes de dimensão $p \times p$ com entradas em \mathbb{R} , isto é,

$$\mathcal{O}(p) = \{ \mathbf{H}(p \times p) : \mathbf{H}\mathbf{H}' = \mathbf{H}'\mathbf{H} = \mathbf{I}_p \}.$$

Definição 2.2.1. Um vetor aleatório $\mathbf{S} \in \mathbb{R}^p$ tem distribuição esférica, com suporte em \mathbb{R}^p , se

para cada matriz $\mathbf{H} \in \mathcal{O}(p)$ a função característica (f.c) de \mathbf{S} satisfaz

$$\phi_{\mathbf{S}}(\mathbf{t}) = \phi_{\mathbf{S}}(\mathbf{H}\mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^p. \quad (2.1)$$

A relação (2.1) implica que a f.c de \mathbf{S} é da forma

$$\phi_{\mathbf{S}}(\mathbf{t}) = \phi(\mathbf{t}'\mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^p,$$

para alguma função ϕ tal que $\phi(u) \in \mathbb{R}$, para todo $u \geq 0$. Se \mathbf{S} tem função densidade de probabilidade (f.d.p) com respeito à medida de Lebesgue, então esta é da forma

$$f_{\mathbf{S}}(\mathbf{s}) = g(\mathbf{s}'\mathbf{s}), \quad \mathbf{s} \in \mathbb{R}^p,$$

para alguma função g tal que $g(u) \geq 0$, para todo $u \geq 0$, e $\int_0^\infty t^{p-1}g(t^2) dt < \infty$; e escrevemos $\mathbf{S} \sim S_p(g)$.

Uma generalização das distribuições esféricas é dada através das distribuições elípticas, apresentadas na Definição 2.2.2. Se $\mathbf{A}(p \times p)$ é uma matriz simétrica, as notações $\mathbf{A} \geq 0$ e $\mathbf{A} > 0$ indicam que \mathbf{A} é definida não-negativa e definida positiva, respectivamente.

Definição 2.2.2. O vetor aleatório $\mathbf{X} \in \mathbb{R}^p$ tem distribuição elíptica com suporte em \mathbb{R}^p , vetor de locação $\boldsymbol{\mu} \in \mathbb{R}^p$ e matriz de dispersão $\boldsymbol{\Sigma}(p \times p) \geq 0$, se a f.c de \mathbf{X} é da forma

$$\phi_{\mathbf{X}}(\mathbf{t}) = \exp(it'\boldsymbol{\mu})\phi(\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}), \quad (2.2)$$

para alguma função ϕ tal que $\phi(u) \in \mathbb{R}$, para todo $u \geq 0$.

A distribuição elíptica pode ser construída a partir da distribuição esférica considerando $\mathbf{S} \in \mathbb{R}^q$ com distribuição esférica e f.c $\phi(\mathbf{s}'\mathbf{s})$, $\mathbf{s} \in \mathbb{R}^q$, e definindo $\mathbf{X} \in \mathbb{R}^p$ através da transformação

$$\mathbf{X} = \boldsymbol{\mu} + \mathbf{A}'\mathbf{S},$$

em que $\boldsymbol{\mu} \in \mathbb{R}^p$ e $\mathbf{A}(q \times p)$ é tal que $\mathbf{A}'\mathbf{A} = \boldsymbol{\Sigma}$ com posto $r(\boldsymbol{\Sigma}) = q$. Com isso é possível mostrar que a f.c de \mathbf{X} tem a forma (2.2). Se a distribuição de \mathbf{X} tem f.d.p com respeito à medida de Lebesgue em \mathbb{R}^p , então $\boldsymbol{\Sigma} > 0$ e a f.d.p é da forma

$$f_{\mathbf{X}}(\mathbf{x}) = \det(\boldsymbol{\Sigma})^{-1/2}g((\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})), \quad \mathbf{x} \in \mathbb{R}^p, \quad (2.3)$$

para alguma função g tal que $g(u) \geq 0$, para todo $u \geq 0$, e $\int_0^\infty t^{p-1}g(t^2) dt < \infty$. Neste caso escrevemos $\mathbf{X} \sim El_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$. Note que a f.d.p (2.3) pode ser obtida usando a transformação

$$\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\mathbf{S}, \quad (2.4)$$

onde $S \sim S_p(g)$. A função g denomina-se função geradora de densidades (f.g.d) e cada membro da classe de distribuições elípticas é determinado pela forma dessa função. A seguir, estão listados alguns membros pertencentes à classe de distribuições elípticas segundo a forma da f.g.d g .

1. Distribuição normal multivariada (Gupta et al. [23]):

$$g(u) = (2\pi)^{-p/2} \exp\left(-\frac{u}{2}\right). \quad (2.5)$$

2. Distribuição de Kotz multivariada (Gupta et al. [23]):

$$g(u) = \frac{sr^{(2q+p-2)/2s}\Gamma(p/2)}{\pi^{p/2}\Gamma((2q+p-2)/2s)} u^{q-1} \exp(-ru^s), \quad (2.6)$$

em que $r > 0$, $s > 0$, $2q + p > 2$ e $\Gamma(z)$ é a função gama definida como

$$\Gamma(z) = \int_0^\infty t^{z-1} \exp(-t) dt, \quad \operatorname{Re}(z) > 0.$$

3. Distribuição de Cauchy multivariada (Gupta et al. [23]):

$$g(u) = \Gamma\left(\frac{p+1}{2}\right) [\pi(1+u)]^{-\frac{p+1}{2}}. \quad (2.7)$$

4. Distribuição t multivariada (Gupta et al. [23]):

$$g(u) = \frac{\Gamma((\tau+p)/2)}{(\tau\pi)^{p/2}\Gamma(\tau/2)} \left(1 + \frac{u}{\tau}\right)^{-\frac{\tau+p}{2}}, \quad (2.8)$$

em que $\tau > 0$.

5. Distribuição exponencial potência multivariada (Gómez et al. [24]):

$$g(u) = \frac{p\Gamma(p/2)}{\pi^{p/2}2^{1+p/(2\beta)}\Gamma(1+p/(2\beta))} \exp\left(-\frac{1}{2}u^\beta\right), \quad (2.9)$$

em que $\beta > 0$.

6. Distribuição de Pearson tipo II multivariada (Gupta et al. [23]):

$$g(u) = \frac{\Gamma(p/2 + q + 1)}{\pi^{p/2}\Gamma(q + 1)} (1 - u)^q, \quad 0 \leq u \leq 1,$$

em que $q > -1$.

7. Distribuição de Pearson tipo VII multivariada (Gupta et al. [23]):

$$g(u) = \frac{\Gamma(\alpha)}{(\nu\pi)^{p/2}\Gamma(\alpha - p/2)} \left(1 + \frac{u}{\nu}\right)^{-\alpha}, \quad (2.10)$$

em que $\nu > 0$ e $\alpha > p/2$.

8. Distribuição logística multivariada (Gupta et al. [23]):

$$g(u) = \left\{ \frac{\pi^{p/2}}{\Gamma(p/2)} \int_0^\infty \frac{t^{p/2-1} \exp(-t)}{[1 + \exp(-t)]^2} dt \right\} \frac{\exp(-u)}{[1 + \exp(-u)]^2}.$$

9. Distribuição de Bessel multivariada (Gupta et al. [23]):

$$g(u) = [2^{p+q-1} \pi^{p/2} r^{p+q} \Gamma(p/2 + q)]^{-1} u^{q/2} K_q \left(\frac{\sqrt{u}}{r} \right),$$

em que $r > 0$, $q > -p/2$ e $K_q(z)$ é a função de Bessel modificada de terceiro tipo dada por

$$K_q(z) = \frac{\pi[I_{-q}(z) - I_q(z)]}{2 \sin(q\pi)}, \quad |\arg(z)| < \pi,$$

sendo

$$I_q(z) = \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(q + k + 1)} \left(\frac{z}{2}\right)^{q+2k}, \quad |z| < \infty, \quad |\arg(z)| < \pi.$$

10. Distribuição slash multivariada (Wang e Genton [25]):

$$g(u) = \begin{cases} \frac{q 2^{q/2-1}}{\pi^{p/2}} u^{-(p+q)/2} \gamma\left(\frac{p+q}{2}, \frac{u}{2}\right), & u > 0, \\ \frac{q}{(2\pi)^{p/2}(p+q)}, & u = 0, \end{cases}$$

em que $q > 0$ e γ é a função gama incompleta dada por

$$\gamma(a, z) = \int_0^z t^{a-1} \exp(-t) dt, \quad \operatorname{Re}(a) > 0, \quad \operatorname{Re}(z) > 0.$$

11. Distribuições de mistura de escala normais (Gupta et al. [23]):

$$g(u) = (2\pi)^{-p/2} \int_0^\infty t^{-p/2} \exp\left(-\frac{u}{2t}\right) dG(t), \quad (2.11)$$

sendo G uma função de distribuição acumulada (f.d.a) definida em $(0, \infty)$.

Seja $\mathbf{X} \sim \text{El}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$. Se a f.g.d g tem a forma (2.5) dizemos que \mathbf{X} tem distribuição normal multivariada com parâmetros $\boldsymbol{\mu} \in \mathbb{R}^p$ e $\boldsymbol{\Sigma}(p \times p) > 0$, e escrevemos $\mathbf{X} \sim$

$N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Se a f.g.d g tem a forma (2.8) dizemos que \mathbf{X} tem distribuição t multivariada com parâmetros $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $\tau > 0$ graus de liberdade, e escrevemos $\mathbf{X} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \tau)$.

Uma variável aleatória U tem distribuição uniforme no intervalo (a, b) , denotada por $U \sim U(a, b)$, se sua f.d.p é dada por

$$f_U(u) = \frac{1}{b-a}, \quad u \in (a, b).$$

Uma variável aleatória $X > 0$ tem distribuição gama inversa com parâmetros $\alpha > 0$ e $\beta > 0$, denotada por $X \sim \text{IG}(\alpha, \beta)$, se sua f.d.p é dada por

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp\left(-\frac{\beta}{x}\right), \quad x > 0. \quad (2.12)$$

Observação 2.2.1. *As distribuições elípticas são determinadas pela forma da f.g.d g . Com base nisso é importante notar que:*

1. *A classe de distribuições elípticas fornece várias alternativas para modelagem estatística, incluindo distribuições de caudas pesadas e leves.*
2. *A f.g.d g pode apresentar parâmetros extras na f.d.p (2.3). Por exemplo, para a distribuição t multivariada tem-se como parâmetro extra o número de graus de liberdade $\tau > 0$, que controla o peso da cauda da distribuição.*
3. *A distribuição normal multivariada é um caso particular das distribuições de Kotz e exponencial potência multivariadas, pois (2.5) é obtida de (2.6) e (2.9) considerando $q = s = 2r = 1$ e $\beta = 1$, respectivamente.*
4. *A distribuição t multivariada é uma mistura de escala normais (2.11), já que (2.8) pode ser expressa como*

$$g(u) = (2\pi)^{-p/2} \int_0^\infty t^{-p/2} \exp\left(-\frac{u}{2t}\right) dG(t), \quad (2.13)$$

em que G é a f.d.a de uma variável aleatória $T \sim \text{IG}(\tau/2, \tau/2)$. A distribuição normal multivariada é um caso limite da distribuição t multivariada quando $\tau \rightarrow \infty$.

5. *A distribuição t multivariada é um caso particular da distribuição de Pearson tipo VII multivariada; basta considerar $\alpha = (\tau + p)/2$ em (2.10) para obter (2.8).*
6. *A distribuição de Cauchy multivariada é um caso particular da distribuição t multivariada, pois (2.7) é obtido de (2.8) tomando $\tau = 1$.*

7. A distribuição slash multivariada também é uma mistura de escala normais (2.11), uma vez que (10) pode ser expressa como

$$g(u) = (2\pi)^{-p/2} \int_0^1 t^{p/q} \exp\left(-\frac{u}{2t^{-2/q}}\right) dG(t), \quad (2.14)$$

sendo G a f.d.a de uma variável aleatória $T \sim U(0, 1)$. A distribuição normal multivariada é um caso limite da distribuição slash multivariada quando $q \rightarrow \infty$.

Ao longo deste texto, quando for necessário apresentaremos alguns resultados sobre as distribuições elípticas.

2.3 A classe de distribuições elípticas truncadas

Nesta seção propomos uma classe geral de distribuições que chamaremos de classe de distribuições elípticas truncadas. Esta classe será necessária para definir e estudar a classe de distribuições Box-Cox elípticas apresentada no Capítulo 3. A classe de distribuições elípticas truncadas tem como membros particulares as famílias de distribuições normal e t multivariadas truncadas, já existentes na literatura. Primeiramente, apresentamos algumas propriedades considerando conjuntos arbitrários mensuráveis em \mathbb{R}^p que determinem as regiões de truncamento, depois derivamos outras propriedades quando os conjuntos de truncamento são retângulos em \mathbb{R}^p . Finalmente, propomos um algoritmo para gerar amostras aleatórias das distribuições elípticas truncadas.

A distribuição esférica truncada provém do truncamento de um vetor aleatório que tem distribuição esférica com função geradora de densidades g sobre um subconjunto mensurável $B \subseteq \mathbb{R}^p$, como é definida a seguir.

Definição 2.3.1. *Seja $B \subseteq \mathbb{R}^p$ mensurável. Dizemos que o vetor aleatório $\mathbf{S} \in \mathbb{R}^p$ tem distribuição esférica truncada com suporte no conjunto B e f.g.d g , escrevemos $\mathbf{S} \sim TS_p(B; g)$, se sua f.d.p é da forma*

$$f_{\mathbf{S}}(\mathbf{s}) = \frac{g(\mathbf{s}'\mathbf{s})}{\int_B g(\mathbf{s}'\mathbf{s}) d\mathbf{s}}, \quad \mathbf{s} \in B, \quad (2.15)$$

em que g é uma função tal que $g(u) \geq 0$, para todo $u \geq 0$, e $\int_0^\infty t^{p-1} g(t^2) dt < \infty$.

A f.d.p (2.15) também pode ser expressa como

$$f_{\mathbf{S}}(\mathbf{s}) = \frac{f_{\mathbf{Z}}(\mathbf{s})}{P(\mathbf{Z} \in B)}, \quad \mathbf{s} \in B, \quad (2.16)$$

onde $f_{\mathbf{Z}}$ é a f.d.p de um vetor aleatório $\mathbf{Z} \sim S_p(g)$.

O caso $p = 1$ corresponde a uma variável aleatória S com distribuição simétrica padrão truncada com suporte no subconjunto mensurável $B \subseteq \mathbb{R}$ e f.g.d g , e escrevemos

$S \sim \text{TS}_1(B; g)$. Neste caso, se $B = (a, b)$, então a f.d.p de S , segundo (2.16), é dada por

$$f_S(s) = \frac{f_Z(s)}{F_Z(b) - F_Z(a)}, \quad s \in (a, b), \quad (2.17)$$

em que f_Z e F_Z são a f.d.p e a f.d.a de uma variável aleatória $Z \sim S_1(g)$, respectivamente.

Se consideramos $B = \mathbb{R}^p$ em (2.16), temos que $P(\mathbf{Z} \in B) = 1$, e portanto $f_S(\mathbf{s}) = f_Z(\mathbf{z})$, com $\mathbf{Z} \sim S_p(g)$. Assim, a classe de distribuições esféricas truncadas é uma generalização da classe de distribuições esféricas.

Se $\mathbf{S} \sim \text{TS}_p(B; g)$, então a f.c de \mathbf{S} é

$$\phi_{\mathbf{S}}(\mathbf{t}) = \frac{\int_B \exp(i\mathbf{t}'\mathbf{s})g(\mathbf{s}'\mathbf{s}) d\mathbf{s}}{\int_B g(\mathbf{s}'\mathbf{s}) d\mathbf{s}}, \quad \mathbf{t} \in \mathbb{R}^p. \quad (2.18)$$

As f.c das distribuições esféricas são invariantes sob transformações ortogonais como estabelece (2.1), mas esta propriedade é perdida quando é considerado um truncamento sobre um subconjunto próprio de \mathbb{R}^p . No teorema a seguir mostramos que a f.d.p (2.15) não é invariante sob transformações ortogonais, mas é igual à f.d.p da distribuição esférica truncada sobre conjuntos obtidos por rotações do conjunto de truncamento; isto é devido à simetria das distribuições esféricas. Do mesmo modo, mostramos que a f.c (2.18) não é invariante sob transformações ortogonais, mas é igual à f.c da variável aleatória esférica truncada sobre conjuntos obtidos por rotações do conjunto de truncamento.

Teorema 2.3.1. *Sejam $\mathbf{S} \sim \text{TS}_p(B; g)$ e $T : \mathbb{R}^p \rightarrow \mathbb{R}^p$ a transformação ortogonal $T(\mathbf{x}) = \mathbf{H}\mathbf{x}$, em que $\mathbf{H} \in \mathcal{O}(p)$. Seja ainda $\mathbf{S}^* \sim \text{TS}_p(T(B); g)$, então*

1. a f.d.p de \mathbf{S} satisfaz a relação $f_{\mathbf{S}}(\mathbf{s}) = f_{\mathbf{S}^*}(T(\mathbf{s}))$, $\mathbf{s} \in B$,
2. a f.c de \mathbf{S} satisfaz a relação $\phi_{\mathbf{S}}(\mathbf{t}) = \phi_{\mathbf{S}^*}(T(\mathbf{t}))$, $\mathbf{t} \in \mathbb{R}^p$.

Demonstração. Ver Apêndice A.1. ■

Na Definição 2.3.2 propomos a classe de distribuições elípticas truncadas, uma generalização da classe de distribuições esféricas truncadas e da classe de distribuições elípticas.

Definição 2.3.2. *Seja $B \subseteq \mathbb{R}^p$ mensurável. O vetor aleatório $\mathbf{X} \in \mathbb{R}^p$ tem distribuição elíptica truncada com suporte no conjunto B , parâmetros $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e f.g.d g , escrevemos $\mathbf{X} \sim \text{TEl}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; B; g)$, se sua f.d.p é dada por*

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{g((\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}))}{\int_B g((\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})) d\mathbf{x}}, \quad \mathbf{x} \in B, \quad (2.19)$$

em que g é uma função tal que $g(u) \geq 0$, para todo $u \geq 0$, e $\int_0^\infty t^{p-1}g(t^2) dt < \infty$.

No caso $p = 1$ dizemos que a variável aleatória X tem distribuição simétrica truncada com suporte no conjunto $B \subseteq \mathbb{R}$, parâmetros $\mu \in \mathbb{R}$, $\sigma^2 > 0$ e f.g.d g , e escrevemos $X \sim \text{TE}\ell_1(\mu, \sigma^2; B; g)$.

Note que a f.d.p (2.19) também pode ser expressa como

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{f_{\mathbf{W}}(\mathbf{x})}{\text{P}(\mathbf{W} \in B)}, \quad \mathbf{x} \in B, \quad (2.20)$$

em que $f_{\mathbf{W}}$ é a f.d.p de um vetor aleatório $\mathbf{W} \sim \text{El}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$. Se $\boldsymbol{\mu} = \mathbf{0}$ e $\boldsymbol{\Sigma} = \mathbf{I}_p$, então a f.d.p (2.19) coincide com a f.d.p (2.15), ou seja, a classe de distribuições esféricas truncadas é um caso particular da classe de distribuições elípticas truncadas. Também, se consideramos $B = \mathbb{R}^p$ em (2.20), temos que $\text{P}(\mathbf{W} \in B) = 1$ e, portanto, $f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{W}}(\mathbf{x})$. Por isto, a classe de distribuições elípticas truncadas é uma generalização da classe de distribuições elípticas.

As f.g.d apresentadas na Seção 2.2 fornecem vários membros pertencentes à classe de distribuições elípticas truncadas segundo a forma da f.g.d g . É importante notar que as f.d.p estão determinadas pelos núcleos das f.d.p elípticas. Além disso, as propriedades mencionadas na Observação 2.2.1 também são satisfeitas na classe de distribuições elípticas truncadas, exceto as representações de mistura de escala normais (2.13) e (2.14).

Dois casos que serão de grande interesse neste trabalho ocorrem quando g assume as formas (2.5) e (2.8). No primeiro caso dizemos que o vetor aleatório $\mathbf{X} \in \mathbb{R}^p$ tem distribuição normal multivariada truncada com suporte no conjunto $B \subseteq \mathbb{R}^p$ e parâmetros $\boldsymbol{\mu} \in \mathbb{R}^p$ e $\boldsymbol{\Sigma}(p \times p) > 0$, escrevemos $\mathbf{X} \sim \text{TN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; B)$, com f.d.p

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)}{\int_B \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) d\mathbf{x}}, \quad \mathbf{x} \in B. \quad (2.21)$$

No segundo caso dizemos que o vetor aleatório $\mathbf{X} \in \mathbb{R}^p$ tem distribuição t multivariada truncada com suporte no conjunto $B \subseteq \mathbb{R}^p$ e parâmetros $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $\tau > 0$ graus de liberdade, escrevemos $\mathbf{X} \sim \text{T}t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \tau; B)$, com f.d.p

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\left(1 + \frac{1}{\tau}(\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)^{-\frac{\tau+p}{2}}}{\int_B \left(1 + \frac{1}{\tau}(\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)^{-\frac{\tau+p}{2}} d\mathbf{x}}, \quad \mathbf{x} \in B. \quad (2.22)$$

Na Figura 2.1 mostramos curvas de nível e gráficos da f.d.p de um vetor aleatório $\mathbf{X} \sim \text{T}t_2(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \tau; B)$ para diferentes parâmetros e conjuntos de truncamento retangu-

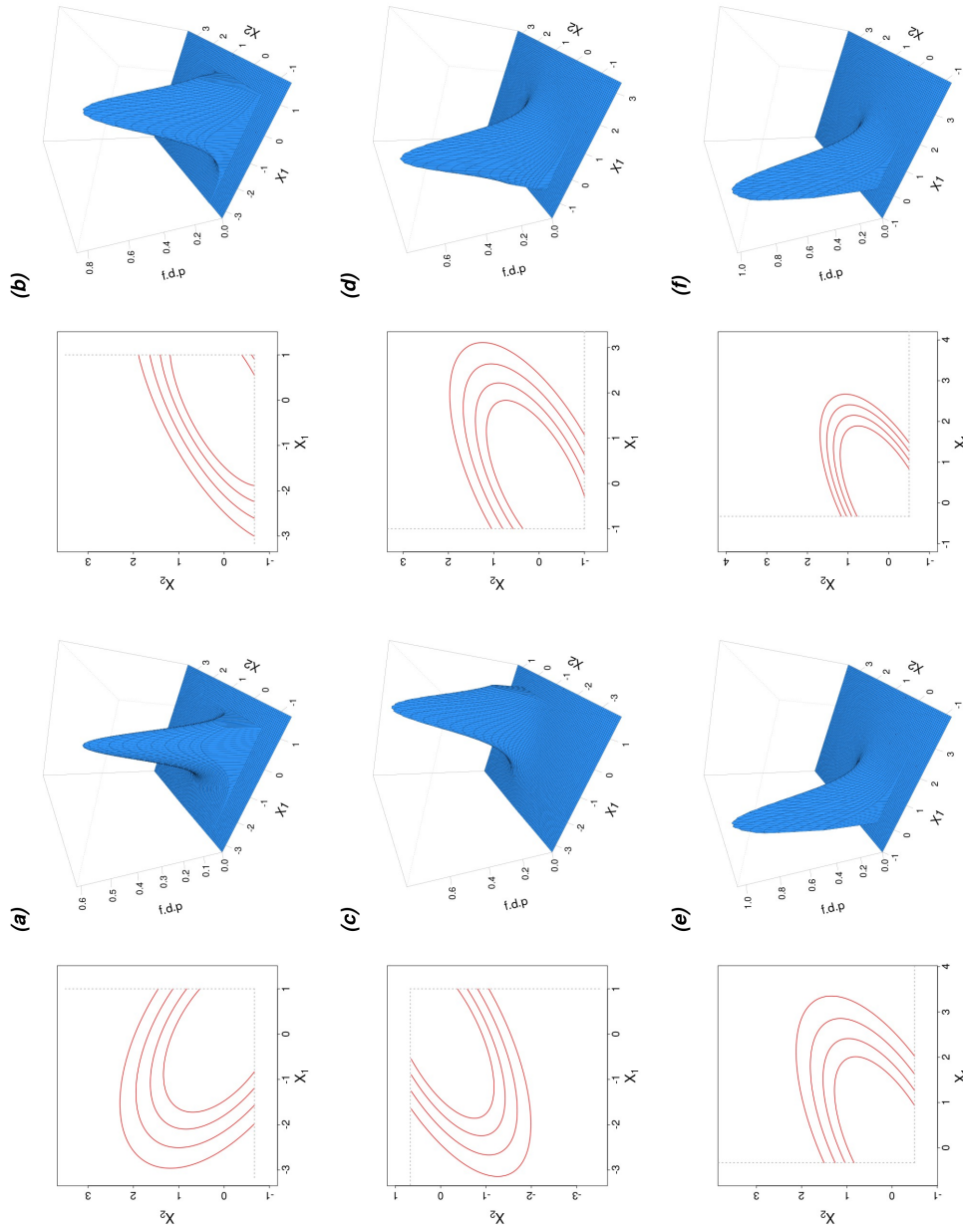


Figura 2.1: Curvas dos núcleos 0,04, 0,02, 0,01, 0,005 e gráficos da f.d.p de $\mathbf{X} \sim Tt_2(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \tau; B)$, em que $\mu_1 = \mu_2 = 0, \sigma_{11} = 0,5$ e (a) $\sigma_{22} = 0,3, \sigma_{12} = -0,2, \tau = 3, B = (-\infty, 1) \times (-2/3, \infty)$, (b) $\sigma_{22} = 0,2, \sigma_{12} = 0,2, \tau = 3, B = (-\infty, 1) \times (-\infty, 2/3)$, (c) $\sigma_{22} = 0,2, \sigma_{12} = 0,2, \tau = 3, B = (-1, \infty) \times (-1/3, \infty)$, (d) $\sigma_{22} = 0,2, \sigma_{12} = 0,2, \tau = 3, B = (-1, \infty) \times (-1/2, \infty)$, (e) $\sigma_{22} = 0,2, \sigma_{12} = 0,2, \tau = 10, B = (-1/3, \infty) \times (-1/2, \infty)$.

lares. Note que em todas as figuras as curvas de nível são elipses projetadas sobre o conjunto de truncamento B . À medida que os graus de liberdade aumentam, as curvas de nível da distribuição t truncada bivariada tendem às curvas de nível da distribuição normal truncada bivariada (Figuras 2.1(e) e 2.1(f)).

2.4 Propriedades das distribuições elípticas truncadas

Nesta seção apresentamos várias propriedades das distribuições elípticas truncadas. Inicialmente mostramos vários resultados distribucionais envolvendo transformações lineares de variáveis esféricas e elípticas truncadas. Depois apresentamos um teorema que fornece condições para que uma variável com distribuição elíptica truncada tenha f.d.p log-côncava. Finalmente, mostramos resultados distribucionais envolvendo distribuições marginal e condicional, e independência.

No Teorema 2.4.1 apresentamos uma caracterização das distribuições elípticas truncadas através das distribuições esféricas truncadas. Este teorema generaliza a caracterização das distribuições elípticas através da transformação (2.4).

Teorema 2.4.1. *Sejam $B \subseteq \mathbb{R}^p$ mensurável, $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $T : \mathbb{R}^p \rightarrow \mathbb{R}^p$ a transformação linear dada por $T(\mathbf{x}) = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\mathbf{x}$. Então, $\mathbf{S} \sim TS_p(B; g)$ se, e somente se, $T(\mathbf{S}) \sim TEL_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; T(B); g)$.*

Demonstração. Ver Apêndice A.2. ■

No Corolário 2.4.1.1 apresentamos formas fechadas para a f.d.p e a f.d.a de uma variável aleatória simétrica truncada sobre um intervalo da forma (a, b) . Estas funções dependem da f.d.p e a f.d.a de uma variável aleatória simétrica padrão.

Corolário 2.4.1.1. *Se $X \sim TEL_1(\mu, \sigma^2; (a, b); g)$, então a f.d.p de X é dada por*

$$f_X(x) = \frac{\frac{1}{\sigma} f_Z\left(\frac{x - \mu}{\sigma}\right)}{F_Z\left(\frac{b - \mu}{\sigma}\right) - F_Z\left(\frac{a - \mu}{\sigma}\right)}, \quad x \in (a, b),$$

e a f.d.a de X é

$$F_X(x) = \frac{F_Z\left(\frac{x - \mu}{\sigma}\right) - F_Z\left(\frac{a - \mu}{\sigma}\right)}{F_Z\left(\frac{b - \mu}{\sigma}\right) - F_Z\left(\frac{a - \mu}{\sigma}\right)}, \quad x \in (a, b), \quad (2.23)$$

em que f_Z e F_Z são a f.d.p e a f.d.a de uma variável aleatória $Z \sim S_1(g)$, respectivamente.

Demonstração. Ver Apêndice A.3. ■

Note que o Corolário 2.4.1.1 também é válido quando $a \rightarrow -\infty$ ou $b \rightarrow \infty$; nesses casos temos que

$$\lim_{a \rightarrow -\infty} F_Z\left(\frac{a - \mu}{\sigma}\right) = 0 \quad \text{ou} \quad \lim_{b \rightarrow \infty} F_Z\left(\frac{b - \mu}{\sigma}\right) = 1.$$

A f.d.a dada em (2.23) pode ser usada para gerar amostras aleatórias da variável aleatória $X \sim \text{TEl}_1(\mu, \sigma^2; (a, b); g)$. Através do método de transformação inversa, basta gerar um número aleatório u de uma variável $U \sim U(0, 1)$ e calcular

$$x = \mu + \sigma F_Z^{-1}\left[u \left\{ F_Z\left(\frac{b - \mu}{\sigma}\right) - F_Z\left(\frac{a - \mu}{\sigma}\right) \right\} + F_Z\left(\frac{a - \mu}{\sigma}\right)\right]. \quad (2.24)$$

A expressão (2.24) será útil na geração de amostras aleatórias de um vetor aleatório apresentando distribuição elíptica truncada sobre retângulos em \mathbb{R}^p (Seção 2.5).

No Corolário 2.4.1.2 apresentamos uma relação entre as f.c das distribuições elípticas truncadas e as f.c das distribuições esféricas truncadas. Esta relação é baseada numa transformação linear do conjunto de truncamento e generaliza (2.2) quando $\Sigma(p \times p) > 0$.

Corolário 2.4.1.2. *Sejam $B \subseteq \mathbb{R}^p$ mensurável, $\boldsymbol{\mu} \in \mathbb{R}^p$, $\Sigma(p \times p) > 0$ e $T : \mathbb{R}^p \rightarrow \mathbb{R}^p$ a transformação linear dada por $T(\mathbf{x}) = \boldsymbol{\mu} + \Sigma^{1/2}\mathbf{x}$. Se $\mathbf{X} \sim \text{TEl}_p(\boldsymbol{\mu}, \Sigma; B; g)$, então a f.c de \mathbf{X} é dada por*

$$\phi_{\mathbf{X}}(\mathbf{t}) = \exp(i\mathbf{t}'\boldsymbol{\mu})\phi_{\mathbf{S}}(\Sigma^{1/2}\mathbf{t}),$$

em que $\phi_{\mathbf{S}}$ é a f.c de $\mathbf{S} \sim \text{TS}_p(T^{-1}(B); g)$.

Demonstração. Ver Apêndice A.4. ■

No Teorema 2.4.2 mostramos que transformações afins de um vetor aleatório com distribuição elíptica truncada também apresenta distribuição elíptica truncada. No caso de um vetor aleatório com distribuição elíptica, as transformações afins da variável são fechadas sob a família de distribuições, porém o vetor de locação e matriz de dispersão são alterados pela transformação. No caso de um vetor aleatório com distribuição elíptica truncada acontecerá o mesmo, mas o conjunto de truncamento também será alterado pela transformação afim.

Teorema 2.4.2. *Sejam $B \subseteq \mathbb{R}^p$ mensurável e $T : \mathbb{R}^p \rightarrow \mathbb{R}^p$ a transformação linear dada por $T(\mathbf{x}) = \mathbf{D}\mathbf{x} + \mathbf{b}$, em que $\mathbf{b} \in \mathbb{R}^p$ e $\mathbf{D}(p \times p)$ é uma matriz tal que $\det(\mathbf{D}) \neq 0$. Se $\mathbf{X} \sim \text{TEl}_p(\boldsymbol{\mu}, \Sigma; B; g)$, então $T(\mathbf{X}) \sim \text{TEl}_p(\mathbf{D}\boldsymbol{\mu} + \mathbf{b}, \mathbf{D}\Sigma\mathbf{D}'; T(B); g)$.*

Demonstração. Ver Apêndice A.5. ■

Uma f.d.p $f : \mathbb{R}^p \rightarrow [0, \infty)$ é log-côncava se a desigualdade

$$f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) \geq [f(\mathbf{x})]^\alpha [f(\mathbf{y})]^{1-\alpha} \quad (2.25)$$

é satisfeita para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$ e para todo $\alpha \in [0, 1]$. Se $\mathbf{X} \in \mathbb{R}^p$ é um vetor aleatório com f.d.p $f_{\mathbf{X}}(\mathbf{x})$ log-côncava, então os conjuntos de nível de $f_{\mathbf{X}}$ são conjuntos convexos fechados, as f.d.p marginais das distribuições log-côncavas são log-côncavas, o produto das f.d.p log-côncavas são log-côncavas. Além disso, $f_{\mathbf{X}}(\mathbf{x})$ é A -unimodal, isto é, o conjunto $\{\mathbf{x} \in \mathbb{R}^p : f_{\mathbf{X}}(\mathbf{x}) \geq \lambda\}$ é convexo para todo $\lambda > 0$. Alguns estudos sobre estes conceitos aparecem em Tong [26].

No Teorema 2.4.3 estabelecemos condições para que um vetor com distribuição elíptica truncada tenha f.d.p log-côncava.

Teorema 2.4.3. *Sejam $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $B \subseteq \mathbb{R}^p$ um conjunto convexo. Seja também $\mathbf{X} \sim \text{TEl}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; B; g)$ com f.d.p dada em (2.20). Se a f.d.p $f_{\mathbf{W}}$ de \mathbf{W} é log-côncava, então a f.d.p $f_{\mathbf{X}}$ de \mathbf{X} também é log-côncava.*

Demonstração. Ver Apêndice A.6. ■

Seja $B \subseteq \mathbb{R}^p$ um conjunto convexo. A f.d.p $f_{\mathbf{X}}$ de $\mathbf{X} \sim \text{TN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; B)$ é log-côncava, já que a f.d.p $f_{\mathbf{W}}$ de $\mathbf{W} \sim \text{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ é log-côncava. Se $\mathbf{X} \sim \text{Tt}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \tau; B)$, não é possível concluir a partir do Teorema 2.4.3 que a f.d.p $f_{\mathbf{X}}$ seja log-côncava, pois a f.d.p $f_{\mathbf{W}}$ de $\mathbf{W} \sim \text{t}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \tau)$ não é log-côncava, a menos que $\tau \rightarrow \infty$. Embora, é possível mostrar diretamente que $f_{\mathbf{X}}$ não satisfaz (2.25) e portanto não é log-côncava, a menos que $\tau \rightarrow \infty$.

Prékopa [27] mostrou que se o vetor aleatório $\mathbf{X} \in \mathbb{R}^p$ tem f.d.p log-côncava, então

$$P(\mathbf{X} \in \alpha A + (1 - \alpha)B) \geq \{P[\mathbf{X} \in A]\}^\alpha \{P[\mathbf{X} \in B]\}^{1-\alpha}, \quad (2.26)$$

para todo $\alpha \in [0, 1]$, em que A e B são subconjuntos de \mathbb{R}^p e

$$\alpha A + (1 - \alpha)B = \{\mathbf{z} \in \mathbb{R}^p : \mathbf{z} = \alpha \mathbf{x} + (1 - \alpha)\mathbf{y}; \mathbf{x} \in A, \mathbf{y} \in B\}.$$

Sejam A_1, \dots, A_m subconjuntos em \mathbb{R}^p e $\alpha_1, \dots, \alpha_m$ números reais tais que $\alpha_i \geq 0$, $i = 1, \dots, m$ e $\sum_{i=1}^m \alpha_i = 1$. Definimos o conjunto $\sum_{i=1}^m \alpha_i A_i$ como

$$\sum_{i=1}^m \alpha_i A_i = \left\{ \mathbf{z} \in \mathbb{R}^p : \mathbf{z} = \sum_{i=1}^m \alpha_i \mathbf{x}_i, \mathbf{x}_i \in A_i, i = 1, \dots, m \right\}.$$

Teorema 2.4.4. *Sejam $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $B \subseteq \mathbb{R}^p$ um conjunto convexo. Se $\mathbf{X} \sim \text{TEl}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; B; g)$ tem f.d.p log-côncava, então*

$$P\left[\mathbf{X} \in \sum_{i=1}^m \alpha_i A_i\right] \geq \prod_{i=1}^m \{P(\mathbf{X} \in A_i)\}^{\alpha_i}$$

Demonstração. Ver Apêndice A.7. ■

O Teorema 2.4.4 nos permite encontrar limitantes para probabilidades envolvendo um vetor aleatório com distribuição elíptica truncada.

Para estabelecer resultados em relação às distribuições marginais e condicionais, e independência das distribuições elípticas truncadas, consideremos partições de $\mathbf{X} \in \mathbb{R}^p$, $\boldsymbol{\mu} \in \mathbb{R}^p$ e $\boldsymbol{\Sigma}(p \times p) > 0$ como

$$\mathbf{X}' = (\mathbf{X}'_1, \mathbf{X}'_2), \quad \boldsymbol{\mu}' = (\boldsymbol{\mu}'_1, \boldsymbol{\mu}'_2), \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}, \quad (2.27)$$

em que $\mathbf{X}_1 \in \mathbb{R}^r$, $\mathbf{X}_2 \in \mathbb{R}^{p-r}$, $\boldsymbol{\mu}_1 \in \mathbb{R}^r$, $\boldsymbol{\mu}_2 \in \mathbb{R}^{p-r}$, $\boldsymbol{\Sigma}_{11}(r \times r) > 0$, $\boldsymbol{\Sigma}_{22}((p-r) \times (p-r)) > 0$ e $\boldsymbol{\Sigma}_{12}(r \times (p-r))$ é tal que $\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}'_{12}$. Consideramos também como conjunto de truncamento o retângulo $R \subseteq \mathbb{R}^p$ dado por $R = I_1 \times \cdots \times I_p$, em que I_1, \dots, I_p são intervalos em \mathbb{R} (finitos ou infinitos). Note que R pode ser expresso como o produto cartesiano dos retângulos $R_1 \subseteq \mathbb{R}^r$ e $R_2 \subseteq \mathbb{R}^{p-r}$, isto é

$$R = R_1 \times R_2, \quad (2.28)$$

sendo $R_1 = I_1 \times \cdots \times I_r$ e $R_2 = I_{r+1} \times \cdots \times I_p$.

Sejam $\mathbf{X} \in \mathbb{R}^p$, $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ particionados como em (2.27) e tal que $\mathbf{X} \sim \text{TEl}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; R; g)$, com R dado em (2.28). As f.d.p marginais de \mathbf{X}_1 e \mathbf{X}_2 são dadas por

$$f_{\mathbf{X}_1}(\mathbf{x}_1) = \frac{\int_{R_2} g((\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})) d\mathbf{x}_2}{\int_R g((\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})) d\mathbf{x}}, \quad \mathbf{x}_1 \in R_1. \quad (2.29)$$

$$f_{\mathbf{X}_2}(\mathbf{x}_2) = \frac{\int_{R_1} g((\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})) d\mathbf{x}_1}{\int_R g((\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})) d\mathbf{x}}, \quad \mathbf{x}_2 \in R_2. \quad (2.30)$$

Note que as f.d.p marginais acima não têm necessariamente a forma (2.19) e, portanto, não pertencem necessariamente à classe de distribuições elípticas truncadas. No Teorema 2.4.5 apresentamos condições sob as quais as distribuições marginais de um vetor aleatório com distribuição elíptica truncada também pertencem à mesma classe de distribuições.

Teorema 2.4.5. *Sejam $\mathbf{X} \in \mathbb{R}^p$, $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ particionados como em (2.27) e tal que $\mathbf{X} \sim \text{TEl}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; R; g)$, com R dado em (2.28). Se $\boldsymbol{\Sigma}_{12} = \mathbf{0}$, então $\mathbf{X}_1 \sim \text{TEl}_r(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}; R_1; g_1)$, em que*

$$g_1(u) = \int_{T(R_2)} g(u + \mathbf{w}'_2 \mathbf{w}_2) d\mathbf{w}_2, \quad u \geq 0,$$

sendo $T : \mathbb{R}^{p-r} \rightarrow \mathbb{R}^{p-r}$ a transformação dada por $T(\mathbf{x}) = \boldsymbol{\Sigma}_{22}^{-1/2}(\mathbf{x} - \boldsymbol{\mu}_2)$ e g_1 tal que $\int_0^\infty t^{r-1} g_1(t^2) dt < \infty$.

Demonstração. Ver Apêndice A.8 ■

No Teorema 2.4.5 mostramos que se $\mathbf{X}' = (\mathbf{X}'_1, \mathbf{X}'_2)$ tem distribuição elíptica truncada, então \mathbf{X}_1 e \mathbf{X}_2 apresentam distribuição elíptica truncada se $\Sigma(p \times p) > 0$ é matriz bloco diagonal. Esta condição não será necessária para as distribuições condicionais de $\mathbf{X}_1|\mathbf{X}_2$ e $\mathbf{X}_2|\mathbf{X}_1$. De fato, estas distribuições condicionais apresentam distribuição elíptica truncada para $\boldsymbol{\mu} \in \mathbb{R}^p$ e $\Sigma(p \times p) > 0$ arbitrários, como mostramos no Teorema 2.4.6.

Teorema 2.4.6. *Sejam $\mathbf{X} \in \mathbb{R}^p$, $\boldsymbol{\mu} \in \mathbb{R}^p$, $\Sigma(p \times p) > 0$ particionados como em (2.27) e tal que $\mathbf{X} \sim TE\ell_p(\boldsymbol{\mu}, \Sigma; R; g)$, com R dado em (2.28). Então, $\mathbf{X}_1|\mathbf{X}_2 = \mathbf{x}_2 \sim TE\ell_r(\boldsymbol{\mu}_1(\mathbf{x}_2), \Sigma_{11.2}; R_1; g_{q(\mathbf{x}_2)})$, em que*

$$\begin{aligned}\boldsymbol{\mu}_1(\mathbf{x}_2) &= \boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \\ \Sigma_{11.2} &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}, \\ g_{q(\mathbf{x}_2)}(u) &= g(u + q(\mathbf{x}_2)), \text{ com } q(\mathbf{x}_2) = (\mathbf{x}_2 - \boldsymbol{\mu}_2)'\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2).\end{aligned}\tag{2.31}$$

Demonstração. Ver Apêndice A.9 ■

Sejam $\mathbf{X} \in \mathbb{R}^p$, $\boldsymbol{\mu} \in \mathbb{R}^p$, $\Sigma(p \times p) > 0$ particionados como em (2.27) e tais que $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$. Sabemos que \mathbf{X}_1 e \mathbf{X}_2 são independentes se, e somente se, $\Sigma_{12} = \mathbf{0}$. No Teorema 2.4.7, mostramos que a única família da classe de distribuições elípticas truncadas que satisfaz esta propriedade é a família de distribuições normais multivariadas truncadas.

Teorema 2.4.7. *Sejam $\mathbf{X} \in \mathbb{R}^p$, $\boldsymbol{\mu} \in \mathbb{R}^p$, $\Sigma(p \times p) > 0$ particionados como em (2.27) e tal que $\mathbf{X} \sim TE\ell_p(\boldsymbol{\mu}, \Sigma; R; g)$, com R dado em (2.28). Então, \mathbf{X}_1 e \mathbf{X}_2 são independentes se, e somente se, $\mathbf{X} \sim TN_p(\boldsymbol{\mu}, \Sigma; R; g)$ e $\Sigma_{12} = \mathbf{0}$.*

Demonstração. Ver Apêndice A.10 ■

2.5 Geração de amostras aleatórias

Nesta seção propomos um algoritmo baseado no amostrador de Gibbs para gerar amostras aleatórias de um vetor aleatório apresentando distribuição elíptica truncada sobre retângulos em \mathbb{R}^p .

Sejam $\mathbf{X}_{-k} \in \mathbb{R}^{p-1}$ e $\boldsymbol{\mu}_{-k} \in \mathbb{R}^{p-1}$, $k = 1, \dots, p$, os subvetores obtidos a partir de $\mathbf{X} \in \mathbb{R}^p$ e $\boldsymbol{\mu} \in \mathbb{R}^p$, respectivamente, tirando seus k -ésimos componentes. Adicionalmente, sejam $\Sigma_{k,-k} \in \mathbb{R}^{p-1}$ o subvetor obtido tirando o k -ésimo componente da k -ésima linha de $\Sigma(p \times p) > 0$, $\Sigma_{-k,k} = \Sigma'_{k,-k}$, e $\Sigma_{-k,-k}$ a submatriz obtida ao excluir a k -ésima linha e a k -ésima coluna de Σ . Além disso, denotamos por $R = I_1 \times \dots \times I_p$ o retângulo em \mathbb{R}^p em que I_1, \dots, I_p são intervalos em \mathbb{R} (finitos ou infinitos). Sem perda de generalidade, supomos que $I_k = (a_k, b_k)$, $k = 1, \dots, p$.

Construímos uma cadeia de Markov amostrando das distribuições condicionais completas $X_k | \mathbf{X}_{-k}$, $k = 1, \dots, p$, as quais são determinadas pelo Teorema 2.4.6. Com efeito, $X_k | \mathbf{X}_{-k} \sim \text{TE}\ell_1(\mu_{k,-k}, \sigma_{k,-k}; (a_k, b_k); g_{k,-k})$, em que

$$\begin{aligned}\mu_{k,-k} &= \mu_k + \Sigma_{k,-k} \Sigma_{-k,-k}^{-1} (\mathbf{x}_{-k} - \boldsymbol{\mu}_{-k}), \\ \sigma_{k,-k} &= \sigma_{kk} - \Sigma_{k,-k} \Sigma_{-k,-k}^{-1} \Sigma_{-k,k}, \\ g_{k,-k}(u) &= g(u + q(\mathbf{x}_{-k})), \text{ com } q(\mathbf{x}_{-k}) = (\mathbf{x}_{-k} - \boldsymbol{\mu}_{-k})' \Sigma_{-k,-k}^{-1} (\mathbf{x}_{-k} - \boldsymbol{\mu}_{-k}),\end{aligned}$$

para $k = 1, \dots, p$.

Denotamos por $\mathbf{x}^{(j)}$ a amostra gerada na j -ésima iteração, $j = 1, \dots, n$.

Algoritmo 2.5.1.

1. Escolha um valor inicial $\mathbf{x}^{(0)}$ da cadeia de Markov.
2. Geramos um valor aleatório u da distribuição $U(0, 1)$.
3. Em cada ciclo $j = 1, \dots, n$, usamos (2.24) para calcular

$$\begin{aligned}x_{k,-k}^{(j)} &= \mu_{k,-k}^{(j)} + \sigma_{k,-k}^{(j)} F_Z^{-1} \left[u \left\{ F_Z \left(\frac{b_k - \mu_{k,-k}^{(j)}}{\sigma_{k,-k}^{(j)}} \right) - F_Z \left(\frac{a_k - \mu_{k,-k}^{(j)}}{\sigma_{k,-k}^{(j)}} \right) \right\} \right. \\ &\quad \left. + F_Z \left(\frac{a_k - \mu_{k,-k}^{(j)}}{\sigma_{k,-k}^{(j)}} \right) \right],\end{aligned}$$

para $k = 1, \dots, p$, em que $Z \sim S_1(g)$. Este é o valor amostrado da distribuição condicional

$$x_k^{(j)} | x_1^{(j)}, \dots, x_{k-1}^{(j)}, x_{k+1}^{(j-1)}, \dots, x_p^{(j-1)}, \quad k = 1, \dots, p.$$

O Algoritmo 2.5.1 é uma proposta geral para gerar vetores aleatórios das distribuições elípticas truncadas, mas podem existir algoritmos alternativos e, possivelmente, mais eficientes para gerar essas amostras. No pacote *tmvtnorm* (Wilhelm e Manjunath [20]) do software livre R [10], está implementado o Algoritmo 2.5.1 para a distribuição normal multivariada truncada sobre retângulos. Esse pacote também inclui uma rotina para gerar vetores aleatórios da distribuição t multivariada truncada, a qual está baseada numa representação estocástica através da distribuição normal multivariada truncada, ao invés de utilizar diretamente o Algoritmo 2.5.1. O pacote *tmvtnorm* também tem implementado o algoritmo de aceitação-rejeição (von Neumann [28]) para geração de vetores aleatórios para ambas as distribuições.

2.6 Conclusões

Neste capítulo definimos e estudamos a classe de distribuições elípticas truncadas, necessária para o estudo da classe de distribuições Box-Cox elípticas apresentada no Capítulo 3. Mostramos que a classe de distribuições elípticas truncadas tem como membros as famílias de distribuições normal e t multivariadas truncadas, e como caso particular a classe de distribuições elípticas. Também, apresentamos uma caracterização desta nova classe através das distribuições esféricas truncadas e mostramos resultados sobre log-concavidade, distribuições marginal e condicional, e independência. Finalmente, propusemos um algoritmo baseado no amostrador de Gibbs para gerar vetores aleatórios das distribuições elípticas truncadas sobre retângulos.

Capítulo 3

A classe de distribuições Box-Cox elípticas

3.1 Introdução

A distribuição normal multivariada é utilizada habitualmente para modelar dados multivariados. Os modelos baseados nesta distribuição pressupõem simetria e apresentam problemas na presença de observações discrepantes, o que pode levar a conclusões erradas. Distribuições como a t multivariada que tem caudas mais pesadas em relação à distribuição normal multivariada são uma alternativa para modelar dados com presença de observações discrepantes (Lange et al. [3]), porém mantém o pressuposto de simetria. Existem várias alternativas para reduzir a falta de normalidade de um conjunto de dados, tal como a transformação de Box-Cox [29], que tem sido extensivamente estudada e aplicada em estatística. No caso univariado, a transformação de Box-Cox é definida para uma variável aleatória $Y > 0$ como

$$Y^{(\lambda)} = \begin{cases} \frac{Y^\lambda - 1}{\lambda}, & \text{se } \lambda \neq 0, \\ \log Y, & \text{se } \lambda = 0, \end{cases} \quad (3.1)$$

em que λ é chamado de parâmetro da transformação. Esta transformação é usada para atender às suposições de normalidade, constância da variância e aditividade em modelos de regressão. Sakia [30] apresenta uma análise e compilação bibliográfica sobre esta transformação. No cenário multivariado a transformação de Box-Cox é definida para um vetor aleatório $\mathbf{Y} = (Y_1, \dots, Y_p) \in \mathbb{R}_+^p$ como

$$Y_k^{(\lambda_k)} = \begin{cases} \frac{Y_k^{\lambda_k} - 1}{\lambda_k}, & \text{se } \lambda_k \neq 0, \\ \log Y_k, & \text{se } \lambda_k = 0, \end{cases} \quad (3.2)$$

* \mathbb{R}_+^p denota ao hiperoctante positivo, definido como $\mathbb{R}_+^p = \{\mathbf{a} \in \mathbb{R}^p : a_k > 0, k = 1, \dots, p\}$.

para $k = 1, \dots, p$. Assim, a transformação de Box-Cox multivariada considera a transformação (3.1) para cada componente de \mathbf{Y} com parâmetro da transformação dado pelo vetor $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_p)$. Em geral, a transformação de Box-Cox multivariada é considerada para corrigir problemas de não normalidade multivariada. Andrews et al. [31] e Velilla [32] consideram a transformação (3.2) e assumem que o vetor transformado $\mathbf{Y}^{(\boldsymbol{\lambda})} = (Y_1^{(\lambda_1)}, \dots, Y_p^{(\lambda_p)})$ tem distribuição normal multivariada. Uma abordagem mais geral é proposta por Quiroz et al. [33], que consideram uma distribuição elíptica para o vetor transformado $\mathbf{Y}^{(\boldsymbol{\lambda})}$.

A abordagem clássica da transformação Box-Cox multivariada tem várias desvantagens. Assumir que o vetor transformado $\mathbf{Y}^{(\boldsymbol{\lambda})}$ tenha distribuição normal multivariada (ou elíptica) é teoricamente incorreto, já que o domínio de $\mathbf{Y}^{(\boldsymbol{\lambda})}$ não é \mathbb{R}^p , a menos que $\boldsymbol{\lambda} = \mathbf{0}$. Com efeito, em (3.2) podemos notar que $Y_k > 0$ implica que $Y_k^{(\lambda_k)} > -1/\lambda_k$, se $\lambda_k > 0$, e $Y_k^{(\lambda_k)} < -1/\lambda_k$, se $\lambda_k < 0$, para $k = 1, \dots, p$. Goto e Hamasaki [34] tratam este problema atribuindo uma distribuição normal multivariada truncada para o vetor transformado $\mathbf{Y}^{(\boldsymbol{\lambda})}$. Esses autores chamaram de distribuição normal potência multivariada a este modelo. Ainda assim, tanto a proposta de Goto e Hamasaki como a abordagem clássica da transformação de Box-cox multivariada têm a desvantagem de que os parâmetros do modelo não são interpretáveis em termos de características do vetor de variáveis originais \mathbf{Y} . Em geral, na literatura aparecem poucos trabalhos de modelagem estatística através da transformação de Box-Cox multivariada. Por exemplo, uma alternativa para a distribuição skew- t multivariada para modelagem estatística (Lin et al. [35]) foi proposta por Lo e Gottardo [36]. Esses autores fornecem um modelo de mistura baseado numa família de distribuições assimétricas, a família de distribuições t multivariadas com a transformação de Box-Cox. Esse modelo permite identificar dados discrepantes usando conjuntamente a transformação de Box-Cox multivariada, mas não leva em consideração o truncamento que esta transformação induz.

Uma extensão da transformação de Box-Cox (3.1), proposta por Cole e Green [6], é definida para uma variável aleatória unidimensional $Y > 0$ como

$$\sigma Y^{(\lambda, \mu)} = \begin{cases} \frac{(Y/\mu)^\lambda - 1}{\lambda}, & \text{se } \lambda \neq 0, \\ \log(Y/\mu), & \text{se } \lambda = 0, \end{cases} \quad (3.3)$$

em que $\mu > 0$ e $\sigma > 0$ são parâmetros de escala e dispersão relativa de Y , respectivamente. Seguindo a abordagem clássica da transformação de Box-Cox, os autores supõem que a variável transformada $Y^{(\lambda, \mu)}$ segue uma distribuição normal padrão e utilizam o método de máxima verossimilhança penalizada para ajustar curvas pelo método de LMS (Cole [37]). Posteriormente, com base na transformação (3.3), Rigby e Stasinopoulos [7, 8] definiram as distribuições Box-Cox exponencial potência e Box-Cox t , supondo que a variável transformada $Y^{(\lambda, \mu)}$ apresenta distribuição exponencial

potência padrão truncada e t truncada, respectivamente. Estas distribuições são construídas levando em consideração o truncamento que a transformação (3.3) induz, e ambas têm como caso particular a distribuição Box-Cox Cole Green (Box-Cox normal). Tanto a distribuição Box-Cox exponencial potência como a distribuição Box-Cox t foram propostas por Rigby e Stasinopoulos para modelar assimetria e curtose em uma variável dependente através do método GAMLSS, que é uma generalização do método LMS. Uma abordagem geral das distribuições introduzidas por Rigby e Stasinopoulos foi proposta por Ferrari e Fumes [4], que definiram a classe de distribuições Box-Cox simétricas supondo que a variável transformada $Y^{(\lambda, \mu)}$ apresenta distribuição simétrica padrão truncada. Esta classe, além das distribuições Box-Cox exponencial potência e Box-Cox t , proporciona uma ampla variedade de distribuições para modelagem estatística, incluindo a classe de distribuições log-simétricas (Vanegas e Paula [5]). Os parâmetros das distribuições Box-Cox simétricas são introduzidos através da transformação (3.3) e, dependendo da família de distribuições considerada, podem aparecer parâmetros extras. Uma vantagem em relação à abordagem clássica da transformação de Box-Cox é que os parâmetros que envolvem as distribuições Box-Cox simétricas podem ser interpretados como características da variável Y . Esta classe fornece uma ferramenta estatística útil para modelagem de dados positivos, assimétricos e, possivelmente, com caudas pesadas.

Ferrari e Fumes [4] analisam um conjunto de dados nutricionais através de ajustes das distribuições Box-Cox simétricas ao consumo de micro e macronutrientes. As autoras mostram que, em geral, o modelo Box-Cox t apresenta melhor ajuste que outros modelos existentes na literatura. Caso o interesse seja analisar o consumo de vários nutrientes conjuntamente, as distribuições Box-Cox simétricas não seriam apropriadas, a menos que se suponha independência entre os consumos de nutrientes. Esta abordagem não é adequada uma vez que é comum a existência de associação entre os consumos de nutrientes. Com base nisso aparece a seguinte pergunta: é possível construir distribuições multivariadas de maneira que as distribuições marginais apresentem características similares a variáveis aleatórias com distribuições Box-Cox simétricas e que, além disso, a associação entre essas variáveis seja levada em consideração? Neste capítulo respondemos a essa pergunta construindo uma classe de distribuições através de uma extensão multivariada da transformação (3.3) e envolvendo a classe de distribuições elípticas truncadas apresentada no Capítulo 2, a qual denominaremos de classe de distribuições Box-Cox elípticas. Esta classe tem como caso particular a classe de distribuições Box-Cox simétricas e a classe de distribuições log-elípticas (Fang et al. [9, Seção 2.8.1]). Alguns parâmetros que determinam as distribuições Box-Cox elípticas são interpretados como características das distribuições marginais e outros parâmetros induzem associação entre essas distribuições. Essa interpretação dos parâmetros como características do vetor de variáveis originais permite modelar dados positivos mul-

tivariados, marginalmente assimétricos e com presença de observações discrepantes. Além disso, evidenciam uma vantagem em relação ao modelo de Goto e Hamasaki [34] e a abordagem clássica da transformação de Box-Cox multivariada. Neste capítulo também apresentamos várias propriedades das distribuições Box-Cox elípticas. Para tratar o problema de estimação dos parâmetros propomos o método de máxima verossimilhança, e verificamos que esta metodologia é adequada através de estudos de simulação. Por fim, são apresentadas aplicações da classe de distribuições Box-Cox elípticas a dados nutricionais e antropométricos.

3.2 A transformação de Box-Cox estendida

Nesta seção definimos e estudamos uma generalização da transformação de Box-Cox multivariada (3.2), a transformação de Box-Cox estendida. Esta transformação é essencial para definir a classe de distribuições Box-Cox elípticas apresentada na Seção 3.3. A transformação de Box-Cox estendida, além do vetor de parâmetros de transformação, inclui um vetor de parâmetros em que cada componente é um parâmetro de escala da respectiva distribuição marginal.

Para cada $\mathbf{a} \in \mathbb{R}^p$, denotamos por $\mathbf{D}_\mathbf{a}$ a matriz diagonal conformada com os componentes de \mathbf{a} na diagonal principal, isto é, $\mathbf{D}_\mathbf{a} = \text{diag}\{a_1, \dots, a_p\}$. Também, denotamos por $A(\mathbf{a})$ ao retângulo em \mathbb{R}^p dado pelo produto cartesiano $A(\mathbf{a}) = I(a_1) \times \dots \times I(a_p)$, em que

$$I(a_k) = \begin{cases} (-1/a_k, \infty), & a_k > 0, \\ (-\infty, -1/a_k), & a_k < 0, \\ (-\infty, \infty), & a_k = 0, \end{cases} \quad (3.4)$$

para $k = 1, \dots, p$.

Definição 3.2.1. *Sejam $\boldsymbol{\lambda} \in \mathbb{R}^p$ e $\boldsymbol{\mu} \in \mathbb{R}_+^p$. Definimos a transformação de Box-Cox estendida $T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})} : \mathbb{R}_+^p \rightarrow A(\boldsymbol{\lambda})$ para o vetor aleatório $\mathbf{Y} \in \mathbb{R}_+^p$ como $T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}(\mathbf{Y}) = \mathbf{W} = (W_1, \dots, W_p)$, em que*

$$W_k = \begin{cases} \frac{(Y_k/\mu_k)^{\lambda_k} - 1}{\lambda_k}, & \lambda_k \neq 0, \\ \log(Y_k/\mu_k), & \lambda_k = 0, \end{cases} \quad (3.5)$$

para $k = 1, \dots, p$.

Note que a transformação de Box-Cox multivariada (3.2) pode ser obtida de (3.5) quando $\boldsymbol{\mu} = \mathbf{1} = (1, \dots, 1)'$. No Teorema 3.2.1 mostramos que se transformamos um vetor aleatório $\mathbf{Y} \in \mathbb{R}_+^p$ através de (3.5), então o k -ésimo componente do vetor $\boldsymbol{\mu}$ é parâmetro de escala do k -ésimo componente de \mathbf{Y} .

Teorema 3.2.1. *Sejam $\boldsymbol{\mu} \in \mathbb{R}_+^p$ e $\boldsymbol{\lambda} \in \mathbb{R}^p$. Sejam ainda $\mathbf{Y} \in \mathbb{R}_+^p$ e $\mathbf{W} \in A(\boldsymbol{\lambda})$ vetores aleatórios tais que $T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}(\mathbf{Y}) = \mathbf{W}$. Então μ_k é parâmetro de escala de Y_k , para $k = 1, \dots, p$.*

Demonstração. Ver Apêndice A.11. ■

No Teorema 3.2.2 mostramos que a composição de determinadas transformações lineares com a transformação de Box-Cox estendida produzem novas transformações de Box-Cox estendidas, mas os vetores de parâmetros de escala são alterados pelas transformações lineares.

Teorema 3.2.2. *Sejam $\boldsymbol{\mu} \in \mathbb{R}_+^p$ e $\boldsymbol{\lambda} \in \mathbb{R}^p$.*

1. *Se $\mathbf{a} \in \mathbb{R}_+^p$ e $U : \mathbb{R}^p \rightarrow \mathbb{R}^p$ é a transformação linear dada por $U(\mathbf{x}) = \mathbf{D}_a \mathbf{x}$, então $T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})} \circ U = T_{(\boldsymbol{\lambda}, \mathbf{D}_a^{-1} \boldsymbol{\mu})}$.*
2. *Se $\mathbf{b} \in A(\boldsymbol{\lambda})$ e $V : \mathbb{R}^p \rightarrow \mathbb{R}^p$ é a transformação linear dada por $V(\mathbf{x}) = \mathbf{D}(\mathbf{x} - \mathbf{b})$, em que $\mathbf{D} = \mathbf{D}_{1+\mathbf{D}\boldsymbol{\lambda}\mathbf{b}}^{-1}$, então $V \circ T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})} = T_{(\boldsymbol{\lambda}, \boldsymbol{\delta})}$, sendo $\boldsymbol{\delta} = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}^{-1}(\mathbf{b})$.*

Demonstração. Ver Apêndice A.12. ■

O Teorema 3.2.2 permite provar algumas propriedades das distribuições Box-Cox elípticas apresentadas na Seção 3.4.

3.3 A classe de distribuições Box-Cox elípticas

Nesta seção definimos uma nova classe de distribuições que chamaremos de classe de distribuições Box-Cox elípticas. Esta classe está definida para vetores aleatórios $\mathbf{Y} \in \mathbb{R}_+^p$ e é caracterizada por meio da classe de distribuições elípticas truncadas estudada no Capítulo 2. Os parâmetros das distribuições Box-Cox elípticas são introduzidos através de uma matriz $\boldsymbol{\Sigma}(p \times p) > 0$ e da transformação de Box-Cox estendida $T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}$ apresentada na Seção 3.2.

Definição 3.3.1. *Dizemos que o vetor aleatório $\mathbf{Y} = (Y_1, \dots, Y_p) \in \mathbb{R}_+^p$ tem distribuição Box-Cox elíptica com parâmetros $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p) \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p$, $\boldsymbol{\Sigma} = (\sigma_{ij})_{p \times p} > 0$ e f.g.d g se $T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}(\mathbf{Y}) \sim TE\ell_p(\mathbf{0}, \boldsymbol{\Sigma}; A(\boldsymbol{\lambda}); g)$, e escrevemos $\mathbf{Y} \sim BCE\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$. A função g é tal que $g(u) \geq 0$, para todo $u \geq 0$, e $\int_0^\infty t^{p-1} g(t^2) dt < \infty$.*

Quando consideramos o caso $\boldsymbol{\lambda} = \mathbf{0}$ na Definição 3.3.1 dizemos que \mathbf{Y} segue uma distribuição log-elíptica com parâmetros $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e f.g.d g (Fang et al. [9, Seção 2.8.1]), e escrevemos $\mathbf{Y} \sim LE\ell(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$.

No Teorema 3.3.1 é dada a f.d.p de um vetor aleatório com distribuição Box-Cox elíptica.

Teorema 3.3.1. *Sejam $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$ e $\boldsymbol{\Sigma}(p \times p) > 0$. Se $\mathbf{Y} \sim BCEl_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$, então a f.d.p de \mathbf{Y} é*

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) \prod_{k=1}^p \frac{y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}}}{\int_{A(\boldsymbol{\lambda})} g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) d\mathbf{w}}, \quad \mathbf{w} = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}(\mathbf{y}), \quad \mathbf{y} \in \mathbb{R}_+^p. \quad (3.6)$$

Demonstração. Ver Apêndice A.13. ■

O caso $p = 1$ em (3.6) corresponde à f.d.p de uma variável aleatória Y com distribuição Box-Cox simétrica de parâmetros $\mu > 0$, $\lambda \in \mathbb{R}$, $\sigma > 0$ e f.g.d g , denotada por $Y \sim BCS(\mu, \lambda, \sigma; g)$. Esta classe de distribuições foi definida e estudada por Ferrari e Fumes [4].

Note que a f.d.p (3.6) pode ser expressa como

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{f_U(\mathbf{w}) \prod_{k=1}^p \frac{y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}}}{P(U \in A(\boldsymbol{\lambda}))}, \quad \mathbf{w} = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}(\mathbf{y}), \quad \mathbf{y} \in \mathbb{R}_+^p, \quad (3.7)$$

em que f_U é a f.d.p de um vetor aleatório $U \sim El(\mathbf{0}, \boldsymbol{\Sigma}; g)$. Se consideramos $\boldsymbol{\lambda} = \mathbf{0}$ em (3.7), isto é $\mathbf{Y} \sim LEl(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$, então $A(\mathbf{0}) = \mathbb{R}^p$, e portanto $P(U \in A(\mathbf{0})) = 1$. Assim, a f.d.p de $\mathbf{Y} \in \mathbb{R}_+^p$ neste caso é dada por

$$f_{\mathbf{Y}}(\mathbf{y}) = \det(\boldsymbol{\Sigma})^{-1/2} g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) \prod_{k=1}^p \frac{1}{y_k}, \quad \mathbf{w} = T_{(\mathbf{0}, \boldsymbol{\mu})}(\mathbf{y}), \quad \mathbf{y} \in \mathbb{R}_+^p. \quad (3.8)$$

O caso $p = 1$ em (3.8) corresponde à f.d.p de uma variável aleatória unidimensional $Y > 0$ com distribuição log-simétrica de parâmetros $\mu > 0$, $\sigma^2 > 0$ e f.g.d g , denotada por $Y \sim LS(\mu, \sigma^2; g)$. Esta classe de distribuições foi estudada por Vanegas e Paula [5].

As f.g.d apresentadas na Seção 2.2 proporcionam vários membros da classe de distribuições Box-Cox elípticas de acordo com a forma da f.g.d g . Todas as propriedades da Observação 2.2.1 também são satisfeitas para a classe de distribuições Box-Cox elípticas. Neste trabalho é de grande interesse a família de distribuições Box-Cox normal multivariadas, gerada pela f.g.d (2.5), e a família de distribuições Box-Cox t multivariadas, gerada pela f.g.d (2.8). Dizemos que o vetor aleatório $\mathbf{Y} \in \mathbb{R}_+^p$ tem distribuição Box-Cox normal multivariada com parâmetros $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$ e $\boldsymbol{\Sigma}(p \times p) > 0$, denotada por $\mathbf{Y} \sim BCN_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma})$, se tem f.d.p dada por

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\exp\left(-\frac{1}{2}\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}\right) \prod_{k=1}^p \frac{y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}}}{\int_{A(\boldsymbol{\lambda})} \exp\left(-\frac{1}{2}\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}\right) d\mathbf{w}}, \quad \mathbf{w} = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}(\mathbf{y}), \quad \mathbf{y} \in \mathbb{R}_+^p. \quad (3.9)$$

No caso de um vetor aleatório $\mathbf{Y} \in \mathbb{R}_+^p$ com distribuição log-normal multivariada com

parâmetros $\boldsymbol{\mu} \in \mathbb{R}_+^p$ e $\boldsymbol{\Sigma}(p \times p) > 0$, denotado por $\mathbf{Y} \sim \text{LN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, a f.d.p de \mathbf{Y} é dada por

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\det(\boldsymbol{\Sigma})^{-1/2}}{(2\pi)^{p/2}} \exp\left(-\frac{1}{2}\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}\right) \prod_{k=1}^p \frac{1}{y_k}, \quad \mathbf{w} = T_{(0,\boldsymbol{\mu})}(\mathbf{y}), \quad \mathbf{y} \in \mathbb{R}_+^p. \quad (3.10)$$

O vetor aleatório $\mathbf{Y} \in \mathbb{R}_+^p$ tem distribuição Box-Cox t multivariada com parâmetros $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $\tau > 0$ graus de liberdade, denotado por $\mathbf{Y} \sim \text{BC}t_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; \tau)$, se sua f.d.p é

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\left(1 + \frac{1}{\tau}\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}\right)^{-\frac{\tau+p}{2}} \prod_{k=1}^p \frac{y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}}}{\int_{A(\boldsymbol{\lambda})} \left(1 + \frac{1}{\tau}\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}\right)^{-\frac{\tau+p}{2}} d\mathbf{w}}, \quad \mathbf{w} = T_{(\boldsymbol{\lambda},\boldsymbol{\mu})}(\mathbf{y}), \quad \mathbf{y} \in \mathbb{R}_+^p. \quad (3.11)$$

No caso de um vetor aleatório $\mathbf{Y} \in \mathbb{R}_+^p$ com distribuição log- t multivariada com parâmetros $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $\tau > 0$ graus de liberdade, denotado por $\mathbf{Y} \sim \text{L}t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \tau)$, a f.d.p de \mathbf{Y} é da forma

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\Gamma((\tau+p)/2) \det(\boldsymbol{\Sigma})^{-1/2}}{(\tau\pi)^{p/2} \Gamma(\tau/2)} \left(1 + \frac{1}{\tau}\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}\right)^{-\frac{\tau+p}{2}} \prod_{k=1}^p \frac{1}{y_k}, \quad \mathbf{y} \in \mathbb{R}_+^p, \quad (3.12)$$

em que $\mathbf{w} = T_{(0,\boldsymbol{\mu})}(\mathbf{y})$.

Na Figura 3.1 são apresentadas curvas de nível e os gráficos da f.d.p de um vetor aleatório $\mathbf{Y} \sim \text{BC}t_2(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; \tau)$ para diferentes valores dos parâmetros. A forma das distribuições muda segundo os valores destes parâmetros. As distribuições apresentadas nas Figuras 3.1(a), 3.1(e), 3.1(f), 3.1(g), 3.1(h) e 3.1(i) são obtidas a partir das distribuições t bivariadas truncadas apresentadas nas Figuras 2.1(a), 2.1(b), 2.1(c), 2.1(d), 2.1(e) e 2.1(f), respectivamente (Definição 3.3.1). Nas Figuras 3.1(a), 3.1(b) e 3.1(c) fixamos os parâmetros exceto σ_{12} que toma os valores $-0,2$, 0 e $0,2$, respectivamente. Note que a associação entre as distribuições marginais de Y_1 e Y_2 é alterada de acordo com os valores que toma este parâmetro. Na Figura 3.1(d) o parâmetro μ_1 é diminuído a 3,5 em relação à Figura 3.1(c). Este é um parâmetro de escala da distribuição marginal de Y_1 . Na Figura 3.1(e) o parâmetro σ_{22} é diminuído a 0,2 em relação à Figura 3.1(d). Este é um parâmetro de dispersão relativa da distribuição marginal de Y_2 . Nas Figuras 3.1(f), 3.1(g) e 3.1(h) os parâmetros λ_1 e λ_2 apresentam diferentes duplas de valores em relação à Figura 3.1(e). Estes parâmetros controlam a assimetria das respectivas distribuições marginais de Y_1 e Y_2 . Na Figura 3.1(g) em que $\lambda_1 = \lambda_2 = 1$, é evidente que as curvas de nível são elipses projetadas sobre \mathbb{R}_+^2 ; este fato é provado no item 3 do Teorema 3.4.2. Na Figura 3.1(i) é aumentado o número de graus de liberdade a 10 em relação à Figura 3.1(h). À medida que os graus de liberdade aumentam, as curvas de nível da distri-

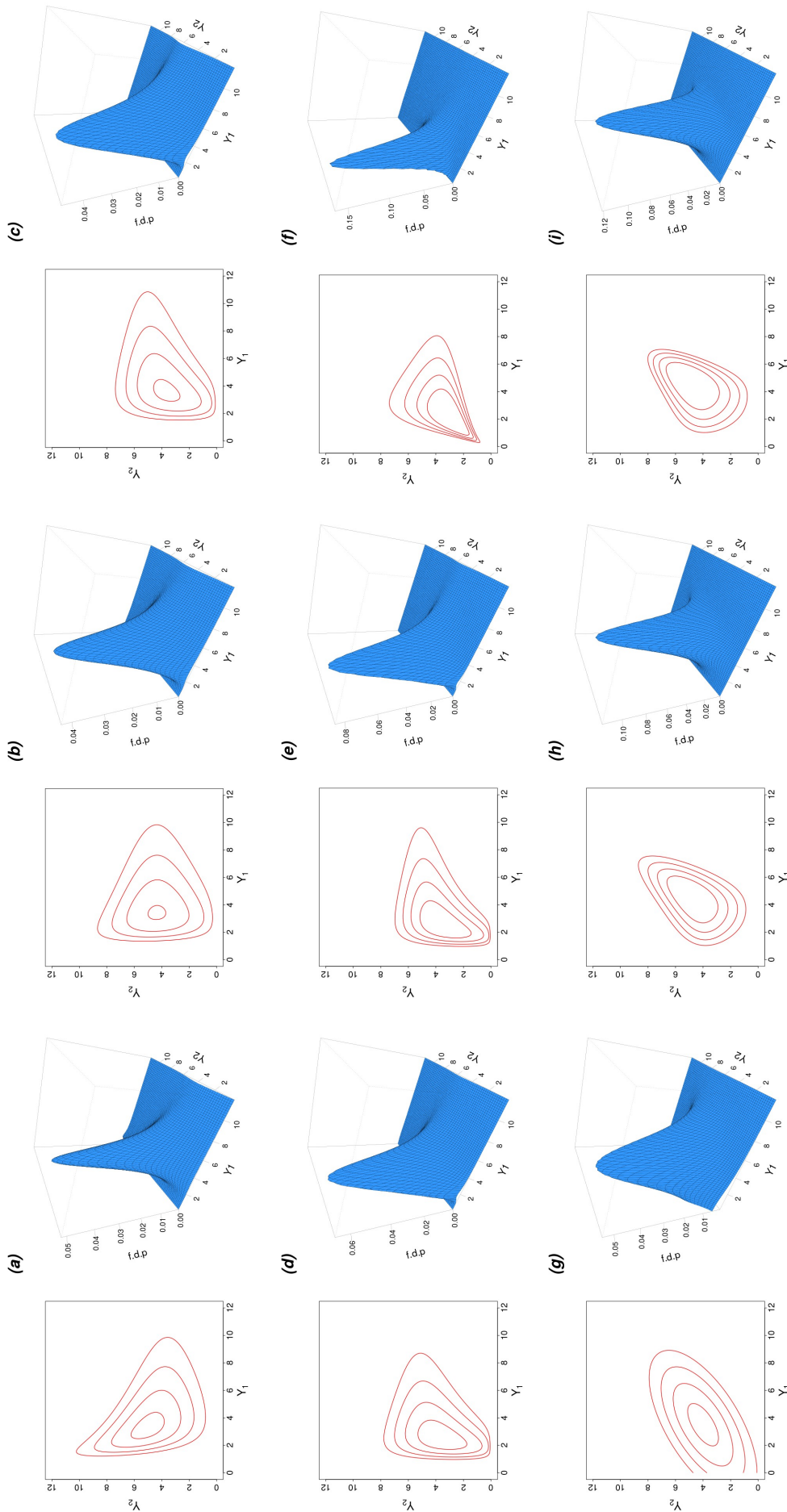


Figura 3.1: Curvas dos níveis 0,04, 0,02, 0,01, 0,005 e gráficos da f.d.p de $\mathbf{Y} \sim BCt_2(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; \tau)$, em que $\mu_2 = 4, \sigma_{11} = 0,5 e$ (a) $\mu_1 = 5, \lambda_1 = -1, \lambda_2 = 1,5, \sigma_{22} = 0,3$, $\sigma_{12} = 0,2, \tau = 3$, (b) $\mu_1 = 3, \mu_2 = 1,5, \sigma_{22} = 0,3, \sigma_{12} = 0,2, \tau = 3$, (c) $\mu_1 = 5, \lambda_1 = -1, \lambda_2 = 1,5, \sigma_{22} = 0,3, \sigma_{12} = 0,2, \tau = 3$, (d) $\mu_1 = 3,5, \lambda_1 = -1, \lambda_2 = 1,5, \sigma_{22} = 0,3, \sigma_{12} = 0,2, \tau = 3$, (e) $\mu_1 = 3, \mu_2 = 1,5, \sigma_{22} = 0,3, \sigma_{12} = 0,2, \tau = 3$, (f) $\mu_1 = 3,5, \lambda_1 = -1, \lambda_2 = 1,5, \sigma_{22} = 0,2, \sigma_{12} = 0,2, \tau = 3$, (g) $\mu_1 = 3,5, \lambda_1 = \lambda_2 = 1, \sigma_{22} = 0,2, \sigma_{12} = 0,2, \tau = 3$, (h) $\mu_1 = 3,5, \lambda_1 = 3, \lambda_2 = 2, \sigma_{22} = 0,2, \sigma_{12} = 0,2, \tau = 10$, (i) $\mu_1 = 3,5, \lambda_1 = 3, \lambda_2 = 2, \sigma_{22} = 0,2, \sigma_{12} = 0,2, \tau = 10$.

buição Box-Cox t bivariada tendem às curvas de nível da distribuição Box-Cox normal bivariada. Além disso, valores pequenos de τ sugerem que as caudas da distribuição Box-Cox t são mais pesadas em relação a valores grandes de τ . Na Seção 3.5 é apresentado um estudo sobre a interpretação dos parâmetros que determinam as distribuições Box-Cox elípticas.

3.4 Propriedades das distribuições Box-Cox elípticas

Nesta seção apresentamos várias propriedades das distribuições Box-Cox elípticas. Inicialmente mostramos uma caracterização através da classe de distribuições elípticas truncadas definida no Capítulo 2. Também provamos resultados distribucionais que envolvem as distribuições marginal e condicional, e independência. Finalmente, derivamos uma fórmula que liga os momentos mistos de vetores aleatórios tendo distribuições Box-Cox elípticas. Nesta seção também apresentamos várias propriedades da classe de distribuições log-elípticas que ainda não estão disponíveis na literatura. Algumas propriedades que apresentamos nesta seção são necessárias para mostrar a interpretação dos parâmetros das distribuições Box-Cox elípticas (Seção 3.5).

No Teorema 3.4.1 mostramos uma caracterização das distribuições Box-Cox elípticas através das distribuições elípticas truncadas. Esta caracterização é uma extensão da Definição 3.3.1.

Teorema 3.4.1. *Sejam $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\mathbf{b} \in A(\boldsymbol{\lambda})$, $\boldsymbol{\delta} = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}^{-1}(\mathbf{b})$, $\mathbf{D} = \mathbf{D}_{\mathbf{1} + \mathbf{D}\boldsymbol{\lambda}\mathbf{b}}^{-1}$ e $\boldsymbol{\Sigma}(p \times p) > 0$. Então, $T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}(\mathbf{Y}) \sim TE\ell_p(\mathbf{b}, \boldsymbol{\Sigma}; A(\boldsymbol{\lambda}); g)$ se, e somente se, $\mathbf{Y} \sim BCE\ell_p(\boldsymbol{\delta}, \boldsymbol{\lambda}, \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}; g)$.*

Demonstração. Ver Apêndice A.14. ■

No Teorema 3.4.2 apresentamos vários resultados distribucionais das distribuições Box-Cox elípticas. Nos itens 1 e 2 mostramos resultados distribucionais quando consideramos alguns tipos de transformações num vetor com distribuição Box-Cox elíptica. No item 3 mostramos que a classe de distribuições elípticas truncadas (2.19), com $B = \mathbb{R}_+^p$ e $\boldsymbol{\mu} \in \mathbb{R}_+^p$, é obtida a partir da classe de distribuições Box-Cox elípticas.

Teorema 3.4.2. *Sejam $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $\mathbf{Y} \sim BCE\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$.*

1. *Se $\mathbf{a} \in \mathbb{R}_+^p$, então $\mathbf{D}_\mathbf{a}\mathbf{Y} \sim BCE\ell_p(\mathbf{D}_\mathbf{a}\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$*
2. *Se $\mathbf{r} \in \mathbb{R}^p \setminus \mathbf{0}$ e $\mathbf{U} \in \mathbb{R}_+^p$ é o vetor aleatório com componentes $U_k = (Y_k/\mu_k)^{r_k}$, $k = 1, \dots, p$, então $\mathbf{U} \sim BCE\ell_p(\mathbf{1}, \mathbf{D}_\mathbf{r}^{-1}\boldsymbol{\lambda}, \mathbf{D}_\mathbf{r}\boldsymbol{\Sigma}\mathbf{D}_\mathbf{r}; g)$.*
3. *Se $\boldsymbol{\lambda} = \mathbf{1}$, então $\mathbf{Y} \sim TE\ell_p(\boldsymbol{\mu}, \mathbf{D}_\boldsymbol{\mu}\boldsymbol{\Sigma}\mathbf{D}_\boldsymbol{\mu}; \mathbb{R}_+^p; g)$.*

Demonstração. Ver Apêndice A.15. ■

Considerando o caso $p = 1$ no Teorema 3.4.2 obtemos os respectivos resultados distribucionais para a classe de distribuições Box-Cox simétricas. Estes resultados já tinham sido provados por Ferrari e Fumes [4].

No Teorema 3.4.3 mostramos que a distribuição log- t multivariada é uma mistura de distribuições log-normais multivariadas.

Teorema 3.4.3. *Sejam $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $\tau > 0$. Se $\mathbf{Y} \sim Lt_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \tau)$, então a f.d.p $f_{\mathbf{Y}}$ de \mathbf{Y} pode ser obtida como*

$$f_{\mathbf{Y}}(\mathbf{y}) = \int_0^\infty f_{\mathbf{Y}|v}(\mathbf{y})f_V(v) dv,$$

em que f_V é a f.d.p de $V \sim IG(\tau/2, \tau/2)$ e $f_{\mathbf{Y}|v}$ é a f.d.p condicional de $\mathbf{Y}|V = v \sim LN_p(\boldsymbol{\mu}, v\boldsymbol{\Sigma})$.

Demonstração. Ver Apêndice A.16. ■

Com o objetivo de estabelecer resultados em relação às distribuições marginais e condicionais, e independência das distribuições Box-Cox elípticas, consideremos partições de $\mathbf{Y} \in \mathbb{R}_+^p$, $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$ e $\boldsymbol{\Sigma}(p \times p) > 0$ como

$$\mathbf{Y}' = (\mathbf{Y}'_1, \mathbf{Y}'_2), \quad \boldsymbol{\mu}' = (\boldsymbol{\mu}'_1, \boldsymbol{\mu}'_2), \quad \boldsymbol{\lambda}' = (\boldsymbol{\lambda}'_1, \boldsymbol{\lambda}'_2), \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}, \quad (3.13)$$

sendo $\mathbf{Y}_1 \in \mathbb{R}_+^r$, $\mathbf{Y}_2 \in \mathbb{R}_+^{p-r}$, $\boldsymbol{\mu}_1 \in \mathbb{R}_+^r$, $\boldsymbol{\mu}_2 \in \mathbb{R}_+^{p-r}$, $\boldsymbol{\lambda}_1 \in \mathbb{R}^r$, $\boldsymbol{\lambda}_2 \in \mathbb{R}^{p-r}$, $\boldsymbol{\Sigma}_{11}(r \times r) > 0$, $\boldsymbol{\Sigma}_{22}((p-r) \times (p-r)) > 0$ e $\boldsymbol{\Sigma}_{12}(r \times (p-r))$ tal que $\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}'_{12}$. O retângulo $A(\boldsymbol{\lambda})$ pode ser escrito como o produto cartesiano entre os retângulos $A(\boldsymbol{\lambda}_1) = I(\lambda_1) \times \cdots \times I(\lambda_r) \in \mathbb{R}^r$ e $A(\boldsymbol{\lambda}_2) = I(\lambda_{r+1}) \times \cdots \times I(\lambda_p) \in \mathbb{R}^{p-r}$, isto é,

$$A(\boldsymbol{\lambda}) = A(\boldsymbol{\lambda}_1) \times A(\boldsymbol{\lambda}_2). \quad (3.14)$$

Sejam $\mathbf{Y} \in \mathbb{R}_+^p$, $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ particionados como em (3.13) e tal que $\mathbf{Y} \sim BCE\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$. As f.d.p marginais dos subvetores aleatórios \mathbf{Y}_1 e \mathbf{Y}_2 são dadas por

$$f_{\mathbf{Y}_1}(\mathbf{y}_1) = \frac{\left\{ \int_{A(\boldsymbol{\lambda}_2)} g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) d\mathbf{w}_2 \right\} \prod_{k=1}^r \frac{y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}}}{\int_{A(\boldsymbol{\lambda})} g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) d\mathbf{w}}, \quad \mathbf{y}_1 \in \mathbb{R}_+^r, \quad (3.15)$$

$$f_{\mathbf{Y}_2}(\mathbf{y}_2) = \frac{\left\{ \int_{A(\boldsymbol{\lambda}_1)} g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) d\mathbf{w}_1 \right\} \prod_{k=r+1}^p \frac{y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}}}{\int_{A(\boldsymbol{\lambda})} g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) d\mathbf{w}}, \quad \mathbf{y}_2 \in \mathbb{R}_+^{p-r}, \quad (3.16)$$

em que $\mathbf{w}_1 = T_{(\boldsymbol{\lambda}_1, \boldsymbol{\mu}_1)}(\mathbf{y}_1)$, $\mathbf{w}_2 = T_{(\boldsymbol{\lambda}_2, \boldsymbol{\mu}_2)}(\mathbf{y}_2)$ e $\mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2)$. Evidentemente, as f.d.p marginais (3.15) e (3.16) não têm necessariamente a estrutura da f.d.p de um vetor

aleatório com distribuição Box-Cox elíptica dada em (3.6). Essa estrutura é possível quando $\Sigma_{12} = \mathbf{0}$, isto é, quando a matriz $\Sigma(p \times p) > 0$ é bloco diagonal. No Teorema 3.4.4 mostramos esse fato.

Teorema 3.4.4. *Sejam $\mathbf{Y} \in \mathbb{R}_+^p$, $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\Sigma(p \times p) > 0$ particionados como em (3.13) e tal que $\mathbf{Y} \sim BCE\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \Sigma; g)$. Se $\Sigma_{12} = \mathbf{0}$, então $\mathbf{Y}_1 \sim BCE\ell_r(\boldsymbol{\mu}_1, \boldsymbol{\lambda}_1, \Sigma_{11}; g_1)$, em que*

$$g_1(u) = \int_{T(A(\boldsymbol{\lambda}_2))} g(u + \mathbf{s}'_2 \mathbf{s}_2) d\mathbf{s}_2, \quad u \geq 0, \quad (3.17)$$

sendo $T : \mathbb{R}^{p-r} \rightarrow \mathbb{R}^{p-r}$ a transformação $T(\mathbf{x}) = \Sigma_{22}^{-1/2} \mathbf{x}$ e g_1 tal que $\int_0^\infty t^{r-1} g_1(t^2) dt < \infty$.

Demonstração. Ver Apêndice A.17. ■

No Corolário 3.4.4.1 mostramos que se a matriz $\Sigma(p \times p)$ é diagonal, então todas as distribuições marginais das distribuições Box-Cox elípticas apresentam distribuições Box-Cox simétricas.

Corolário 3.4.4.1. *Sejam $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\Sigma(p \times p) > 0$ e $\mathbf{Y} \sim BCE\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \Sigma; g)$, onde $\Sigma = \text{diag}\{\sigma_{11}, \dots, \sigma_{pp}\}$. Então $Y_k \sim BCS(\mu_k, \sigma_{kk}, \lambda_k; g_k)$, $k = 1, \dots, p$, sendo*

$$g_k(u) = \int_{A(\Sigma_{-k,-k}^{-1/2} \boldsymbol{\lambda}_{-k})} g(u + \mathbf{s}'_{-k} \mathbf{s}_{-k}) d\mathbf{s}_{-k}, \quad u \geq 0,$$

uma função tal que $\int_0^\infty g_k(t^2) dt < \infty$.

Demonstração. Basta considerar $\mathbf{Y}_1 = Y_k$, $\mathbf{Y}_2 = \mathbf{Y}_{-k}$, $\boldsymbol{\mu}_1 = \mu_k$, $\boldsymbol{\mu}_2 = \boldsymbol{\mu}_{-k}$, $\boldsymbol{\lambda}_1 = \lambda_k$, $\boldsymbol{\lambda}_2 = \boldsymbol{\lambda}_{-k}$, $\Sigma_{11} = \sigma_{kk}$ e $\Sigma_{22} = \Sigma_{-k,-k}$, $k = 1, \dots, p$, no Teorema 3.4.4. ■

No Teorema 3.4.4 mostramos que se $\mathbf{Y}' = (\mathbf{Y}'_1, \mathbf{Y}'_2) \sim BCE\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \Sigma; g)$, então o subvetor \mathbf{Y}_1 também tem distribuição Box-Cox elíptica se $\Sigma(p \times p) > 0$ é bloco diagonal. Note que, sob essa condição, \mathbf{Y}_1 apresenta distribuição na classe de distribuições Box-Cox elípticas, mas não na mesma família de distribuições de \mathbf{Y} , a menos que a f.g.d g_1 dada em (3.17) tenha a mesma estrutura da f.g.d g . Ainda que a condição estabelecida no Teorema 3.4.4 seja suficiente para que todos os membros na classe de distribuições Box-Cox elípticas apresentem distribuições marginais na mesma classe, nas distribuições log-normal e log- t multivariadas não é necessário impor essa condição para que as distribuições marginais pertençam à classe de distribuições log-elípticas. Além disso, as distribuições marginais nestas duas famílias pertencem à mesma família de distribuições, como mostramos no Teorema 3.4.5. Esse resultado está baseado no fato de que se $\mathbf{X}' = (\mathbf{X}'_1, \mathbf{X}'_2)$ tem distribuição elíptica, então a distribuição marginal de \mathbf{X}_1 também tem distribuição elíptica (Fang et al. [9]). Em geral, a distribuição marginal de \mathbf{X}_1 não pertence à mesma família de distribuições de \mathbf{X} , como ocorre, por exemplo, com a distribuição exponencial potência multivariada (Gómez et al. [24]),

mas nas distribuições normal e t multivariadas as distribuições marginais pertencem à mesma família de distribuições. Com efeito, sejam $\mathbf{X} \in \mathbb{R}^p$, $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ particionados como em (2.27). Se $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, então $\mathbf{X}_1 \sim N_r(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$. Se $\mathbf{X} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \tau)$, então $\mathbf{X}_1 \sim t_r(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}; \tau)$.

Teorema 3.4.5. *Sejam $\mathbf{Y} \in \mathbb{R}_+^p$, $\boldsymbol{\mu} \in \mathbb{R}_+^p$ e $\boldsymbol{\Sigma}(p \times p) > 0$ particionados como em (3.13). Seja ainda $\tau > 0$.*

1. *Se $\mathbf{Y} \sim LN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, então $\mathbf{Y}_1 \sim LN_r(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$*
2. *Se $\mathbf{Y} \sim Lt_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \tau)$, então $\mathbf{Y}_1 \sim Lt_r(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}; \tau)$.*

Demonstração. Ver Apêndice A.18. ■

No Teorema 3.4.6 mostramos que se $\mathbf{Y} = (\mathbf{Y}'_1, \mathbf{Y}'_2)'$ tem distribuição Box-Cox elíptica, então as distribuições condicionais de $\mathbf{Y}_1|\mathbf{Y}_2$ e $\mathbf{Y}_2|\mathbf{Y}_1$ também apresentem distribuição Box-Cox elíptica sob determinada condição.

Teorema 3.4.6. *Sejam $\mathbf{Y} \in \mathbb{R}_+^p$, $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ particionados como em (3.13) e tal que $\mathbf{Y} \sim BCE\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$. Seja $\boldsymbol{\mu}_1(\mathbf{w}_2) = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{w}_2 \in A(\boldsymbol{\lambda}_1)$, sendo $\mathbf{w}_2 = T_{(\lambda_2, \mu_2)}(\mathbf{y}_2)$, então $\mathbf{Y}_1|\mathbf{Y}_2 = \mathbf{y}_2 \sim BCE\ell_r(\boldsymbol{\delta}_1, \boldsymbol{\lambda}_1, \mathbf{D}_1\boldsymbol{\Sigma}_{11.2}\mathbf{D}_1; g_{q(\mathbf{w}_2)})$, em que $\boldsymbol{\delta}_1 = T_{(\lambda_1, \mu_1)}^{-1}(\boldsymbol{\mu}_1(\mathbf{w}_2))$ e*

$$\begin{aligned} \mathbf{D}_1 &= \mathbf{D}_{1+\mathbf{D}\lambda_1\mu_1(\mathbf{w}_2)}^{-1} \\ \boldsymbol{\Sigma}_{11.2} &= \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}, \\ g_{q(\mathbf{w}_2)}(u) &= g(u + q(\mathbf{w}_2)), \quad u \geq 0, \quad \text{com} \quad q(\mathbf{w}_2) = \mathbf{w}'_2\boldsymbol{\Sigma}_{22}^{-1}\mathbf{w}_2. \end{aligned} \tag{3.18}$$

Demonstração. Ver Apêndice A.19. ■

No Teorema 3.4.7 mostramos um resultado similar ao Teorema 2.4.7 para a classe de distribuições Box-Cox elípticas. Este teorema é uma caracterização da independência entre dois subvetores de um vetor aleatório com distribuição Box-Cox elíptica.

Teorema 3.4.7. *Sejam $\mathbf{Y} \in \mathbb{R}_+^p$, $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ particionados como em (3.13) e tal que $\mathbf{Y} \sim BCE\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$. Então, \mathbf{Y}_1 e \mathbf{Y}_2 são independentes se, e somente se, $\mathbf{Y} \sim BCN_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$ e $\boldsymbol{\Sigma}_{12} = \mathbf{0}$.*

Demonstração. Similar à prova do Teorema 2.4.7. ■

No Teorema 3.4.8 derivamos uma expressão que liga os momentos mistos de vetores aleatórios tendo distribuições Box-Cox elípticas.

Teorema 3.4.8. *Sejam $\mathbf{h} \in \mathbb{R}^p$, $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $\mathbf{Y} \sim BCE\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$. Então*

$$E\left(\prod_{k=1}^p Y_k^{h_k}\right) = \left(\prod_{k=1}^p \mu_k^{h_k}\right) E\left(\prod_{k=1}^p U_k^{h_k}\right), \quad (3.19)$$

sempre que $E\left(\prod_{k=1}^p U_k^{h_k}\right) < \infty$, sendo $\mathbf{U} = (U_1, \dots, U_p) \sim BCE\ell_p(\mathbf{1}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$.

Demonstração. Ver Apêndice A.20. ■

Pela Definição 3.3.1 podemos expressar o vetor aleatório \mathbf{U} que aparece em (3.19) como $\mathbf{U} = T_{(\boldsymbol{\lambda}, \mathbf{1})}^{-1}(\mathbf{W})$, em que $\mathbf{W} \sim TEL_p(\mathbf{0}, \boldsymbol{\Sigma}; A(\boldsymbol{\lambda}); g)$. Isto pode facilitar o cálculo de (3.19) através de integração de Monte Carlo. Usando o Algoritmo 2.5.1 geramos uma amostra aleatória de tamanho n do vetor aleatório $\mathbf{W} \sim TEL_p(\mathbf{0}, \boldsymbol{\Sigma}; A(\boldsymbol{\lambda}); g)$, chamemos $\mathbf{w}_1, \dots, \mathbf{w}_n$, em que $\mathbf{w}_i = (w_{i1}, \dots, w_{ip})$, $i = 1, \dots, n$, e calculamos

$$E\left(\prod_{k=1}^p Y_k^{h_k}\right) \approx \frac{1}{n} \sum_{i=1}^n \prod_{k=1}^p (\mu_k u_k(w_{ik}))^{h_k},$$

sendo $u_k(w_{ik}) = T_{(\lambda_k, 1)}^{-1}(w_{ik})$, $i = 1, \dots, n$; $k = 1, \dots, p$.

No Corolário 3.4.8.1 apresentamos uma interessante fórmula para o cálculo dos momentos mistos das distribuições log-elípticas. Neste caso a fórmula depende da função geradora de momentos (f.g.m) de um vetor aleatório com distribuição elíptica. Este resultado já tinha sido provado por Fang et al. [9, Seção 2.8.1].

Corolário 3.4.8.1. *Sejam $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $\mathbf{Y} \sim LE\ell_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$. Seja $\mathbf{h} \in \mathbb{R}^p$, então*

$$E\left(\prod_{k=1}^p Y_k^{h_k}\right) = \left(\prod_{k=1}^p \mu_k^{h_k}\right) M_{\mathbf{W}}(\mathbf{h}), \quad (3.20)$$

sempre que $M_{\mathbf{W}}$, a f.g.m de $\mathbf{W} \sim EL_p(\mathbf{0}, \boldsymbol{\Sigma}; g)$, exista.

Demonstração. Basta fazer $\boldsymbol{\lambda} = \mathbf{0}$ no Teorema 3.4.8. ■

Note que se uma família de distribuições elípticas apresenta forma fechada para sua f.g.m, então a fórmula para calcular os momentos mistos da respectiva família de distribuições log-elíptica também apresentará forma fechada. Por exemplo, se $\mathbf{Y} \sim LN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, então, usando (3.20), temos que

$$E\left(\prod_{k=1}^p Y_k^{h_k}\right) = \left(\prod_{k=1}^p \mu_k^{h_k}\right) \exp\left(\frac{1}{2} \mathbf{h}' \boldsymbol{\Sigma} \mathbf{h}\right). \quad (3.21)$$

No Corolário 3.4.8.2 mostramos uma relação entre matrizes de covariâncias envolvendo vetores aleatórios que têm distribuições Box-Cox elípticas.

Corolário 3.4.8.2. *Sejam $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $\mathbf{Y} \sim BCE\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$. A matriz de covariâncias de \mathbf{Y} , $\boldsymbol{\Sigma}_{\mathbf{Y}}$, satisfaz a relação*

$$\boldsymbol{\Sigma}_{\mathbf{Y}} = \mathbf{D}_{\boldsymbol{\mu}} \boldsymbol{\Sigma}_{\mathbf{U}} \mathbf{D}_{\boldsymbol{\mu}}, \quad (3.22)$$

em que $\boldsymbol{\Sigma}_{\mathbf{U}}$ é a matriz de covariâncias do vetor aleatório $\mathbf{U} \sim BCE\ell_p(\mathbf{1}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$. Além disso, as matrizes de correlação de \mathbf{Y} e \mathbf{U} são iguais, isto é,

$$\boldsymbol{\rho}_{\mathbf{Y}} = \boldsymbol{\rho}_{\mathbf{U}}. \quad (3.23)$$

Demonstração. Ver Apêndice A.21. ■

3.5 Interpretação dos parâmetros

Nesta seção apresentamos a interpretação dos parâmetros que determinam as distribuições Box-Cox elípticas.

Pela Definição 3.3.1 a distribuição de um vetor aleatório $\mathbf{Y} \sim BCE\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$ é construída a partir da transformação de Box-Cox estendida (3.5), levando em consideração a região de truncamento que ela induz. Isto faz com que as distribuições Box-Cox elípticas sejam caracterizadas através das distribuições elípticas truncadas. Nessa caracterização os vetores de parâmetros $\boldsymbol{\mu} \in \mathbb{R}_+^p$ e $\boldsymbol{\lambda} \in \mathbb{R}^p$ são introduzidos de maneira que os componentes μ_k e λ_k , $k = 1, \dots, p$, sejam parâmetros que transformem unicamente a variável Y_k e, portanto, esses parâmetros são características da distribuição de Y_k . As distribuições marginais das distribuições elípticas truncadas que caracterizam as distribuições Box-Cox elípticas estão associados através da matriz $\boldsymbol{\Sigma}(p \times p) > 0$. Assim, essa matriz também introduz associação entre as distribuições marginais das distribuições Box-Cox elípticas. Portanto, μ_k , $k = 1, \dots, p$, é parâmetro de escala (Teorema 3.2.1) da distribuição de Y_k ; λ_k , $k = 1, \dots, p$, é parâmetro de assimetria (transformação potência para simetria marginal) da distribuição de Y_k ; σ_{jk} , $j, k = 1, \dots, p$, é parâmetro de associação entre Y_j e Y_k .

Os parâmetros μ_k e σ_{kk} , $k = 1, \dots, p$, estão relacionados a percentis. Com o intuito de estabelecer essas relações, escrevemos a f.d.p marginal de Y_k , $k = 1, \dots, p$, como

$$f_{Y_k}(y_k) = \frac{g_{\boldsymbol{\Upsilon}_k}(s_k) \frac{y_k^{\lambda_k - 1}}{\sqrt{\sigma_{kk} \mu_k^{\lambda_k}}}}{\int_{I(\lambda_k \sqrt{\sigma_{kk}})} g_{\boldsymbol{\Upsilon}_k}(s_k) ds_k}, \quad s_k = \sigma_{kk}^{-1/2} T_{(\lambda_k, \mu_k)}(y_k), \quad y_k > 0, \quad (3.24)$$

em que $I(\lambda_k \sqrt{\sigma_{kk}})$ está definido segundo (3.4) e

$$g_{\boldsymbol{\Upsilon}_k}(u_k) = \int_{A(\boldsymbol{\lambda}_{-k})} g((1 + \boldsymbol{\Upsilon}_k \boldsymbol{\Upsilon}'_k) u_k^2 - 2 \boldsymbol{\Upsilon}_k \boldsymbol{\Psi}_k \mathbf{u}_{-k} u_k + \mathbf{u}'_{-k} \boldsymbol{\Psi}'_k \boldsymbol{\Psi}_k \mathbf{u}_{-k}) d\mathbf{u}_{-k}, \quad (3.25)$$

em que $u_k \in I(\lambda_k \sqrt{\sigma_{kk}})$, $\Psi_k = (\Sigma_{-k,-k} - \sigma_{kk}^{-1} \Sigma_{-k,k} \Sigma_{k,-k})^{-1/2}$ e $\Upsilon_k = \sigma_{kk}^{-1/2} \Sigma_{k,-k} \Psi_k$.

A f.d.p (3.24) pode ser construída a partir de uma variável aleatória definida em \mathbb{R} . Com efeito, note primeiramente que a função g_{Υ_k} dada em (3.25) pode ser definida em \mathbb{R} . Desta maneira, podemos definir uma variável aleatória $U_k \in \mathbb{R}$ através da f.d.p

$$f_{U_k}(u_k) = c_k g_{\Upsilon_k}(u_k), \quad u_k \in \mathbb{R}, \quad (3.26)$$

em que $c_k^{-1} = \int_{-\infty}^{\infty} g_{\Upsilon_k}(u_k) du_k$. A f.d.a de U_k é dada por

$$F_{U_k}(u_k) = c_k \int_{-\infty}^{u_k} g_{\Upsilon_k}(t_k) dt_k. \quad (3.27)$$

Definindo a variável aleatória S_k como a variável aleatória U_k truncada sobre o conjunto $I(\lambda_k \sqrt{\sigma_{kk}})$, temos que a f.d.p de S_k é dada por

$$f_{S_k}(s_k) = \frac{g_{\Upsilon_k}(s_k)}{\int_{I(\lambda_k \sqrt{\sigma_{kk}})} g_{\Upsilon_k}(s_k) ds_k}, \quad s_k \in I(\lambda_k \sqrt{\sigma_{kk}}). \quad (3.28)$$

Considerando a transformação $S_k = \sigma_{kk}^{-1/2} T_{(\lambda_k, \mu_k)}(Y_k)$, com Jacobiano

$$J(s_k \rightarrow y_k) = \frac{y_k^{\lambda_k - 1}}{\sqrt{\sigma_{kk} \mu_k^{\lambda_k}}},$$

obtemos a f.d.p de Y_k dada em (3.24).

Um caso de interesse da f.d.p (3.26) ocorre quando a integral que envolve a função g_{Υ_k} tem região de integração sobre \mathbb{R}^{p-1} . Nesse caso, a partir da relação

$$\det(\Sigma)^{-1/2} g_{\Upsilon_k}(u_k) = \frac{1}{\sqrt{\sigma_{kk}}} g_k(u_k^2), \quad u_k \in \mathbb{R}, \quad (3.29)$$

em que

$$g_k(u) = \int_{\mathbb{R}^{p-1}} g(u + \mathbf{w}'_{-k} \mathbf{w}_{-k}) d\mathbf{w}_{-k}, \quad u \geq 0, \quad (3.30)$$

segue que

$$f_{U_k}(u_k) = g_k(u_k^2), \quad u_k \in \mathbb{R}, \quad (3.31)$$

que corresponde à f.d.p de uma variável simétrica padrão com f.g.d dada por (3.30).

No Teorema 3.5.1 mostramos que todos os quantis das distribuições marginais univariadas das distribuições Box-Cox elípticas são proporcionais aos respectivos componentes do parâmetro μ .

Teorema 3.5.1. *Sejam $\mu \in \mathbb{R}_+^p$, $\lambda \in \mathbb{R}^p$, $\Sigma(p \times p) > 0$ e $Y \sim BCE\ell_p(\mu, \lambda, \Sigma; g)$. O α -quantil*

$y_{k,\alpha}$ de Y_k , $\alpha \in (0, 1)$, $k = 1, \dots, p$, satisfaz

$$y_{k,\alpha} = \begin{cases} \mu_k (1 + \lambda_k \sqrt{\sigma_{kk}} s_{k,\alpha})^{1/\lambda_k}, & \text{se } \lambda_k \neq 0, \\ \mu_k \exp(\sqrt{\sigma_{kk}} s_{k,\alpha}), & \text{se } \lambda_k = 0, \end{cases} \quad (3.32)$$

em que

$$s_{k,\alpha} = \begin{cases} F_{U_k}^{-1}(\alpha + (1 - \alpha)F_{U_k}(-1/\lambda_k \sqrt{\sigma_{kk}})), & \text{se } \lambda_k > 0, \\ F_{U_k}^{-1}((1 + \alpha)F_{U_k}(-1/\lambda_k \sqrt{\sigma_{kk}}) - 1), & \text{se } \lambda_k < 0, \\ F_{U_k}^{-1}(\alpha), & \text{se } \lambda_k = 0, \end{cases} \quad (3.33)$$

onde F_{U_k} é a f.d.a dada em (3.27).

Demonstração. Ver Apêndice A.22. ■

No Teorema 3.5.1 provamos que se $\mathbf{Y} \sim \text{BCEl}_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$, então todos os quantis de Y_k , $k = 1, \dots, p$, em particular, a mediana, são proporcionais a μ_k . Isto torna a classe de distribuições Box-Cox elípticas atrativa para modelagem de regressão (Capítulo 4). No Corolário 3.5.1.1 mostramos condições para que os quantis de Y_k , $k = 1, \dots, p$, de $\mathbf{Y} \sim \text{BCEl}_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$ sejam calculados em termos de quantis de distribuições simétricas padrão.

Corolário 3.5.1.1. *Sejam $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $\mathbf{Y} \sim \text{BCEl}_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$.*

1. *Se $\boldsymbol{\lambda} = \mathbf{0}$ (isto é $\mathbf{Y} \sim \text{LEl}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$), então o α -quantil $y_{k,\alpha}$ de Y_k , $\alpha \in (0, 1)$, $k = 1, \dots, p$, é dado por $y_{k,\alpha} = \mu_k \exp(\sqrt{\sigma_{kk}} s_{k,\alpha})$, em que $s_{k,\alpha}$ é o α -quantil de uma distribuição simétrica padrão com f.g.d dada por (3.30).*
2. *Se $\lambda_j \sqrt{\sigma_{jj}} \rightarrow 0$, $j = 1, \dots, p$, então o α -quantil $y_{k,\alpha}$ de Y_k , $\alpha \in (0, 1)$, $k = 1, \dots, p$, é dado por $y_{k,\alpha} = \mu_k \exp(\sqrt{\sigma_{kk}} s_{k,\alpha})$, em que $s_{k,\alpha}$ é o α -quantil de uma distribuição simétrica padrão com f.g.d dada por (3.30).*

Demonstração. Ver Apêndice A.23. ■

Como indicamos anteriormente, se um vetor aleatório tem distribuição na família normal ou t multivariada, então as distribuições marginais pertencem à mesma família de distribuições. Isso faz com que, sob as condições impostas no Corolário 3.5.1.1, os quantis de Y_k , $k = 1, \dots, p$, de um vetor aleatório \mathbf{Y} com distribuição nas famílias log-normal, log- t , Box-Cox normal ou Box-Cox t multivariadas possam ser calculados com facilidade através das distribuições normal padrão ou t , conforme o caso. Com efeito, se $\mathbf{Y} \sim \text{LN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ou $\mathbf{Y} \sim \text{BCN}_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma})$, com $\lambda_j \sqrt{\sigma_{jj}} \rightarrow 0$, $j = 1, \dots, p$, então o α -quantil $y_{k,\alpha}$ de Y_k , $\alpha \in (0, 1)$, $k = 1, \dots, p$, nesses casos é dado

por $y_{k,\alpha} = \mu_k \exp(\sqrt{\sigma_{kk}}s_\alpha)$, em que s_α é o α -quantil da distribuição $N(0, 1)$. Isto vem do fato de que a f.d.p (3.26) é calculada a partir de (3.29) e (3.30) como

$$f_{U_k}(u_k) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u_k^2}{2}\right), \quad u_k \in \mathbb{R},$$

quando é considerada a f.g.d (2.5). De forma semelhante, se $\mathbf{Y} \sim Lt_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \tau)$ ou $\mathbf{Y} \sim BCt_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; \tau)$, com $\lambda_j \sqrt{\sigma_{jj}} \rightarrow 0, j = 1, \dots, p$, então o α -quantil $y_{k,\alpha}$ de $Y_k, \alpha \in (0, 1), k = 1, \dots, p$, nesses casos é dado por $y_{k,\alpha} = \mu_k \exp(\sqrt{\sigma_{kk}}s_\alpha)$, em que s_α é o α -quantil da distribuição t com τ graus de liberdade.

Seja $\mathbf{Y} \sim BCEl_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$. O Corolário 3.5.1.1 garante as relações a seguir dos parâmetros $\mu_k, k = 1, \dots, p$: se $\boldsymbol{\lambda} = \mathbf{0}$ (isto é $\mathbf{Y} \sim LEl(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$), então $\mu_k = y_{k,1/2}$. Também, se $\boldsymbol{\lambda} \approx \mathbf{0}$ ou $\lambda_j \sqrt{\sigma_{jj}} \approx 0, j = 1, \dots, p$, então $\mu_k \approx y_{k,1/2}$. Portanto, nesses casos, $\boldsymbol{\mu}$ é igual ou aproximadamente igual ao vetor das medianas das distribuições marginais de $\mathbf{Y} \sim BCEl(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}; g)$.

Um coeficiente de variação baseado nos percentis (Rigby e Stasinopoulos [8]) para $Y_k, k = 1, \dots, p$, é definido como

$$CV_{Y_k} = \frac{3(y_{k,3/4} - y_{k,1/4})}{4y_{k,1/2}}. \quad (3.34)$$

O Corolário 3.5.1.1 garante as relações a seguir dos parâmetros $\sigma_{kk}, k = 1, \dots, p$, de $\mathbf{Y} \sim BCEl_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$: se $\boldsymbol{\lambda} = \mathbf{0}$ (isto é $\mathbf{Y} \sim LEl(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$), então $CV_{Y_k} = 1,5 \sinh(\sqrt{\sigma_{kk}}s_{k,3/4})$, em que $s_{k,3/4}$ é o terceiro quartil de uma distribuição simétrica padrão com f.g.d dada por (3.30). Se $\boldsymbol{\lambda} \approx \mathbf{0}$ ou $\lambda_j \sqrt{\sigma_{jj}} \approx 0, j = 1, \dots, p$, então $CV_{Y_k} \approx 1,5 \sinh(\sqrt{\sigma_{kk}}s_{k,3/4})$. Assim, nesses casos, CV_{Y_k} depende de σ_{kk} através da função seno hiperbólico, que é uma função crescente. Portanto, σ_{kk} pode ser visto como um parâmetro de dispersão relativa da distribuição de Y_k . Em particular, se $\mathbf{Y} \sim LN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ($\mathbf{Y} \sim Lt_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \tau)$), então $CV_{Y_k} = 1,5 \sinh(\sqrt{\sigma_{kk}}s_{3/4})$, em que $s_{3/4}$ é o terceiro quartil da distribuição normal padrão (t com τ graus de liberdade). Se $\mathbf{Y} \sim BCN_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma})$ ($\mathbf{Y} \sim BCt_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; \tau)$), com $\boldsymbol{\lambda} \approx \mathbf{0}$ ou $\lambda_j \sqrt{\sigma_{jj}} \approx 0, j = 1, \dots, p$, então $CV_{Y_k} \approx 1,5 \sinh(\sqrt{\sigma_{kk}}s_{3/4})$, em que $s_{3/4}$ é o terceiro quartil da distribuição normal padrão (t com τ graus de liberdade).

Os parâmetros $\sigma_{jk}, j \neq k$, induzem associação entre as distribuições marginais das distribuições Box-Cox elípticas. As Figuras 2.1(a), 2.1(b), 2.1(c) mostram um exemplo de como a associação entre as distribuições marginais de uma distribuição Box-Cox t bivariada é alterada segundo os valores de σ_{12} . No entanto, não é fácil estabelecer a maneira em que esses parâmetros modelam a associação na classe, pois depende de cada família considerada. Por exemplo, se $\mathbf{Y} \sim BCN_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma})$, com $\boldsymbol{\lambda} \approx \mathbf{0}$, então de (3.21) podemos mostrar que

$$\text{Cov}(Y_j, Y_k) \approx \mu_j \mu_k \exp\left[\frac{1}{2}(\sigma_{jj} + \sigma_{kk})\right] [\exp(\sigma_{jk}) - 1], \quad j \neq k.$$

Portanto, se $\sigma_{jk} > 0$, então $\text{Cov}(Y_j, Y_k) > 0$; se $\sigma_{jk} < 0$, então $\text{Cov}(Y_j, Y_k) < 0$; e se $\sigma_{jk} = 0$, então $\text{Cov}(Y_j, Y_k) = 0$.

3.6 Estimação de máxima verossimilhança

Neste seção apresentamos alguns aspectos teóricos e computacionais associados ao método de máxima verossimilhança, utilizado na estimação dos parâmetros das distribuições Box-Cox elípticas.

Para estimar os parâmetros das distribuições Box-Cox elípticas propomos o método de máxima verossimilhança. Sejam $\mathbf{y}_1, \dots, \mathbf{y}_n$ os valores observados de uma amostra aleatória $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ de $\mathbf{Y} \sim \text{BCE}_{\ell_p}(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$, sendo $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{ip})$, $i = 1, \dots, n$. Seja $\boldsymbol{\eta} = (\eta_1, \dots, \eta_q)$ o vetor de parâmetros extras induzidos pela f.g.d g . Os estimadores de máxima verossimilhança (EMV) de $\boldsymbol{\mu}$, $\boldsymbol{\lambda}$, $\boldsymbol{\Sigma}$ e $\boldsymbol{\eta}$ denotados por $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\lambda}}$, $\hat{\boldsymbol{\Sigma}}$ e $\hat{\boldsymbol{\eta}}$, respectivamente, serão tais que maximizem o logaritmo da função de verossimilhança

$$\ell = \sum_{i=1}^n \ell_i, \quad (3.35)$$

sendo ℓ_i , $i = 1, \dots, n$, o logaritmo da função de verossimilhança de \mathbf{y}_i , dado por

$$\ell_i = -\log \left\{ \int_{A(\boldsymbol{\lambda})} g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) d\mathbf{w} \right\} + \log \{ g(\mathbf{w}_i'\boldsymbol{\Sigma}^{-1}\mathbf{w}_i) \} + \sum_{k=1}^p (\lambda_k - 1) \log y_{ik} - \sum_{k=1}^p \lambda_k \log \mu_k,$$

em que $\mathbf{w}_i = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}(\mathbf{y}_i)$. Não existe forma fechada para $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\lambda}}$, $\hat{\boldsymbol{\Sigma}}$ e $\hat{\boldsymbol{\eta}}$, mas estes estimadores podem ser obtidos através de métodos otimização implementados em programas computacionais.

Sejam $\boldsymbol{\mu}^{(0)}$, $\boldsymbol{\lambda}^{(0)}$, $\boldsymbol{\Sigma}^{(0)}$ e $\boldsymbol{\eta}^{(0)}$ os valores iniciais no processo de estimação de $\boldsymbol{\mu}$, $\boldsymbol{\lambda}$, $\boldsymbol{\Sigma}$ e $\boldsymbol{\eta}$, respectivamente. Para a escolha de $\mu_k^{(0)}$, $\lambda_k^{(0)}$ e $\sigma_{kk}^{(0)}$, $k = 1, \dots, p$, sugerimos as estimativas geradas ao ajustar a distribuição Box-Cox simétrica no k -ésimo componente de \mathbf{Y} , isto é, os valores estimados dos parâmetros de $Y_k \sim \text{BCS}(\mu_k, \lambda_k, \sqrt{\sigma_{kk}}; g)$. Como valores iniciais para os parâmetros σ_{jk} , sugerimos $\sigma_{jk}^{(0)} = 0$, $j \neq k$. A proposta para os valores iniciais dos parâmetros extras, $\eta_j^{(0)}$, $j = 1, \dots, q$, dependerá da família de distribuições considerada na classe. Por exemplo, para a distribuição Box-Cox t multivariada propomos como valor inicial do número de graus de liberdade, $\tau^{(0)}$, o número de graus de liberdade estimado no ajuste de uma distribuição t multivariada no vetor $\mathbf{X} = T_{(\boldsymbol{\lambda}^{(0)}, \boldsymbol{\mu}^{(0)})}(\mathbf{Y})$. Note que não é considerado o truncamento que induz a transformação de Box-Cox estendida nesse ajuste.

O processo de estimação que envolve as distribuições Box-Cox elípticas é complicado devido, principalmente, à necessidade de calcular eficientemente a integral $\int_{A(\boldsymbol{\lambda})} g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) d\mathbf{w}$ que aparece em cada ℓ_i , $i = 1, \dots, n$. Esta integral depende da

complexidade e estrutura da f.g.d g e deve ser calculada sobre o conjunto $A(\boldsymbol{\lambda})$. Assim, o vetor de parâmetros extras $\boldsymbol{\eta}$, a matriz $\boldsymbol{\Sigma}$ e o vetor $\boldsymbol{\lambda}$ estão envolvidos através desta integral no processo de estimação. Genz e Bretz [22] propõem algoritmos para calcular esse tipo de integrais quando g é a f.g.d das famílias normal e t multivariadas dadas em (2.5) e (2.8), respectivamente. Esses algoritmos estão implementados no pacote `mvtnorm` (Genz et al. [21]) do software livre R [10] através das rotinas `pmvnorm` e `pmvt`, que permitem calcular as f.d.a das distribuições normal e t multivariadas sobre retângulos. No entanto, a rotina `pmvt` admite cálculos apenas com graus de liberdade inteiros. No Apêndice B disponibilizamos a função `pmvt1`, que permite calcular a f.d.a da distribuição t multivariada sobre retângulos com graus de liberdade no conjunto dos reais positivos. O cálculo eficiente dessas integrais para outras f.g.d é problema em aberto. No Apêndice B também disponibilizamos a programação em R que permite obter os EMV nas famílias de distribuições Box-Cox normal e Box-Cox t multivariadas por meio das funções `mlemvbcnorm` e `mlemvbct`, respectivamente. Os métodos iterativos disponíveis em nossos programas que permite maximizar (3.35) para estas duas famílias são Nelder-Mead, BFGS, CG, SANN e L-BFGS-B (ver [38]). Esses métodos estão implementados na rotina `optim` do pacote `stats` do R.

Observação 3.6.1. *Note que na classe de distribuições log-elípticas ($\boldsymbol{\lambda} = \mathbf{0}$) não é necessário calcular a integral $\int_{A(\boldsymbol{\lambda})} g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) d\mathbf{w}$, o que facilita o processo de estimação nessa classe. De fato, nesse caso, o logaritmo da função de verossimilhança é dado por (3.35), em que cada ℓ_i , $i = 1, \dots, n$, é*

$$\ell_i = -\frac{1}{2} \log(\det(\boldsymbol{\Sigma})) + \log\{g(\mathbf{w}'_i\boldsymbol{\Sigma}^{-1}\mathbf{w}_i)\} - \sum_{k=1}^p \log y_{ik},$$

sendo $\mathbf{w}_i = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}(\mathbf{y}_i)$. Nesse caso os parâmetros a estimar são $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ e $\boldsymbol{\eta}$.

Com o objetivo de verificar se a metodologia de estimação é adequada realizamos estudos de simulação para estimar parâmetros das distribuições log-normal, log- t , Box-Cox normal e Box-Cox t bivariadas a partir de amostras aleatórias com diferentes tamanhos. As amostras aleatórias de tamanho n de $\mathbf{Y} \sim \text{BCE}\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$ são geradas de acordo com o Algoritmo 3.6.1, o qual foi implementado no R para as famílias de distribuições Box-Cox normal e Box-Cox t multivariadas através das funções `rmvbcnorm` e `rmvbct`, respectivamente (ver Apêndice B).

Algoritmo 3.6.1.

1. A partir do Algoritmo 2.5.1 gerar uma amostra aleatória de tamanho n do vetor aleatório $\mathbf{W} \sim \text{TE}\ell_p(\mathbf{0}, \boldsymbol{\Sigma}; A(\boldsymbol{\lambda}); g)$, chamemos $\mathbf{w}_1, \dots, \mathbf{w}_n$.
2. Pela Definição 3.3.1 temos que $\mathbf{y}_1 = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}^{-1}(\mathbf{w}_1), \dots, \mathbf{y}_n = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}^{-1}(\mathbf{w}_n)$ é uma amostra aleatória de $\mathbf{Y} \sim \text{BCE}\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$.

Conduzimos quatro estudos de simulação considerando $N = 5000$ réplicas para amostras de tamanho $n \in \{125, 250, 500\}$ geradas a partir do vetor aleatório \mathbf{Y} tendo as distribuições a seguir: $\text{LN}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\text{Lt}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \tau)$, $\text{BCN}_2(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma})$ e $\text{BCt}_2(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; \tau)$. Os parâmetros considerados em cada estudo de simulação aparecem na Tabela 3.1.

Tabela 3.1: Valores dos parâmetros nos estudos de simulação.

Estudo de simulação	Distri- buição	Parâmetros verdadeiros							
		μ_1	μ_2	λ_1	λ_2	σ_{11}	σ_{12}	σ_{22}	τ
1	LN_2	8	8	–	–	0,8	–0,5	1	–
2	Lt_2	7	10	–	–	1,2	0,6	1,4	5
3	BCN_2	20	15	0,4	0,3	0,4	0,1	0,3	–
4	BCt_2	5	4	–1	0,5	0,6	0,2	0,8	6

O cálculo dos valores iniciais $\mu_k^{(0)}$, $\lambda_k^{(0)}$ e $\sigma_{kk}^{(0)}$, $k = 1, \dots, p$, nos quatro estudo de simulação, foi feito ajustando as distribuições log-normal, log- t , Box-Cox normal e Box-Cox t , conforme o caso, utilizando a rotina `gamlss` do pacote com o mesmo nome (Rigby e Stasinopoulos [39]) do R. O cálculo do valor inicial do número de graus de liberdade, $\tau^{(0)}$, para as distribuições Box-Cox t e log- t bivariadas, foi feito ajustando a distribuição t multivariada no vetor $\mathbf{X} = T_{(\boldsymbol{\lambda}^{(0)}, \boldsymbol{\mu}^{(0)})}(\mathbf{Y})$ através da rotina `selm` do pacote `sn` (Azzalini [40]) do R. Em cada estudo de simulação utilizamos o método de otimização BFGS, que não permite restrições sobre os problemas de otimização em que seja aplicado. Os espaços paramétricos das distribuições consideradas nos estudos de simulação têm restrições determinadas por $\boldsymbol{\mu} \in \mathbb{R}_+^2$, $\boldsymbol{\lambda} \in \mathbb{R}^2$, $\boldsymbol{\Sigma}(2 \times 2) > 0$ e $\tau > 0$, conforme o caso. Por meio de reparametrizações removemos essas restrições e, assim, podemos aplicar o método BFGS[†]. As reparametrizações consideradas em cada estudo de simulação, conforme o caso, são mostradas a seguir:

$$\begin{aligned}
 \mu_k &= \exp(\xi_k), \quad \xi_k \in \mathbb{R}, \quad k = 1, 2, \\
 \boldsymbol{\Sigma} &= \mathbf{L}\mathbf{L}', \\
 \tau &= \exp(\nu), \quad \nu \in \mathbb{R},
 \end{aligned} \tag{3.36}$$

onde $\mathbf{L}(2 \times 2)$ é uma matriz triangular inferior. A fatoração de $\boldsymbol{\Sigma}(2 \times 2)$ dada em (3.36) é denominada decomposição de Cholesky (Seber [41]). Essa fatoração garante que a matriz $\boldsymbol{\Sigma}(2 \times 2)$ seja definida positiva durante o processo de estimação. Note que não é necessário considerar reparametrizações sobre o vetor $\boldsymbol{\lambda}$, já que está definido em \mathbb{R}^2 .

[†]Para utilizar os métodos Nelder-Mead, CG e SANN também é necessário considerar reparametrizações. O método L-BFGS-B permite impor restrições intervalares para cada parâmetro, que seria útil em nosso caso exceto para a matriz $\boldsymbol{\Sigma}$ que deve ser definida positiva. A decomposição de Cholesky da matriz $\boldsymbol{\Sigma}$ fornece uma reparametrização útil para poder aplicar os métodos de otimização disponíveis em nossos programas.

Sejam $\hat{\theta}_1, \dots, \hat{\theta}_N$ os valores ordenados das estimativas de um parâmetro θ num estudo de simulação. Denotamos por $M(\hat{\theta})$ a mediana de $\{\hat{\theta}_1, \dots, \hat{\theta}_N\}$. O viés mediano (MB, siglas em inglês) das estimativas, denotado por $MB(\hat{\theta})$, é dado por $MB(\hat{\theta}) = M(\hat{\theta}) - \theta$. O desvio absoluto mediano (MAD, siglas em inglês) das estimativas, denotado por $MAD(\hat{\theta})$, é definido como a mediana de $\{|\hat{\theta}_1 - M(\hat{\theta})|, \dots, |\hat{\theta}_N - M(\hat{\theta})|\}$. Também denotamos por $IQR(\hat{\theta})$ a amplitude interquartil (IQR, siglas em inglês) de $\{\hat{\theta}_1, \dots, \hat{\theta}_N\}$. Nas Tabelas 3.2, 3.3, 3.4 e 3.5 apresentamos as medianas das estimativas e os valores dos vieses medianos, desvios absolutos medianos e amplitudes interquartis das estimativas dos parâmetros dos estudos de simulação 1, 2, 3 e 4, respectivamente. Estas medidas são robustas a textitoutliers extremos (muito altos) das estimativas do parâmetro de graus de liberdade, que correspondem a ajustes das distribuições Box-Cox normal ou log-normal bivariadas, conforme o caso. Note que em todos os casos os valores destas medidas indicam bom comportamento dos estimadores, já que os vieses medianos são próximos de zero e os valores dos desvios medianos, desvios absolutos medianos e amplitudes interquartis diminuem quando n aumenta.

3.7 Aplicações a dados reais

Nesta seção apresentamos aplicações a dados reais através de ajustes das distribuições Box-Cox elípticas a conjuntos de dados nutricionais e antropométricos.

Inicialmente analisamos os consumos de nutrientes em pessoas idosas considerando uma amostra coletada no ano 2011 por meio da aplicação de recordatório 24 horas (R24h)[‡]. Examinamos os consumos de vitaminas B2 (em mg), B3 (em mg), B12 (em mcg) e D (em mcg) obtidos do primeiro R24h em 136 homens idosos. As estatísticas descritivas e as correlações amostrais das variáveis consideradas são apresentadas nas Tabelas 3.6 e 3.7, respectivamente. Na Figura 3.2 apresentamos a matriz de bagplots (Rousseeuw et al. [42]) de duplas de variáveis nutricionais, com os boxplots de cada variável na diagonal. O bagplot é uma versão bivariada do boxplot, é conformado por um *bag* (região da cor cinza escuro) que contém 50% das observações e no seu interior está localizado o dado com maior profundidade (ponto de cor vermelho). O *loop* (região da cor cinza claro) contém as observações que não pertencem ao *bag* e estão dentro da *fence* (linha que separa os dados discrepantes, indicados com pontos vermelhos). O bagplot permite visualizar a localização (dado com maior profundidade), dispersão (tamanho do *bag*), correlação (orientação do *bag*), assimetria (forma do *bag* e do *loop*) e dados discrepantes (pontos fora da *fence*). A Figura 3.2 indica que as duplas de variáveis nutricionais têm várias observações bem discrepantes, são positivamente correlacionadas e apresentam assimetria multivariada.

[‡]Este conjunto de dados faz parte do projeto de pesquisa “Avaliação da adequação nutricional na terceira idade” de José Eduardo Corrente, processo FAPESP 08/10261-8.

Tabela 3.2: Valores da mediana (M), viés mediano (MB), desvio absoluto mediano (MAD) e amplitude interquartil (IQR) associados ao estudo de simulação 1.

n		$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_{11}$	$\hat{\sigma}_{12}$	$\hat{\sigma}_{22}$
125	M	7,98	8,00	0,78	-0,48	0,97
	MB	-0,02	0,00	-0,02	0,02	-0,03
	MAD	0,59	0,66	0,07	0,07	0,09
	IQR	1,19	1,33	0,15	0,14	0,18
LN ₂	M	7,99	8,00	0,79	-0,49	0,99
	MB	-0,01	0,00	-0,01	0,01	-0,01
	MAD	0,42	0,47	0,05	0,05	0,07
	IQR	0,85	0,93	0,10	0,10	0,13
500	M	8,00	8,00	0,80	-0,50	1,00
	MB	0,00	0,00	0,00	0,00	0,00
	MAD	0,30	0,33	0,04	0,04	0,05
	IQR	0,60	0,66	0,08	0,07	0,09

Tabela 3.3: Valores da mediana (M), viés mediano (MB), desvio absoluto mediano (MAD) e amplitude interquartil (IQR) associados ao estudo de simulação 2.

n		$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_{11}$	$\hat{\sigma}_{12}$	$\hat{\sigma}_{22}$	$\hat{\tau}$
125	M	7,01	10,02	1,20	0,59	1,39	5,25
	MB	0,01	0,02	0,00	-0,01	-0,01	0,25
	MAD	0,97	1,45	0,22	0,17	0,26	1,76
	IQR	1,31	1,98	0,30	0,22	0,35	2,63
Lt ₂	M	6,99	9,99	1,20	0,60	1,40	5,15
	MB	-0,01	-0,01	0,00	0,00	0,00	0,15
	MAD	0,69	1,05	0,16	0,12	0,19	1,18
	IQR	0,93	1,43	0,22	0,16	0,25	1,66
500	M	7,00	10,00	1,20	0,60	1,40	5,07
	MB	0,00	0,00	0,00	0,00	0,00	0,07
	MAD	0,46	0,72	0,12	0,08	0,13	0,82
	IQR	0,62	0,97	0,16	0,11	0,18	1,13

Tabela 3.4: Valores da mediana (M), viés mediano (MB), desvio absoluto mediano (MAD) e amplitude interquartil (IQR) associados ao estudo de simulação 3.

n		$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\sigma}_{11}$	$\hat{\sigma}_{12}$	$\hat{\sigma}_{22}$	
BCN ₂	125	M	19,83	14,94	0,41	0,30	0,40	0,10	0,29
		MB	-0,17	-0,06	0,01	0,00	0,00	0,00	-0,01
		MAD	0,91	0,58	0,08	0,09	0,04	0,02	0,03
		IQR	1,80	1,15	0,15	0,17	0,08	0,05	0,05
	250	M	19,93	14,96	0,40	0,30	0,40	0,10	0,30
		MB	-0,07	-0,04	0,00	0,00	0,00	0,00	0,00
		MAD	0,63	0,41	0,05	0,06	0,03	0,02	0,02
		IQR	1,26	0,82	0,10	0,12	0,05	0,03	0,04
	500	M	19,98	14,99	0,40	0,30	0,40	0,10	0,30
		MB	-0,02	-0,01	0,00	0,00	0,00	0,00	0,00
		MAD	0,45	0,29	0,04	0,04	0,02	0,01	0,01
		IQR	0,90	0,58	0,07	0,09	0,04	0,02	0,03

Tabela 3.5: Valores da mediana (M), viés mediano (MB), desvio absoluto mediano (MAD) e amplitude interquartil (IQR) associados ao estudo de simulação 4.

n		$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\sigma}_{11}$	$\hat{\sigma}_{12}$	$\hat{\sigma}_{22}$	$\hat{\tau}$	
BC t_2	125	M	5,08	3,99	-1,04	0,49	0,66	0,26	0,85	8,89
		MB	0,08	-0,01	-0,04	-0,01	0,06	0,06	0,05	2,89
		MAD	0,72	0,42	0,25	0,13	0,31	0,16	0,16	4,62
		IQR	2,96	0,84	0,51	0,27	2,80	0,61	0,36	18,42
	250	M	5,05	4,00	-1,02	0,49	0,64	0,23	0,82	7,48
		MB	0,05	0,00	-0,02	-0,01	0,04	0,03	0,02	1,48
		MAD	0,50	0,29	0,19	0,10	0,24	0,10	0,11	2,70
		IQR	1,46	0,58	0,38	0,20	0,96	0,28	0,24	6,60
	500	M	5,07	4,01	-1,02	0,49	0,65	0,22	0,82	7,07
		MB	0,07	0,01	-0,02	-0,01	0,05	0,02	0,02	1,07
		MAD	0,40	0,20	0,13	0,07	0,20	0,07	0,08	1,67
		IQR	0,94	0,40	0,27	0,15	0,56	0,17	0,16	3,74

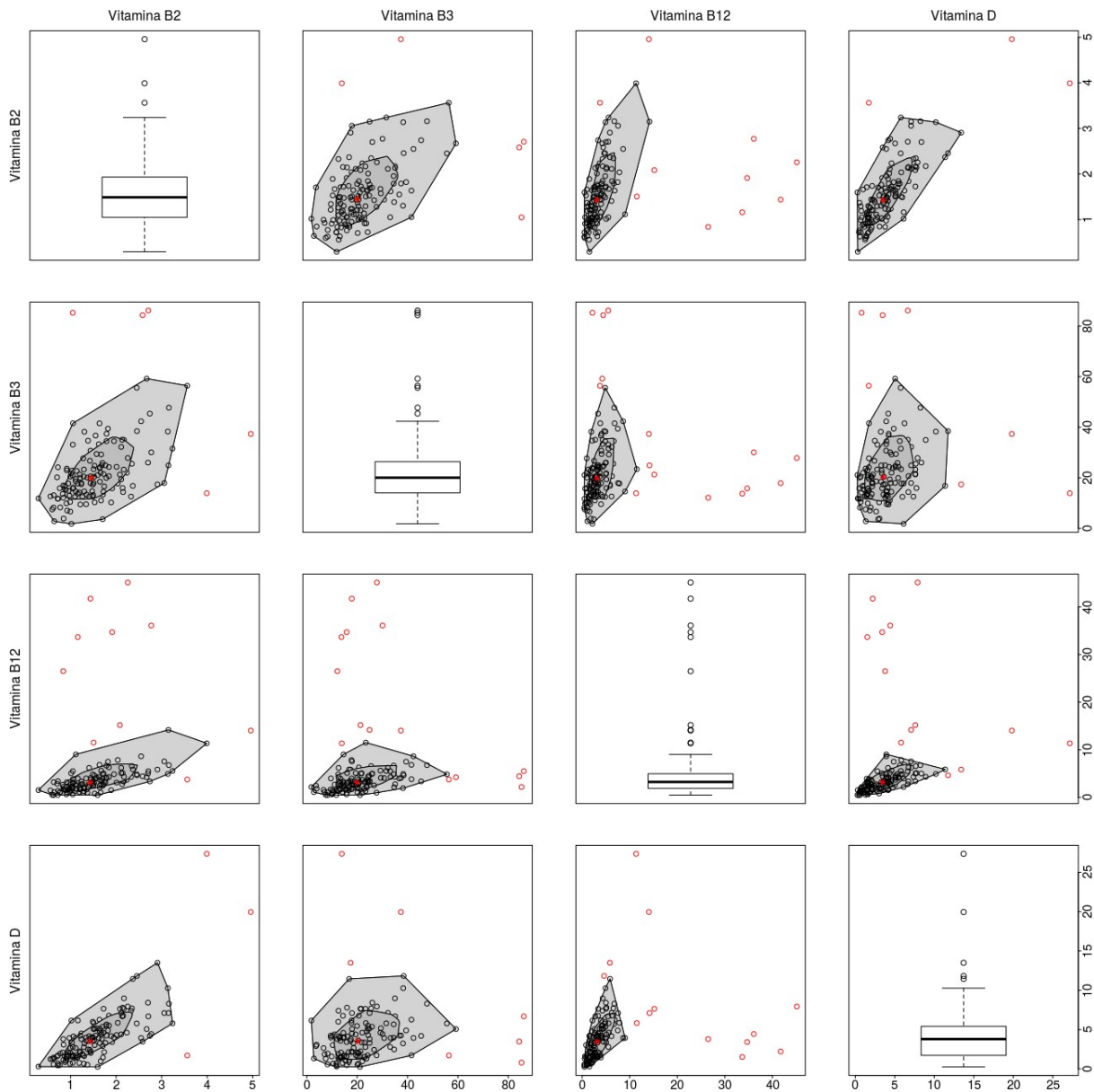


Figura 3.2: Matriz de bagplots das variáveis nutricionais.

Tabela 3.6: Estatísticas descritivas das variáveis nutricionais.

Variável	Min.	Q ₁	Mediana	Q ₃	Max.	CV [¶]	Média	SD
vit. B2	0,29	1,05	1,49	1,92	4,96	0,44	1,60	0,75
vit. B3	1,84	14,14	19,99	26,20	86,13	0,45	22,89	14,27
vit. B12	0,48	1,88	3,24	4,94	45,12	0,71	5,13	7,32
vit. D	0,26	1,73	3,80	5,39	27,35	0,72	4,16	3,51

[¶]Coeficiente de variação baseado em percentis definido em (3.34) (Rigby e Stasinopoulos [8]).

Tabela 3.7: Correlações amostrais entre as variáveis nutricionais.

Variável	vit. B2	vit. B3	vit. B12	vit. D
vit. B2	1,00	0,46	0,28	0,76
vit. B3	0,46	1,00	0,06	0,16
vit. B12	0,28	0,06	1,00	0,23
vit. D	0,76	0,16	0,23	1,00

Na Tabela 3.8 apresentamos os valores do AIC para avaliar a qualidade do ajuste de várias distribuições bivariadas em duplas de variáveis nutricionais. Consideramos ajustes de distribuições Box-Cox normal e Box-Cox t marginalmente independentes e das distribuições skew normal (SN_p), skew- t^{\S} (St_p), log-normal, log- t , Box-Cox normal e Box-Cox t bivariadas. Note que a distribuição Box-Cox t bivariada apresenta o melhor ajuste nos consumos das duplas de vitaminas B2-D, B3-D e B12-D, enquanto que a distribuição log- t bivariada apresenta melhor ajuste nos consumos das duplas de vitaminas B2-B3, B2-B12 e B3-B12. As distribuições Box-Cox t e log- t bivariadas levam em consideração a associação existente entre as variáveis, caso que não é considerado ao ajustar distribuições Box-Cox t marginalmente independentes. Além disso, a presença de observações discrepantes (Figura 3.2) faz com que as distribuições Box-Cox t e log- t apresentem melhor ajuste do que as distribuições log-normal e Box-Cox normal bivariadas. Na Tabela 3.9 mostramos os valores das estimativas (e erros padrão) dos parâmetros correspondentes aos modelos que apresentam o melhor ajuste na Tabela 3.8. Para o cálculo dos erros padrão das estimativas utilizamos o método delta multivariado (Sen et al. [45]), já que usamos as reparametrizações dadas em (3.36).

Os valores das medianas e os coeficientes de variação baseados em percentis mostrados na Tabela 3.6 obedecem às interpretações dos parâmetros estimados (Seção 3.5). Como a distribuição ajustada nos consumos da dupla de vitaminas B2-B3 é log- t bivariada, então $\hat{\mu}_1 = 1,45$ e $\hat{\mu}_2 = 19,91$ são aproximadamente iguais aos consumos medianos de vitaminas B2 e B3, respectivamente. Também, os coeficientes de variação baseados em percentis dos consumos de vitaminas B2 e B3 são aproximadamente iguais às quantidades $1,5 \sinh(\sqrt{\hat{\sigma}_{11}} s_{3/4})$ e $1,5 \sinh(\sqrt{\hat{\sigma}_{22}} s_{3/4})$, respectivamente; sendo $s_{3/4}$ o terceiro quartil da distribuição t com 6,22 graus de liberdade. Devido a que $\hat{\sigma}_{11} = 0,16$ e $\hat{\sigma}_{22} = 0,23$, então a dispersão relativa do consumo de vitamina B3 é levemente maior do que a do consumo de vitamina B2. Para os consumos da dupla de vitaminas B12-D a distribuição ajustada é Box-Cox t bivariada com $\hat{\lambda}_1 \sqrt{\hat{\sigma}_{11}} = -0,13$ e $\hat{\lambda}_2 \sqrt{\hat{\sigma}_{22}} = 0,07$, que são valores próximos de zero. Portanto, $\hat{\mu}_1 = 3,10$ e $\hat{\mu}_2 = 3,42$ são aproximadamente iguais aos consumos medianos de vitaminas B12 e D, respectivamente. Além

[§]As distribuições skew normal e skew- t multivariadas fazem parte da classe de distribuições skew-elípticas (Branco e Dey [43], Azzalini [44]).

Tabela 3.8: Valores do AIC das distribuições bivariadas ajustadas a duplas de variáveis nutricionais.

Variáveis	Ind. BCN ₁	Ind. BC _{t₁}	SN ₂	St ₂	LN ₂	L _{t₂}	BCN ₂	BC _{t₂}
(vit. B2, vit. B3)	1328,90	1323,15	1342,54	1299,81	1291,88	1279,44	1289,57	1281,66
(vit. B2, vit. B12)	945,13	944,11	1092,86	938,18	900,99	883,24	895,20	885,68
(vit. B2, vit. D)	914,99	918,08	851,08	813,67	814,27	802,72	803,79	796,84
(vit. B3, vit. B12)	1720,67	1710,27	1904,91	1732,81	1710,38	1689,32	1702,24	1690,79
(vit. B3, vit. D)	1690,52	1684,75	1793,66	1698,34	1689,65	1679,20	1683,20	1676,79
(vit. B12, vit. D)	1306,76	1305,23	1506,20	1272,96	1250,09	1226,64	1233,10	1223,88

Tabela 3.9: Estimativas (e erros padrão) dos parâmetros dos modelos ajustados a duplas de variáveis nutricionais.

Variáveis	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\sigma}_{11}$	$\hat{\sigma}_{12}$	$\hat{\sigma}_{22}$	$\hat{\tau}$
(vit. B2, vit. B3)	1,45 (0,05)	19,91 (0,90)	-	-	0,16 (0,02)	0,10 (0,02)	0,23 (0,04)	6,22 (2,18)
(vit. B2, vit. B12)	1,46 (0,06)	3,10 (0,20)	-	-	0,15 (0,02)	0,16 (0,03)	0,43 (0,08)	4,57 (1,36)
(vit. B2, vit. D)	1,45 (0,06)	3,42 (0,23)	0,19 (0,17)	0,31 (0,11)	0,16 (0,03)	0,22 (0,04)	0,48 (0,08)	7,96 (3,57)
(vit. B3, vit. B12)	20,10 (0,91)	3,13 (0,20)	-	-	0,20 (0,04)	0,13 (0,03)	0,42 (0,08)	3,96 (1,12)
(vit. B3, vit. D)	19,86 (0,94)	3,30 (0,23)	0,15 (0,14)	0,24 (0,12)	0,24 (0,04)	0,12 (0,03)	0,51 (0,08)	7,42 (3,03)
(vit. B12, vit. D)	3,10 (0,20)	3,42 (0,24)	-0,19 (0,10)	0,15 (0,11)	0,45 (0,08)	0,31 (0,06)	0,47 (0,08)	5,50 (1,99)

disso, os coeficientes de variação baseados em percentis dos consumos de vitaminas B12 e D são aproximadamente iguais a $1,5 \sinh(\sqrt{\hat{\sigma}_{11}}s_{3/4})$ e $1,5 \sinh(\sqrt{\hat{\sigma}_{22}}s_{3/4})$, respectivamente; em que $s_{3/4}$ o terceiro quartil da distribuição t com 5,5 graus de liberdade. Como $\hat{\sigma}_{11} = 0,45$ e $\hat{\sigma}_{11} = 0,47$, então a dispersão relativa dos consumos de vitaminas B12 e D são próximos. Em todos os ajustes considerados as estimativas do parâmetro de graus de liberdade indicam ajustes de distribuições de caudas pesadas, isto é devido à presença de observações discrepantes como mostram os bagplots da Figura 3.2.

Na Figura 3.3 mostramos para cada dupla de variáveis nutricionais curvas de níveis das distribuições ajustadas sobrepostas nos diagramas de dispersão e os respectivos gráficos das f.d.p. Note que as curvas de nível mostram um ajuste satisfatório em todas as duplas de variáveis nutricionais.

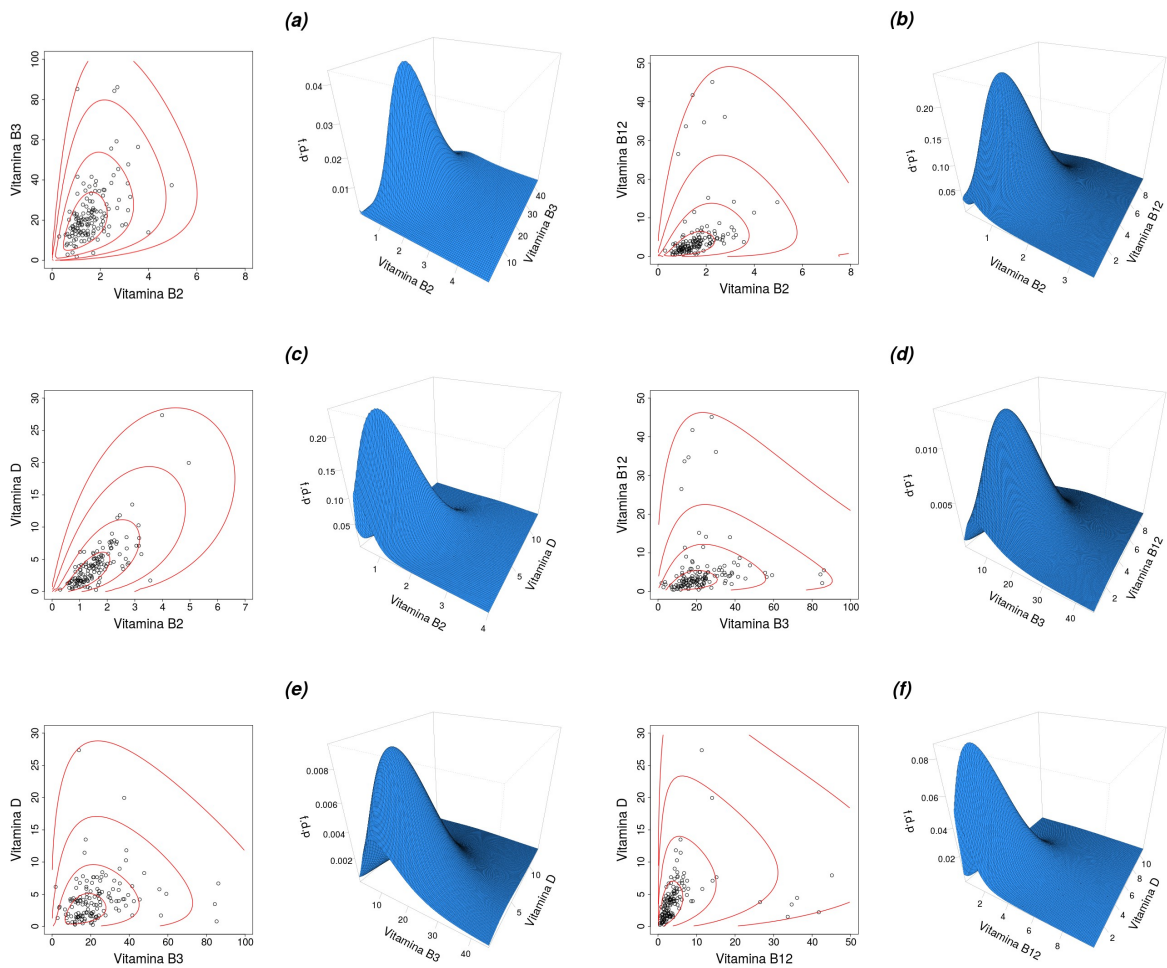


Figura 3.3: Curvas de níveis e gráficos das f.d.p das distribuições ajustadas a dados nutricionais.

Em seguida, consideramos medidas antropométricas de indivíduos na faixa etária entre 20 e 30 anos^{||}. Consideramos as medições dos perímetros da cintura (em cm), peito (em cm), quadril (em cm) e o índice de massa corporal (i.m.c, em kg/m^2) em

^{||} Este conjunto de dados está disponível na página web <http://ww2.amstat.org/publications/jse/datasets/body.dat.txt>, e é analisado em Heinz et al. [46]

247 homens. As estatísticas descritivas e as correlações amostrais dessas variáveis são apresentadas nas Tabelas 3.10 e 3.11, respectivamente. A matriz de bagplots apresentada na Figura 3.4 indica baixa presença de observações discrepantes nas duplas de variáveis antropométricas. Além disso, todas as duplas de variáveis são positivamente correlacionadas e apresentam assimetria multivariada.

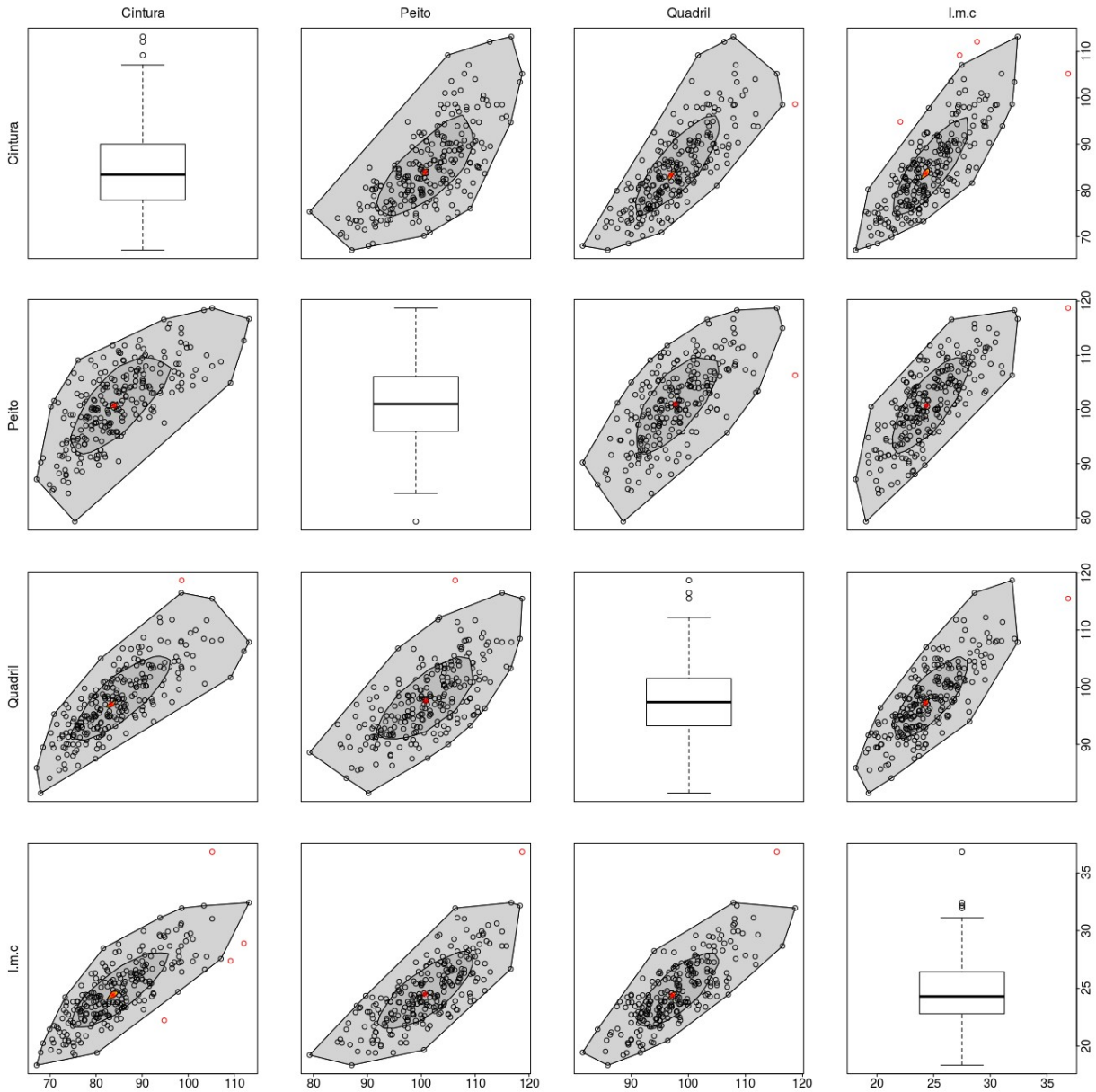


Figura 3.4: Matriz de bagplots das variáveis antropométricas.

Tabela 3.10: Estatísticas descritivas das variáveis antropométricas.

Variável	Min.	Q ₁	Mediana	Q ₃	Max.	CV	Média	SD
cintura	67,10	77,90	83,40	90,00	113,20	0,11	84,53	8,78
peito	79,30	95,95	101,00	106,00	118,70	0,07	101,00	7,21
quadril	81,50	93,25	97,40	101,60	118,70	0,06	97,76	6,23
i.m.c	18,35	22,81	24,30	26,43	36,82	0,11	24,71	2,84

Tabela 3.11: Correlações amostrais entre as variáveis antropométricas.

Variable	cintura	peito	quadril	i.m.c
cintura	1,00	0,71	0,80	0,80
peito	0,71	1,00	0,68	0,77
quadril	0,80	0,68	1,00	0,79
i.m.c	0,80	0,77	0,79	1,00

Conduzimos uma análise similar como fizemos com as variáveis nutricionais. A Tabela 3.12 apresenta os valores do AIC dos ajustes das distribuições Box-Cox normal e Box-Cox t marginalmente independentes, e das distribuições skew normal, skew- t , log-normal, log- t , Box-Cox normal e Box-Cox t bivariadas. A distribuição Box-Cox normal bivariada apresentou melhor ajuste que as outras distribuições em todos os casos, exceto para a dupla de variáveis quadril-i.m.c, em que o melhor ajuste foi dado através da distribuição log-normal bivariada. As distribuições Box-Cox normal e log-normal bivariadas modelam a associação existente entre as variáveis, caso que não é considerado ao ajustar distribuições Box-Cox normal marginalmente independentes. Também, a baixa presença de observações discrepantes (Figura 3.4) faz com que as distribuições Box-Cox normal e log-normal apresentem melhor ajuste que as distribuições Box-Cox t e log- t bivariadas. Na Tabela 3.13 apresentamos os valores das estimativas (e erros padrão) dos parâmetros correspondentes aos modelos que apresentam o melhor ajuste na Tabela 3.12.

As medianas e os coeficientes de variação baseados em percentis mostrados na Tabela 3.10 estão relacionados com as interpretações dos parâmetros estimados. A distribuição ajustada nas medições da dupla de variáveis cintura-peito é Box-Cox normal bivariada, sendo que $\hat{\lambda}_1\sqrt{\hat{\sigma}_{11}} = -0,146$ e $\hat{\lambda}_2\sqrt{\hat{\sigma}_{22}} = 0,113$ são próximos de zero. Portanto, $\hat{\mu}_1 = 83,474$ e $\hat{\mu}_2 = 101,143$ são aproximadamente iguais às medianas das medições de cintura e peito, respectivamente. Por outro lado, os coeficientes de variação baseados em percentis das medições de cintura e peito são aproximadamente iguais a $1,5 \sinh(\sqrt{\hat{\sigma}_{11}}s_{3/4})$ e $1,5 \sinh(\sqrt{\hat{\sigma}_{22}}s_{3/4})$, respectivamente; onde $s_{3/4}$ o terceiro quartil da distribuição normal padrão. Devido a que $\hat{\sigma}_{11} = 0,01$ e $\hat{\sigma}_{22} = 0,005$, então

Tabela 3.12: Valores do AIC das distribuições bivariadas ajustadas a duplas de variáveis antropométricas.

Variáveis	Ind. BCN ₁	Ind. BCt ₁	SN ₂	St ₂	LN ₂	Lt ₂	BCN ₂	BCt ₂
(cintura, peito)	3440,50	3444,50	3261,95	3263,95	3271,82	3273,82	3258,11	3260,11
(cintura, quadril)	3361,87	3365,87	3108,30	3108,62	3109,06	3111,06	3097,68	3099,68
(cintura, i.m.c)	2969,73	2973,73	2716,92	2715,93	2713,93	2715,93	2706,70	2708,70
(peito, quadril)	3283,10	3287,10	3128,21	3130,22	3132,98	3134,98	3126,16	3128,16
(peito, i.m.c)	2890,96	2894,96	2682,87	2673,51	2674,77	2676,78	2665,73	2667,74
(quadril, i.m.c)	2812,33	2816,33	2575,62	2577,22	2569,77	2571,77	2572,36	2574,36

Tabela 3.13: Estimativas (e erros padrão) dos parâmetros dos modelos ajustados a duplas de variáveis antropométricas.

Variáveis	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\sigma}_{11}$	$\hat{\sigma}_{12}$	$\hat{\sigma}_{22}$
(cintura, peito)	83,474 (0,564)	101,143 (0,487)	-1,460 (0,487)	1,596 (0,670)	0,010 (0,001)	0,005 (0,001)	0,005 (4,6 × 10 ⁻⁴)
(cintura, quadril)	83,397 (0,557)	97,293 (0,407)	-1,649 (0,453)	-1,430 (0,653)	0,010 (0,001)	0,005 (0,001)	0,004 (3,5 × 10 ⁻⁴)
(cintura, i.m.c)	83,455 (0,560)	24,463 (0,184)	-1,507 (0,465)	-0,572 (0,367)	0,010 (0,001)	0,009 (0,001)	0,013 (0,001)
(peito, quadril)	101,165 (0,488)	97,313 (0,411)	1,687 (0,687)	-1,325 (0,724)	0,005 (4,6 × 10 ⁻⁴)	0,003 (3,4 × 10 ⁻⁴)	0,004 (3,5 × 10 ⁻⁴)
(peito, i.m.c)	101,255 (0,485)	24,488 (0,185)	2,032 (0,664)	-0,410 (0,382)	0,005 (4,6 × 10 ⁻⁴)	0,006 (0,001)	0,013 (0,001)
(quadril, i.m.c)	97,569 (0,391)	24,553 (0,177)	-	-	0,004 (3,6 × 10 ⁻⁴)	0,006 (0,001)	0,013 (0,001)

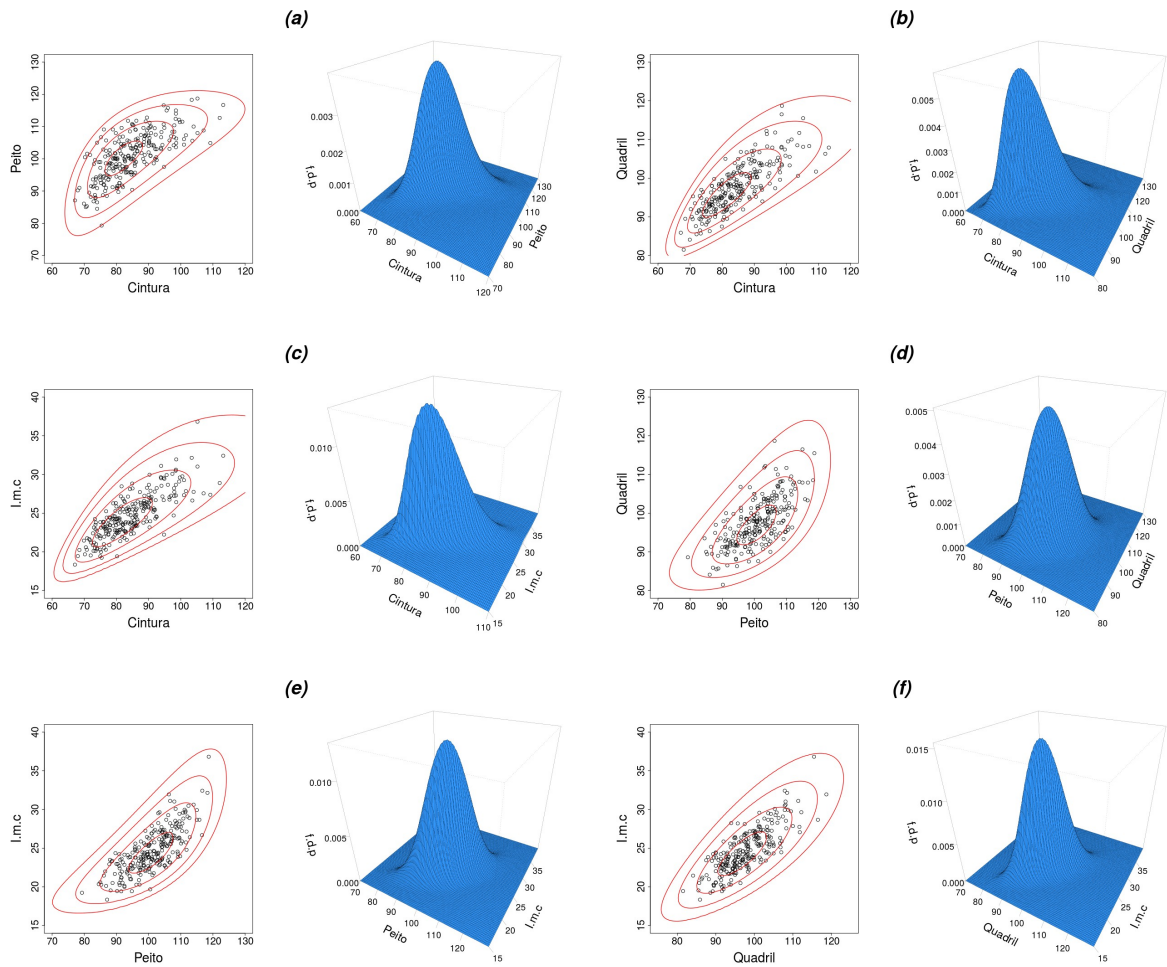


Figura 3.5: Curvas de níveis e gráficos das f.d.p das distribuições ajustadas a dados antropométricos.

as dispersões relativas das medidas de peito e cintura são próximas. A distribuição ajustada nas medições da dupla de variáveis quadril-i.m.c é log-normal bivariada, e assim $\hat{\mu}_1 = 97,569$ e $\hat{\mu}_2 = 24,553$ são aproximadamente iguais às medições medianas de quadril e i.m.c, respectivamente. Também, os coeficientes de variação baseados em percentis dessas medições são aproximadamente iguais a $1,5 \sinh(\sqrt{\hat{\sigma}_{11}}s_{3/4})$ e $1,5 \sinh(\sqrt{\hat{\sigma}_{22}}s_{3/4})$, respectivamente; em que $s_{3/4}$ o terceiro quartil da distribuição normal padrão. Como $\hat{\sigma}_{11} = 0,004$ e $\hat{\sigma}_{22} = 0,013$, então as dispersões relativas das medidas de i.m.c e quadril são próximas. Na Figura 3.5 apresentamos para cada dupla de variáveis antropométricas curvas de níveis das distribuições ajustadas sobrepostas nos diagramas de dispersão e os respectivos gráficos das f.d.p. Evidentemente, as curvas de nível mostram um ajuste satisfatório em todas as duplas de variáveis antropométricas.

3.8 Conclusões

Neste capítulo definimos uma nova classe de distribuições chamada de classe de distribuições Box-Cox elípticas. Para isto foi necessário definir uma nova família de transformações que denominamos transformações de Box-Cox estendida. Mostramos que a classe de distribuições Box-Cox elípticas é caracterizada por meio da classe de distribuições elípticas truncadas estudada no Capítulo 2. Apresentamos propriedades distribucionais desta nova classe através de transformações, assim como propriedades envolvendo distribuições marginais e condicionais, e independência. Também estabelecemos relações para calcular os momentos mistos, a matriz de covariâncias e os quantis marginalmente. Mostramos que alguns parâmetros que determinam as distribuições Box-Cox elípticas podem ser interpretados como características das distribuições marginais e outros parâmetros induzem associação entre essas distribuições. Além disso, mostramos que alguns parâmetros estão relacionados a percentis das distribuições marginais. Verificamos por meio de estudos de simulação que o método de máxima verossimilhança é adequado para estimar os parâmetros das distribuições Box-Cox normal e Box-Cox t multivariadas. Finalmente, apresentamos aplicações a dados nutricionais e antropométricos, mostrando que as distribuições Box-Cox elípticas são úteis para modelar dados positivos multivariados, marginalmente assimétricos e com presença de observações discrepantes.

Capítulo 4

Modelos de regressão lineares Box-Cox elípticos

4.1 Introdução

O modelo de regressão linear normal multivariado tem sido de grande importância na modelagem estatística. Este modelo é dado por

$$\begin{cases} \mathbf{Y}_i \sim N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}), \\ \boldsymbol{\mu}_i = \mathbf{B}'\mathbf{x}_i, \end{cases} \quad (4.1)$$

para $i = 1, \dots, n$, em que $\mathbf{Y}'_i = (Y_{i1}, \dots, Y_{ip})$ é o vetor de respostas do i -ésimo indivíduo para p variáveis Y_1, \dots, Y_p , sendo \mathbf{Y}_i e \mathbf{Y}_j independentes, $i \neq j$; $\mathbf{x}'_i = (x_{i1}, \dots, x_{ir})$ é o vetor com os valores do i -ésimo indivíduo para r variáveis explicativas X_1, \dots, X_r ; $\mathbf{B} = (\beta_{lk})_{r \times p}$ é a matriz de parâmetros e $\boldsymbol{\Sigma}(p \times p) > 0$.

Uma das principais desvantagens do modelo de regressão linear normal multivariado descrito acima é sua sensibilidade à presença de observações discrepantes. Considerar distribuições elípticas com caudas mais pesadas em relação à distribuição normal multivariada para os vetores de respostas \mathbf{Y}_i (modelo de regressão linear elíptico) é uma possível solução para este problema, como por exemplo a distribuição t multivariada (Lange et al. [3]). No entanto, os modelos de regressão lineares elípticos não levam em consideração a assimetria multivariada dos dados, fazendo com que esses modelos sejam inadequados para modelar dados com essa característica. A abordagem habitualmente utilizada para atender o problema de assimetria multivariada é considerar transformações dos dados, porém a interpretação dos parâmetros em termos do vetor de variáveis de interesse é perdida. Os modelos de regressão lineares skew-elípticos permitem modelar dados multivariados assimétricos com presença de observações discrepantes (Azzalini [40]). Nesses modelos os vetores de respostas

Y_i se supõem com distribuição skew-elíptica. Azzalini [40] estuda o método de estimação de máxima verossimilhança nos modelos de regressão lineares skew-normal e skew- t multivariados, que são casos particulares dos modelos de regressão lineares skew-elípticos. Também propõe métodos de diagnóstico e apresenta aplicações a dados multivariados.

Os modelos de regressão lineares elípticos e skew-elípticos podem ser inadequados para analisar dados positivos multivariados, especialmente se apresentam assimetria e observações discrepantes. Isto pode ser devido a que os vetores de variáveis resposta se supõem com uma distribuição com suporte em \mathbb{R}^p . Uma alternativa é usar esses modelos com os dados transformados por meio de uma transformação com contradomínio em \mathbb{R}^p , porém a interpretação dos parâmetros em termos do vetor de variáveis de interesse é perdida. Considerar distribuições multivariadas com suporte em \mathbb{R}_+^p para os vetores de respostas Y_i é uma alternativa para a modelagem desse tipo de dados, mas esta abordagem tem sido pouco estudada e utilizada. Exemplos deste tipo de abordagem aparecem em Iwasaki e Tsubaki [47] e Bonat e Jørgensen [48].

Neste capítulo definimos os modelos de regressão lineares Box-Cox elípticos, que permitem analisar dados positivos multivariados, marginalmente assimétricos e com presença de observações discrepantes. Mostramos como os quantis das distribuições marginais do vetor de variáveis resposta são afetados pelas covariáveis, o que facilita a interpretação dos parâmetros. Propomos o método de máxima verossimilhança para estimar os parâmetros envolvidos nestes modelos. Também apresentamos um método de diagnóstico baseado em distâncias de Mahalanobis (Healy [49]) para os modelos de regressão lineares log-normal e log- t multivariados. Finalmente, apresentamos aplicações a dados nutricionais e antropométricos.

4.2 Modelos de regressão lineares Box-Cox elípticos

Neste seção definimos e estudamos alguns aspectos dos modelos de regressão lineares Box-Cox elípticos.

Sejam Y_1, \dots, Y_n vetores aleatórios em \mathbb{R}_+^p dados por $Y_i' = (Y_{i1}, \dots, Y_{ip})$, $i = 1, \dots, n$. O componente Y_{ik} , $i = 1, \dots, n$, $k = 1, \dots, p$, representa a resposta do i -ésimo indivíduo para a k -ésima variável. Assumimos que para cada indivíduo i , os componentes do vetor Y_i são correlacionados, porém os vetores Y_1, \dots, Y_n são independentes. Os modelos de regressão lineares Box-Cox elípticos são dados por

$$\begin{cases} Y_i \sim \text{BCE}\ell(\mu_i, \lambda, \Sigma; g), \\ h(\mu_{ik}) = \mathbf{x}_i' \beta_k, \end{cases} \quad (4.2)$$

para $i = 1, \dots, n$, $k = 1, \dots, p$, em que $\mu_i' = (\mu_{i1}, \dots, \mu_{ip}) \in \mathbb{R}_+^p$, $\lambda' = (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p$,

$\Sigma = (\sigma_{jk})_{p \times p} > 0$ e g é a f.g.d, que deve ser tal que $g(u) \geq 0$, para todo $u \geq 0$, e $\int_0^\infty t^{p-1} g(t^2) dt < \infty$. A função h é monótona e diferenciável, chamada de função de ligação, $\beta_k = (\beta_{k1}, \dots, \beta_{kr})'$ é a k -ésima coluna da matriz de parâmetros $\mathbf{B} = (\beta_1, \dots, \beta_p)$ de dimensão $r \times p$, $\mathbf{x}'_i = (x_{i1}, \dots, x_{ir})$ é o vetor com os valores do i -ésimo indivíduo para as r variáveis explicativas X_1, \dots, X_r . O caso $\lambda = \mathbf{0}$ em (4.2) corresponde aos modelos de regressão lineares log-elípticos, que são dados por

$$\begin{cases} Y_i \sim \text{LEl}(\mu_i, \Sigma; g), \\ h(\mu_{ik}) = \mathbf{x}'_i \beta_k, \end{cases} \quad (4.3)$$

para $i = 1, \dots, n$, $k = 1, \dots, p$. O caso $p = 1$ em (4.2) corresponde aos modelos de regressão lineares Box-Cox simétricos dados por

$$\begin{cases} Y_i \sim \text{BCS}(\mu_i, \lambda, \sigma; g), \\ h(\mu_i) = \mathbf{x}'_i \beta, \end{cases} \quad (4.4)$$

para $i = 1, \dots, n$, em que $\mu_i > 0$, $\beta \in \mathbb{R}^r$, $\lambda \in \mathbb{R}$, $\sigma > 0$ e g é uma função tal que $g(u) \geq 0$, para todo $u \geq 0$, e $\int_0^\infty g(t^2) dt < \infty$. Estes modelos envolvem a classe de distribuições Box-Cox simétricas* estudada por Ferrari e Fumes [4]. Neste caso, se $\lambda = 0$, obtemos os modelos de regressão lineares log-simétricos† como

$$\begin{cases} Y_i \sim \text{LS}(\mu_i, \sigma; g), \\ h(\mu_i) = \mathbf{x}'_i \beta, \end{cases} \quad (4.5)$$

para $i = 1, \dots, n$, que corresponde ao caso $p = 1$ de (4.3).

Os modelos de regressão lineares Box-Cox elípticos são determinados pela f.g.d g (Seção 2.2), fornecendo várias alternativas para modelagem estatística, incluindo distribuições de caudas pesadas e leves. Denominamos (4.2) e (4.3) de modelos de regressão lineares Box-Cox normal multivariado e log-normal multivariado, respectivamente, quando a f.g.d g é dada por (2.5). Também, chamamos (4.2) e (4.3) de modelos de regressão lineares Box-Cox t multivariado e log- t multivariado, respectivamente, quando a f.g.d g é dada por (2.8). Neste último caso os modelos apresentam como parâmetro extra o número de graus de liberdade $\tau > 0$. Os modelos de regressão lineares Box-Cox normal, log-normal, Box-Cox t e log- t são casos particulares dos modelos de regressão lineares Box-Cox simétricos, e correspondem ao caso $p = 1$ dos modelos de regressão lineares Box-Cox normal, log-normal, Box-Cox t e log- t multivariados, res-

*Existe na literatura trabalhos de modelagem estatística através de algumas famílias pertencentes à classe de distribuições Box-Cox simétricas. Esses trabalhos envolvem as distribuições de Box-Cox normal, Box-Cox exponencial potência e Box-Cox t , e aparecem em Cole e Green [6], e Rigby e Stasinopoulos [7, 8], respectivamente.

†Uma extensão dos modelos de regressão lineares log-simétricos aparecem em Vanegas e Paula [50].

pectivamente.

Pelo Teorema 3.5.1 temos que no modelo (4.2) o α -quantil $y_{ik,\alpha}$ de Y_{ik} , $\alpha \in (0, 1)$, $i = 1, \dots, n$, $k = 1, \dots, p$, é dado por

$$y_{ik,\alpha} = \begin{cases} \mu_{ik} (1 + \lambda_k \sqrt{\sigma_{kk}} s_{k,\alpha})^{1/\lambda_k}, & \text{se } \lambda_k \neq 0, \\ \mu_{ik} \exp(\sqrt{\sigma_{kk}} s_{k,\alpha}), & \text{se } \lambda_k = 0, \end{cases} \quad (4.6)$$

em que $s_{k,\alpha}$ é dado por (3.33). Note que $s_{k,\alpha}$ é o mesmo para cada $i = 1, \dots, n$, já que se seguimos a construção apresentada na Seção 3.5 para obter a f.d.p marginal de Y_{ik} como em (3.24), chegamos a que $S_{ik} = \sigma_{kk}^{-1/2} T_{(\lambda_k, \mu_{ik})}(Y_{ik})$ tem a mesma f.d.p para cada i , a qual é dada por (3.28). Isto é devido a que a função $g_{\mathbf{r}_k}$ dada em (3.25) é determinada por $\boldsymbol{\lambda}$, $\boldsymbol{\Sigma}$ e g , que não dependem de i . A relação (4.6) indica que todos os quantis das distribuições marginais do vetor de variáveis resposta são afetados pelas variáveis explicativas, o que permite estabelecer a interpretação dos parâmetros β_{kl} , $k = 1, \dots, p$; $l = 1, \dots, r$. No entanto, essa interpretação depende da função de ligação considerada em (4.2). Em particular, se h é a função de ligação logarítmica, isto é, $h(x) = \log(x)$, $x > 0$, então $\mu_{ik} = \exp(\mathbf{x}'_i \boldsymbol{\beta}_k)$, $i = 1, \dots, n$, $k = 1, \dots, p$. Portanto, se X_{il} é acrescido em uma unidade, mantendo as demais variáveis explicativas fixadas, obtemos que o α -quantil $y_{ik,\alpha}$ de Y_{ik} , $\alpha \in (0, 1)$, é multiplicado pelo fator $\exp(\beta_{kl})$. Um fato interessante ocorre quando $\boldsymbol{\lambda} = \mathbf{0}$ (isto é $\mathbf{Y} \sim \text{LE}(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$) ou $\lambda_j \sqrt{\sigma_{jj}} \rightarrow 0$, $j = 1, \dots, p$. Neste caso as medianas de Y_{ik} são dadas $y_{ik,1/2} = \exp(\mathbf{x}'_i \boldsymbol{\beta}_k)$, o que indica que as medianas são afetadas unicamente pelas variáveis explicativas através da função exponencial, valendo também a interpretação dada acima em termos dos parâmetros β_{kl} . A interpretação dos parâmetros $\boldsymbol{\lambda}$ e $\boldsymbol{\Sigma}$ é a mesma para cada \mathbf{Y}_i , $i = 1, \dots, n$, como descrevemos na Seção 3.5.

4.3 Estimação de máxima verossimilhança

Neste seção descrevemos o método de máxima verossimilhança aplicado na estimação dos parâmetros que determinam os modelos de regressão lineares Box-Cox elípticos.

Sejam $\mathbf{y}_1, \dots, \mathbf{y}_n$ os valores observados de uma amostra aleatória $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ de acordo ao modelo (4.2), com respectivos vetores $\mathbf{x}_1, \dots, \mathbf{x}_n$ de valores para as r variáveis explicativas. Seja $\boldsymbol{\eta}' = (\eta_1, \dots, \eta_q)$ o vetor de parâmetros extras induzidos pela f.g.d g . Os EMV de \mathbf{B} , $\boldsymbol{\lambda}$, $\boldsymbol{\Sigma}$ e $\boldsymbol{\eta}$, representados por $\widehat{\mathbf{B}}$, $\widehat{\boldsymbol{\lambda}}$, $\widehat{\boldsymbol{\Sigma}}$ e $\widehat{\boldsymbol{\eta}}$, respectivamente, serão tais que maximizem o logaritmo da função de verossimilhança

$$\ell = \sum_{i=1}^n \ell_i, \quad (4.7)$$

sendo $\ell_i, i = 1, \dots, n$, o logaritmo da funo de verossimilhana de \mathbf{y}_i , dado por

$$\ell_i = -\log \left\{ \int_{A(\boldsymbol{\lambda})} g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) \, d\mathbf{w} \right\} + \log \{g(\mathbf{w}'_i\boldsymbol{\Sigma}^{-1}\mathbf{w}_i)\} + \\ \sum_{k=1}^p (\lambda_k - 1) \log y_{ik} - \sum_{k=1}^p \lambda_k \log \{h^{-1}(\mathbf{x}'_i\boldsymbol{\beta}_k)\},$$

sendo $\mathbf{w}_i = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu}_i)}(\mathbf{y}_i)$, com $\boldsymbol{\mu}_i = (h^{-1}(\mathbf{x}'_i\boldsymbol{\beta}_1), \dots, h^{-1}(\mathbf{x}'_i\boldsymbol{\beta}_p))$. Devido a que no existe forma fechada para $\widehat{\mathbf{B}}, \widehat{\boldsymbol{\lambda}}, \widehat{\boldsymbol{\Sigma}}$ e $\widehat{\boldsymbol{\eta}}$, sugerimos obter estes estimadores por meio de mtodos otimizaco implementados em programas computacionais.

Sejam $\mathbf{B}^{(0)}, \boldsymbol{\lambda}^{(0)}, \boldsymbol{\Sigma}^{(0)}$ e $\boldsymbol{\eta}^{(0)}$ os valores iniciais de $\mathbf{B}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}$ e $\boldsymbol{\eta}$, respectivamente. Para os valores de $\boldsymbol{\beta}_k^{(0)}$ (k -sima coluna de $\mathbf{B}^{(0)}$), $\lambda_k^{(0)}$ e $\sigma_{kk}^{(0)}, k = 1, \dots, p$, sugerimos as estimativas geradas ao ajustar o modelo de regresso Box-Cox simtrico (4.4) na k -sima linha de $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_n)$, isto , os valores dos parmetros estimados ao ajustar o modelo $Y_{ik} \sim \text{BCS}(\mu_{ik}, \lambda_k, \sqrt{\sigma_{kk}}; g)$, em que $h(\mu_{ik}) = \mathbf{x}'_i\boldsymbol{\beta}_k, i = 1, \dots, n$ [‡]. Para os valores iniciais dos parmetros σ_{jk} sugerimos $\sigma_{jk}^{(0)} = 0, j \neq k$. A proposta para os valores iniciais dos parmetros extras, $\eta_j^{(0)}, j = 1, \dots, q$, depender da famlia de distribuo considerada no modelo. No caso do modelo de regresso linear Box-Cox t multivariado, sugerimos como valor inicial do nmero de graus de liberdade, $\tau^{(0)}$, o nmero de graus de liberdade estimado no ajuste do modelo de regresso t multivariado[§] com vetor de variveis resposta $\mathbf{W}_i = T_{(\boldsymbol{\lambda}^{(0)}, \mathbf{B}^{(0)'}\mathbf{x}_i)}(\mathbf{Y}_i), i = 1, \dots, n$, em que $T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}$  a transformao de Box-Cox estendida definida na Seo 3.2.

Desenvolvemos as funoes `mlregmvtbcnorm` e `mlregmvtbct` em R que permitem obter os EMV dos parmetros nos modelos de regresso lineares Box-Cox normal e Box-Cox t multivariados considerando h como a funo de ligao logarmica, respectivamente (Apndice B). Os mtodos iterativos disponveis em nossos programas que permitem maximizar (4.7) para estes modelos so Nelder-Mead, BFGS, CG, SANN e L-BFGS-B (ver [38]), implementados na rotina `optim` do pacote `stats` do R. Alguns destes mtodos requerem reparametrizao para remover as restrioes induzidas pelas condioes $\boldsymbol{\Sigma}(p \times p) > 0$ e $\tau > 0$, conforme o caso. Essas reparametrizaoes so dadas por

$$\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}', \\ \tau = \exp(\nu), \quad \nu \in \mathbb{R},$$

sendo $\mathbf{L}(p \times p)$ uma matriz triangular inferior. No  necessrio considerar reparame-

[‡]O pacote `gamlss` (Rigby e Stasinopoulos [39]) do R permite ajustar os modelos de regresso lineares Box-Cox Cole Green (Box-Cox normal), Box-Cox t e Box-Cox exponencial potncia atravs do mtodo GAMLSS.

[§]Este modelo est implementado no pacote `sn` (Azzalini [40]) do R atravs do modelo de regresso linear skew- t multivariado.

trizações sobre os parâmetros \mathbf{B} e $\boldsymbol{\lambda} \in \mathbb{R}^p$.

4.4 Diagnóstico para os modelos log-normal e log- t multivariados

Nesta seção apresentamos um método de diagnóstico para avaliar ajustes dos modelos de regressão log-normal e log- t multivariados.

Uma variável aleatória unidimensional $X > 0$ tem distribuição qui-quadrado com $\nu > 0$ graus de liberdade, denotada por $X \sim \chi_\nu^2$, se sua f.d.p é dada por

$$f_X(x) = \frac{2^{-\nu/2}}{\Gamma(\nu/2)} x^{\nu/2-1} \exp\left(-\frac{x}{2}\right), \quad x > 0.$$

Uma variável aleatória unidimensional $Y > 0$ tem distribuição F com graus de liberdade $\nu_1 > 0$ e $\nu_2 > 0$, denotada por $Y \sim F_{\nu_1, \nu_2}$, se sua f.d.p é

$$f_Y(y) = \frac{\Gamma[(\nu_1 + \nu_2)/2]}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} y^{\nu_1/2-1} \left(1 + \frac{\nu_1}{\nu_2}y\right)^{-(\nu_1+\nu_2)/2}, \quad y > 0.$$

O método de diagnóstico que propomos para avaliar ajustes dos modelos de regressão log-normal e log- t multivariados estão baseados no Teorema 4.4.1, em que estabelecemos as distribuições de determinadas formas quadráticas de vetores aleatórios envolvidos nas distribuições log-normal e log- t multivariadas.

Teorema 4.4.1. *Sejam $\boldsymbol{\mu} \in \mathbb{R}_+^p$, $\boldsymbol{\Sigma}(p \times p) > 0$ e $\tau > 0$.*

1. *Seja $\mathbf{Y} \sim LN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ e $\mathbf{W} = T_{(0, \boldsymbol{\mu})}(\mathbf{Y})$, então $\mathbf{W}'\boldsymbol{\Sigma}^{-1}\mathbf{W} \sim \chi_p^2$.*
2. *Seja $\mathbf{Y} \sim Lt_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \tau)$ e $\mathbf{W} = T_{(0, \boldsymbol{\mu})}(\mathbf{Y})$, então $p^{-1}\mathbf{W}'\boldsymbol{\Sigma}^{-1}\mathbf{W} \sim F_{p, \tau}$.*

Demonstração. Ver Apêndice A.24. ■

Depois de ajustar o modelo de regressão linear log-normal ou log- t multivariado calculando $\widehat{\mathbf{B}}$, $\widehat{\boldsymbol{\lambda}}$, $\widehat{\boldsymbol{\Sigma}}$ e $\widehat{\tau}$ através do método de máxima verossimilhança, conforme o caso, avaliamos se o ajuste é adequado fazendo a comparação das distâncias de Mahalanobis empíricas (Healy [49]) com a respectiva distribuição sugerida pelo Teorema 4.4.1. As distâncias de Mahalanobis empíricas são dadas por $D_i = \widehat{\mathbf{W}}_i' \widehat{\boldsymbol{\Sigma}}^{-1} \widehat{\mathbf{W}}_i$, em que $\widehat{\mathbf{W}}_i = T_{(0, \widehat{\boldsymbol{\mu}}_i)}(\mathbf{Y}_i)$, sendo $\widehat{\boldsymbol{\mu}}_i = (h^{-1}(\mathbf{x}'_i \widehat{\boldsymbol{\beta}}_1), \dots, h^{-1}(\mathbf{x}'_i \widehat{\boldsymbol{\beta}}_p))$, $i = 1, \dots, n$. Os *QQ-plots* fornecem uma ferramenta gráfica útil para avaliar se o ajuste dos modelos log-normal ou log- t multivariado é adequado, através da comparação entre as distâncias de Mahalanobis empíricas ordenadas (quantis observados) e os α_i -quantis (quantis teóricos), em que $\alpha_i = i/(n+1)$, $i = 1, \dots, n$, das distribuições χ_p^2 ou $pF_{p, \widehat{\tau}}$, conforme o caso. No caso $p = 1$ as distâncias de Mahalanobis empíricas são dadas por $D_i = \sigma^{-2} \{\log[Y_i/h(\mathbf{x}'_i \widehat{\boldsymbol{\beta}})]\}^2$.

Neste caso, o *QQ-plot* compara os quantis observados com os quantis teóricos das distribuições χ_1^2 ou $F_{1,\hat{\tau}}$, conforme o caso.

4.5 Aplicações a dados reais

Nesta seção apresentamos aplicações a dados reais através de ajustes dos modelos de regressão lineares log-normal e log-*t* multivariados.

Primeiramente, consideramos o conjunto de dados nutricionais descrito na Seção 3.7. Na Tabela 4.1 apresentamos as estatísticas descritivas do i.m.c (em kg/m²) e dos consumos de vitaminas B2 (em mg), B3 (em mg) e D (em mcg) obtidos do primeiro R24h em 136 homens e 230 mulheres.

Tabela 4.1: Estatísticas descritivas das variáveis nutricionais segundo o sexo.

	Variável	Min.	Q ₁	Mediana	Q ₃	Max.	CV	Média	SD
Homens (<i>n</i> = 136)	i.m.c	17,56	23,81	26,51	30,48	47,58	0,19	27,24	4,91
	vit. B2	0,29	1,05	1,49	1,92	4,96	0,44	1,60	0,75
	vit. B3	1,84	14,14	19,99	26,20	86,13	0,45	22,89	14,27
	vit. D	0,26	1,73	3,80	5,39	27,35	0,72	4,16	3,51
Mulheres (<i>n</i> = 230)	i.m.c	15,21	24,60	27,67	31,57	44,04	0,19	28,29	5,40
	vit. B2	0,23	1,04	1,36	1,72	8,90	0,38	1,52	0,85
	vit. B3	3,41	11,76	16,79	22,74	83,80	0,49	18,99	11,46
	vit. D	0,05	1,99	3,28	5,12	25,36	0,72	3,94	3,02

O vetor de variáveis resposta é conformado pelos consumos de vitaminas B2, B3 e D. As variáveis explicativas são sexo (homem=1, mulher=0) e i.m.c. Ajustamos os modelos de regressão lineares skew-normal, skew-*t*, log-normal e log-*t* multivariados. Na Tabela 3.12 apresentamos os valores do AIC dos modelos ajustados. Estes valores mostram que o modelo de regressão linear log-*t* multivariado apresenta o melhor ajuste (isto é, tem o menor AIC). Na Figura 4.1 apresentamos *QQ-plots* das distâncias de Mahalanobis empíricas calculadas a partir de cada modelo ajustado[¶]. Esta figura sugere que, entre os quatro modelos considerados, o modelo de regressão linear log-*t* multivariado apresenta o melhor ajuste, sendo que o correspondente *QQ-plot* não apresenta fugas significativas dos pontos em torno da linha identidade.

[¶]No caso dos modelos de regressão lineares skew-normal e skew-*t* multivariados, Azzalini [40] propõe como métodos de diagnóstico comparar por meio de *QQ-plots* distâncias de Mahalanobis empíricas ordenadas e os α_i -quantis, em que $\alpha_i = i/(n+1)$, $i = 1, \dots, n$, das distribuições χ_p^2 ou $pF_{p,\hat{\tau}}$, respectivamente. Neste caso, $\hat{\tau}$ corresponde ao parâmetro de graus de liberdade estimado no modelo de regressão skew-*t* multivariado.

Tabela 4.2: Valores do AIC dos modelos de regressão lineares multivariados ajustados a dados nutricionais.

skew-normal	skew- t	log-normal	log- t
5303,86	4866,11	4847,26	4770,29

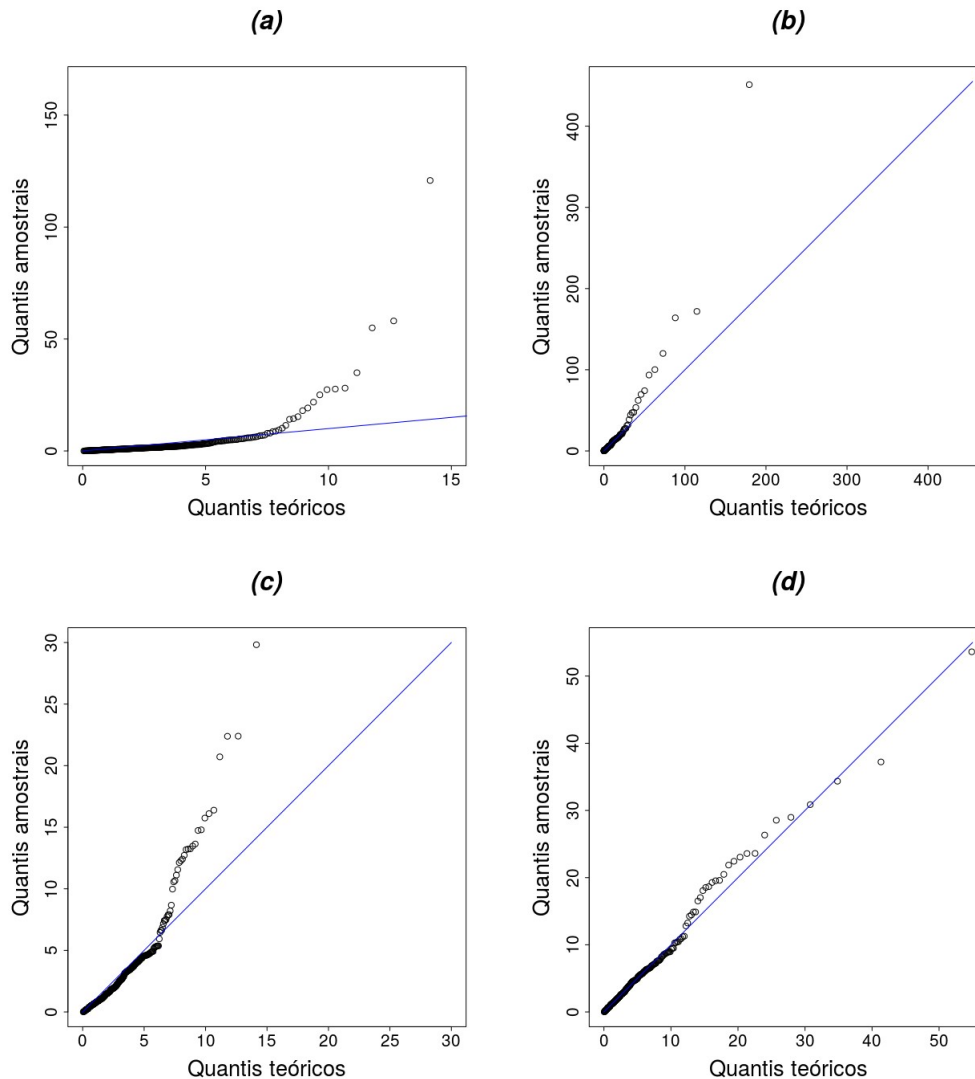


Figura 4.1: QQ-plot das distâncias de Mahalanobis empíricas do ajuste do modelo de regressão linear (a) skew-normal multivariado, (b) skew- t multivariado, (c) log-normal multivariado, (d) log- t multivariado.

Na Tabela 4.3 apresentamos as estimativas (e erros padrão) dos parâmetros envolvidos na estrutura de regressão correspondentes ao modelo que apresenta o melhor ajuste na Tabela 4.2 (log- t multivariado). Também apresentamos o valor- p do teste da razão de verossimilhanças (Sen et al. [45]) para testar $H_0 : \beta_{lk} = 0$, contra $H_1 : \beta_{lk} \neq 0$, $l = 1, 2, 3$, $k = 1, 2, 3$. Note que para o consumo de vitamina B3 o coeficiente da variável explicativa sexo é estatisticamente diferente de zero. As estimativas (e erros padrão) dos outros parâmetros envolvidos no modelo são $\hat{\sigma}_{11} = 0,139(0,013)$,

$\hat{\sigma}_{12} = 0,078 (0,012)$, $\hat{\sigma}_{13} = 0,172 (0,020)$, $\hat{\sigma}_{22} = 0,210 (0,021)$, $\hat{\sigma}_{23} = 0,055 (0,018)$, $\hat{\sigma}_{33} = 0,462 (0,046)$ e $\hat{\tau} = 5,545 (0,939)$. A estimativa do parâmetro de graus de liberdade indica ajuste de um modelo de cauda pesada; isto é devido à presença de observações discrepantes.

Tabela 4.3: Estimativas (e erros padrão) dos parâmetros do modelo de regressão linear log- t multivariado ajustado a dados nutricionais.

Resposta	Parâmetro	Estimativa	E. padrão	Valor- p
Vit. B2	Intercepto	0,382	0,120	0,000
	Sexo	0,053	0,046	0,250
	I.m.c	-0,002	0,004	0,607
Vit. B3	Intercepto	2,784	0,145	0,000
	Sexo	0,210	0,056	0,000
	I.m.c	$1,3 \times 10^{-4}$	0,005	0,982
Vit. D	Intercepto	1,422	0,215	0,000
	Sexo	0,007	0,083	0,932
	I.m.c	-0,009	0,007	0,243

Em seguida, analisamos o conjunto de dados antropométricos descrito na Seção 3.7. As estatísticas descritivas do i.m.c (em kg/m^2) e dos perímetros da cintura (em cm) e peito (em cm) de 247 homens e 260 mulheres são apresentadas na Tabela 4.4.

Tabela 4.4: Estatísticas descritivas das variáveis antropométricas segundo o sexo.

	Variável	Min.	Q_1	Mediana	Q_3	Max.	CV	Média	SD
Homens ($n = 247$)	I.m.c	18,35	22,81	24,30	26,43	36,82	0,11	24,71	2,84
	Cintura	67,10	77,90	83,40	90,00	113,20	0,11	84,53	8,78
	Peito	79,30	95,95	101,00	106,00	118,70	0,07	101,00	7,21
Mulheres ($n = 260$)	I.m.c	16,87	20,02	21,78	23,63	38,19	0,12	22,28	3,22
	Cintura	57,90	64,75	68,30	72,75	101,50	0,09	69,80	7,59
	Peito	72,60	81,98	85,50	89,50	109,00	0,07	86,06	6,17

Como vetor de variáveis resposta consideramos os perímetros da cintura e peito. As variáveis explicativas são sexo (homem=1, mulher=0) e i.m.c. Realizamos uma análise similar como fizemos com as variáveis nutricionais, e assim ajustamos os modelos de regressão linear skew-normal, skew- t , log-normal e log- t bivariados. Os valores do AIC apresentados na Tabela 4.5 indicam que o modelo que apresenta melhor ajuste é o modelo de regressão linear log- t bivariado. Note que dos QQ -plots da Figura 4.2 aquele que corresponde ao modelo de regressão linear skew- t bivariado não apresenta fugas significativas dos pontos em torno da linha identidade.

Tabela 4.5: Valores do AIC dos modelos de regressão lineares multivariados ajustados a dados antropométricos.

skew-normal	skew- t	log-normal	log- t
5783,76	5774,11	5718,92	5710,75

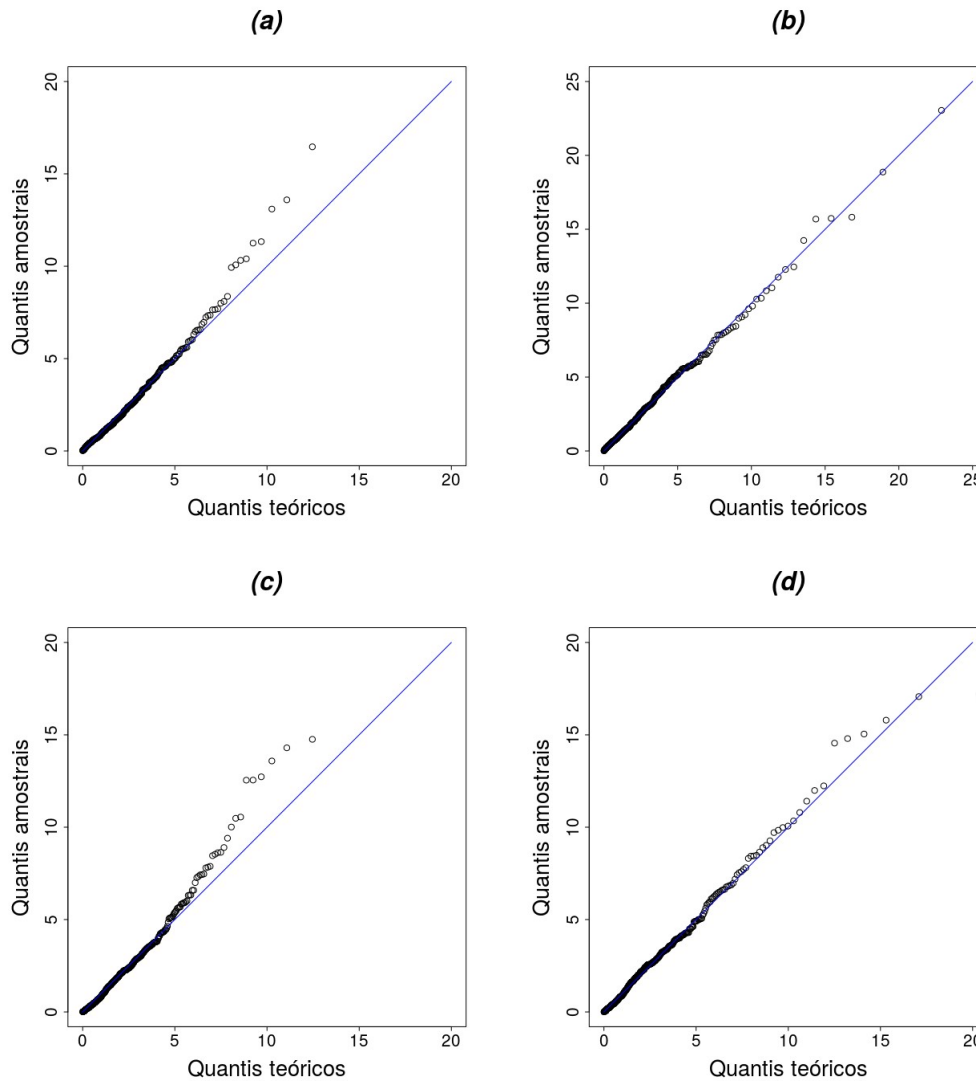


Figura 4.2: QQ-plot das distâncias de Mahalanobis empíricas do ajuste do modelo de regressão linear (a) skew-normal bivariado, (b) skew- t bivariado, (c) log-normal bivariado, (d) log- t bivariado.

Na Tabela 4.6 apresentamos as estimativas (e erros padrão) dos parâmetros envolvidos na estrutura de regressão do modelo de regressão log- t bivariado, sendo o modelo que apresenta o melhor ajuste na Tabela 4.5. Também apresentamos o valor- p do teste da razão de verossimilhanças para testar $H_0 : \beta_{lk} = 0$, contra $H_1 : \beta_{lk} \neq 0$, $l = 1, 2, 3$, $k = 1, 2$. Note que as covariáveis sexo e i.m.c apresentam coeficientes estatisticamente diferentes de zero. As estimativas dos outros parâmetros (e erros padrão) envolvidos no modelo são $\hat{\sigma}_{11} = 0,003 (2,3 \times 10^{-4})$, $\hat{\sigma}_{12} = 0,001 (1,1 \times 10^{-4})$, $\hat{\sigma}_{22} = 0,002 (1,4 \times 10^{-4})$ e

$\hat{\tau} = 13,638 (5,057)$. A estimativa do parâmetro de graus de liberdade sugere ajuste de um modelo de cauda pesada em relação ao modelo log-normal bivariado.

Tabela 4.6: Estimativas (e erros padrão) dos parâmetros do modelo de regressão linear log- t bivariado ajustado a dados antropométricos.

Resposta	Parâmetro	Estimativa	E. padrão	Valor- p
	Intercepto	3,608	0,019	0,000
Cintura	Sexo	0,122	0,005	0,000
	I.m.c	0,028	0,001	0,000
	Intercepto	4,042	0,015	0,000
Peito	Sexo	0,115	0,004	0,000
	I.m.c	0,018	0,001	0,000

4.6 Conclusões

Neste capítulo definimos os modelos de regressão lineares Box-Cox elípticos e mostramos que são úteis na análise de dados positivos multivariados, marginalmente assimétricos e com presença de observações discrepantes. Mostramos como as covariáveis afetam os quantis das distribuições marginais do vetor de variáveis resposta, facilitando a interpretação dos parâmetros. Descrevemos o método de máxima verossimilhança para estimar os parâmetros envolvidos nestes modelos. Propusemos um método de diagnóstico baseado em distâncias de Mahalanobis para os modelos de regressão linear log-normal e log- t multivariados. Finalmente, apresentamos aplicações a dados nutricionais e antropométricos.

Capítulo 5

Considerações finais e propostas futuras

Neste trabalho propusemos a classe de distribuições Box-Cox elípticas como alternativa para a modelagem de dados positivos multivariados através da transformação de Box-Cox multivariada. Mostramos que esta classe tem como casos particulares a classe de distribuições log-elípticas e Box-Cox simétricas. Provamos várias propriedades estatísticas e mostramos que os parâmetros envolvidos são interpretáveis em termos de características do vetor de variáveis originais, uma vez que alguns parâmetros modelam características das distribuições marginais e outros modelam associação. Além disso, mostramos que alguns parâmetros estão relacionados a quantis das distribuições marginais. Essa interpretação permite modelar dados positivos multivariados, marginalmente assimétricos e com presença de observações discrepantes, o que exemplificamos por meio de aplicações a dados reais. Estudamos aspectos teóricos e computacionais do método de máxima verossimilhança, que foi a metodologia que propusemos para a estimação dos parâmetros e que verificamos por meio de estudos de simulação.

O estudo da classe de distribuições Box-Cox elípticas está fundamentado na classe de distribuições elípticas truncadas, que definimos e estudamos previamente. Mostramos que a classe de distribuições elípticas é um caso particular da classe de distribuições elípticas truncadas. Provamos várias propriedades estatísticas e propusemos um algoritmo para gerar amostras aleatórias, que é útil na geração de amostras aleatórias das distribuições Box-Cox elípticas.

Neste trabalho também definimos os modelos de regressão lineares Box-Cox elípticos e seus casos particulares: modelos de regressão lineares log-elípticos e Box-Cox simétricos. Estabelecemos uma interpretação dos parâmetros envolvidos no modelo, que está baseada na forma em que as covariáveis afetam os quantis das distribuições marginais do vetor de variáveis resposta. Descrevemos o método de máxima verossimilhança como metodologia de estimação dos parâmetros. Propusemos métodos de diagnóstico para avaliar ajustes dos modelos de regressão lineares log-normal e log- t multivariados. Mostramos a utilidade dos modelos de regressão Box-Cox elípticos por

meio de aplicações a dados reais.

Disponibilizamos a programação em R relacionada a alguns cálculos computacionais das famílias de distribuições Box-Cox normal e Box-Cox t multivariadas, e que foi necessária para o desenvolvimento deste trabalho. Entre esses programas computacionais estão as funções que permitem gerar amostras aleatórias e estimar os parâmetros destas duas famílias, incluindo modelos de regressão lineares.

Para estudos futuros pretendemos:

1. Usar a teoria de cópulas como uma outra abordagem para construir distribuições multivariadas que marginalmente tenham características similares a variáveis aleatórias com distribuições Box-Cox simétricas, de modo que a associação entre essas variáveis seja levada em consideração.
2. Desenvolver métodos de diagnóstico para avaliar ajustes dos modelos de regressão Box-Cox elípticos.
3. Estender os modelos de regressão lineares Box-Cox elípticos colocando estrutura de regressão nos outros parâmetros.
4. Desenvolver o pacote `bcell` no *software* R , em que sejam considerados vários aspectos de cálculo computacional em relação a algumas famílias de distribuições pertencentes à classe de distribuições Box-Cox elípticas.

Apêndice A

Provas dos teoremas e corolários

A.1 Prova Teorema 2.3.1

1. Fazendo a transformação $\mathbf{u} = T(\mathbf{s}) = \mathbf{H}\mathbf{s}$ em (2.15), com Jacobiano $J(\mathbf{s} \rightarrow \mathbf{u}) = \pm 1$ e usando a propriedade $\mathbf{H}'\mathbf{H} = \mathbf{I}_p$, temos que

$$f_{\mathbf{S}}(\mathbf{s}) = \frac{g(\mathbf{u}'\mathbf{u})}{\int_{T(B)} g(\mathbf{u}'\mathbf{u}) d\mathbf{u}}, \quad \mathbf{u} = T(\mathbf{s}), \quad \mathbf{s} \in B.$$

Isto prova o resultado.

2. Considerando a mesma transformação do item 1 em (2.18), temos que

$$\begin{aligned} \phi_{\mathbf{S}}(\mathbf{t}) &= \frac{\int_{T(B)} \exp(i(\mathbf{H}\mathbf{t})'\mathbf{u}) g(\mathbf{u}'\mathbf{u}) d\mathbf{u}}{\int_{T(B)} g(\mathbf{u}'\mathbf{u}) d\mathbf{u}} \\ &= \phi_{\mathbf{S}^*}(T(\mathbf{t})), \quad \mathbf{t} \in \mathbb{R}^p. \end{aligned}$$

A.2 Prova Teorema 2.4.1

(\Rightarrow) A f.d.p. de \mathbf{S} é dada em (2.15). Transformando $\mathbf{X} = T(\mathbf{S}) = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\mathbf{S}$ em (2.15), com Jacobiano $J(\mathbf{s} \rightarrow \mathbf{x}) = \det(\boldsymbol{\Sigma})^{-1/2}$, e notando que $\mathbf{s}'\mathbf{s} = (\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})$, obtemos a f.d.p de \mathbf{X} como

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{g((\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}))}{\int_{T(B)} g((\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})) d\mathbf{x}}, \quad \mathbf{x} \in T(B). \quad (\text{A.1})$$

Portanto, $\mathbf{X} \sim \text{TE}\ell_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; T(B); g)$.

(\Leftarrow) A f.d.p de \mathbf{X} é dada por (A.1). Transformando $\mathbf{S} = T^{-1}(\mathbf{X}) = \boldsymbol{\Sigma}^{-1/2}(\mathbf{X} - \boldsymbol{\mu})$, com Jacobiano $J(\mathbf{x} \rightarrow \mathbf{s}) = \det(\boldsymbol{\Sigma})^{1/2}$, e notando que $(\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) = \mathbf{s}'\mathbf{s}$, obtemos o resultado.

A.3 Prova Corolário 2.4.1.1

Pelo Teorema 2.4.1 temos que

$$S = \frac{X - \mu}{\sigma} \sim \text{TS}_1\left(\left(\frac{a - \mu}{\sigma}, \frac{b - \mu}{\sigma}\right); g\right).$$

Portanto,

$$f_X(x) = \frac{1}{\sigma} f_S\left(\frac{x - \mu}{\sigma}\right), \quad x \in (a, b).$$

Aplicando a relação (2.17) obtemos a primeira parte do Corolário. A segunda parte do Corolário é imediata ao ver que

$$\frac{1}{\sigma} \int_a^x f_Z\left(\frac{t - \mu}{\sigma}\right) dt = F_Z\left(\frac{x - \mu}{\sigma}\right) - F_Z\left(\frac{a - \mu}{\sigma}\right).$$

A.4 Prova Corolário 2.4.1.2

Como $\mathbf{X} \sim \text{TEl}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; B; g)$, então pelo Teorema 2.4.1 segue que $\mathbf{S} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) \sim \text{TS}_p(T^{-1}(B); g)$. Assim,

$$\begin{aligned} \phi_{\mathbf{X}}(\mathbf{t}) &= \text{E}(\exp(i\mathbf{t}'\mathbf{X})) \\ &= \exp(i\mathbf{t}'\boldsymbol{\mu})\text{E}((\boldsymbol{\Sigma}^{1/2}\mathbf{t})'\mathbf{S}) \\ &= \exp(i\mathbf{t}'\boldsymbol{\mu})\phi_{\mathbf{S}}(\boldsymbol{\Sigma}^{1/2}\mathbf{t}). \end{aligned}$$

A.5 Prova Teorema 2.4.2

Fazendo a transformação $\mathbf{Y} = T(\mathbf{X}) = \mathbf{b} + \mathbf{D}\mathbf{X}$ em (2.19), com Jacobiano dado por $J(\mathbf{x} \rightarrow \mathbf{y}) = \det(\mathbf{D})^{-1}$, e notando que

$$(\mathbf{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) = (\mathbf{y} - (\mathbf{D}\boldsymbol{\mu} + \mathbf{b}))'(\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}')^{-1}(\mathbf{y} - (\mathbf{D}\boldsymbol{\mu} + \mathbf{b})),$$

obtemos o resultado.

A.6 Prova Teorema 2.4.3

Sejam $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$. Provaremos a desigualdade (2.25) para $f_{\mathbf{X}}$ dada em (2.20). Se $\mathbf{x} \notin B$ ou $\mathbf{y} \notin B$ temos que $[f_{\mathbf{X}}(\mathbf{x})]^\alpha [f_{\mathbf{X}}(\mathbf{y})]^{1-\alpha} = 0$, para todo $\alpha \in [0, 1]$, logo (2.25) é satisfeita, pois $f_{\mathbf{X}}(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \geq 0$. Se $\mathbf{x}, \mathbf{y} \in B$, então pela convexidade de B segue que $\alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in B$, para todo $\alpha \in [0, 1]$; logo $f_{\mathbf{X}}(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) > 0$ e, portanto,

$f_{\mathbf{W}}(\alpha \mathbf{x} + (1 - \alpha)\mathbf{y}) > 0$. Como $f_{\mathbf{W}}$ é log-côncava, então

$$\frac{f_{\mathbf{W}}(\alpha \mathbf{x} + (1 - \alpha)\mathbf{y})}{\mathbb{P}(\mathbf{W} \in B)} \geq \frac{[f_{\mathbf{W}}(\mathbf{x})]^\alpha [f_{\mathbf{W}}(\mathbf{y})]^{1-\alpha}}{\mathbb{P}(\mathbf{W} \in B)}, \quad \alpha \in [0, 1].$$

A equação (2.25) é satisfeita notando que, para $\alpha \in [0, 1]$, é satisfeita a relação

$$\mathbb{P}(\mathbf{W} \in B) = [\mathbb{P}(\mathbf{W} \in B)]^\alpha [\mathbb{P}(\mathbf{W} \in B)]^{1-\alpha}, \quad \alpha \in [0, 1].$$

A.7 Prova Teorema 2.4.4

Seguimos a prova de Tong [26], que apresenta o resultado no caso de um vetor aleatório com distribuição normal multivariada. Como $\mathbf{X} \sim \text{TE}_{\ell_p}(\boldsymbol{\mu}, \boldsymbol{\Sigma}; B; g)$ tem f.d.p log-côncava, sendo B um conjunto convexo, então usando iteradamente a desigualdade (2.26) temos que

$$\begin{aligned} \mathbb{P}\left[\mathbf{X} \in \sum_{i=1}^m \alpha_i A_i\right] &= \mathbb{P}\left[\mathbf{X} \in \alpha_1 A_1 + (1 - \alpha_1) \sum_{i=2}^m \frac{\alpha_i}{1 - \alpha_1} A_i\right] \\ &\geq \{\mathbb{P}[\mathbf{X} \in A_1]\}^{\alpha_1} \left\{ \mathbb{P}\left[\mathbf{X} \in \sum_{i=2}^m \frac{\alpha_i}{1 - \alpha_1} A_i\right] \right\}^{1-\alpha_1} \\ &\geq \{\mathbb{P}[\mathbf{X} \in A_1]\}^{\alpha_1} \{\mathbb{P}[\mathbf{X} \in A_2]\}^{\alpha_2} \left\{ \mathbb{P}\left[\mathbf{X} \in \sum_{i=3}^m \frac{\alpha_i}{1 - \alpha_1 - \alpha_2} A_i\right] \right\}^{1-\alpha_1-\alpha_2} \\ &\vdots \\ &\geq \prod_{i=1}^m \{\mathbb{P}(\mathbf{X} \in A_i)\}^{\alpha_i} \end{aligned}$$

A.8 Prova Teorema 2.4.5

Considerando $\boldsymbol{\Sigma}_{12} = \mathbf{0}$ em (2.29) temos que a f.d.p marginal de \mathbf{X}_1 é

$$f_{\mathbf{X}_1}(\mathbf{x}_1) = \frac{\int_{R_2} g((\mathbf{x}_1 - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) + (\mathbf{x}_2 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}_{22}^{-1} (\mathbf{x}_2 - \boldsymbol{\mu}_2)) d\mathbf{x}_2}{\int_{R_1} \int_{R_2} g((\mathbf{x}_1 - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) + (\mathbf{x}_2 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}_{22}^{-1} (\mathbf{x}_2 - \boldsymbol{\mu}_2)) d\mathbf{x}_2 d\mathbf{x}_1},$$

em que $\mathbf{x}_1 \in R_1$. Fazendo $\mathbf{w}_2 = T(\mathbf{x}_2) = \boldsymbol{\Sigma}_{22}^{-1/2} (\mathbf{x}_2 - \boldsymbol{\mu}_2)$, com Jacobiano dado por $J(\mathbf{x}_2 \rightarrow \mathbf{w}_2) = \det(\boldsymbol{\Sigma}_{22})^{1/2}$, segue que

$$\begin{aligned} f_{\mathbf{X}_1}(\mathbf{x}_1) &= \frac{\int_{T(R_2)} g((\mathbf{x}_1 - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) + \mathbf{w}_2' \mathbf{w}_2) d\mathbf{w}_2}{\int_{R_1} \int_{T(R_2)} g((\mathbf{x}_1 - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) + \mathbf{w}_2' \mathbf{w}_2) d\mathbf{w}_2 d\mathbf{x}_1} \\ &= \frac{g_1((\mathbf{x}_1 - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1))}{\int_{R_1} g_1((\mathbf{x}_1 - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1)) d\mathbf{x}_1}, \quad \mathbf{x}_1 \in R_1. \end{aligned}$$

Por outro lado, note que

$$g_1(u) = \int_{T(R_2)} g(u + \mathbf{w}'_2 \mathbf{w}_2) d\mathbf{w}_2 \leq \int_{\mathbb{R}^{p-r}} g(u + \mathbf{w}'_2 \mathbf{w}_2) d\mathbf{w}_2 = h_1(u), \quad u \geq 0.$$

A função h_1 é tal que $\int_0^\infty t^{r-1} h_1(t^2) dt < \infty$ (Fang et al. [9]). Isto completa a prova.

A.9 Prova Teorema 2.4.6

Usando (2.19) e (2.30) obtemos a f.d.p condicional de $\mathbf{X}_1 | \mathbf{X}_2$ como

$$f_{\mathbf{X}_1 | \mathbf{X}_2}(\mathbf{x}_1) = \frac{g((\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}))}{\int_{R_1} g((\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})) d\mathbf{x}_1}, \quad \mathbf{x}_1 \in R_1.$$

Pela identidade (Seber [41])

$$(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) = (\mathbf{x}_1 - \boldsymbol{\mu}_1(\mathbf{x}_2))' \boldsymbol{\Sigma}_{11.2}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1(\mathbf{x}_2)) + q(\mathbf{x}_2),$$

com $\boldsymbol{\mu}_1(\mathbf{x}_2)$, $\boldsymbol{\Sigma}_{11.2}$ e $q(\mathbf{x}_2)$ dados em (2.31), temos que

$$f_{\mathbf{X}_1 | \mathbf{X}_2}(\mathbf{x}_1) = \frac{g_{q(\mathbf{x}_2)}((\mathbf{x}_1 - \boldsymbol{\mu}_1(\mathbf{x}_2))' \boldsymbol{\Sigma}_{11.2}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1(\mathbf{x}_2)))}{\int_{R_1} g_{q(\mathbf{x}_2)}((\mathbf{x}_1 - \boldsymbol{\mu}_1(\mathbf{x}_2))' \boldsymbol{\Sigma}_{11.2}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1(\mathbf{x}_2))) d\mathbf{x}_1}, \quad \mathbf{x}_1 \in R_1,$$

onde $g_{q(\mathbf{x}_2)}$ é dada em (2.31). Isto prova o resultado.

A.10 Prova Teorema 2.4.7

(\Rightarrow) Sejam \mathbf{X}_1 e \mathbf{X}_2 independentes, então $f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{X}_1}(\mathbf{x}_1) f_{\mathbf{X}_2}(\mathbf{x}_2)$. Portanto,

$$\frac{g((\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}))}{\int_{R_1} \int_{R_2} g((\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})) d\mathbf{x}_1 d\mathbf{x}_2} = \frac{g((\mathbf{x}_1 - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1)) g((\mathbf{x}_2 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}_{22}^{-1} (\mathbf{x}_2 - \boldsymbol{\mu}_2))}{\int_{R_1} g((\mathbf{x}_1 - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1)) d\mathbf{x}_1 \int_{R_2} g((\mathbf{x}_2 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}_{22}^{-1} (\mathbf{x}_2 - \boldsymbol{\mu}_2)) d\mathbf{x}_2},$$

em que $\mathbf{x}_1 \in R_1$ e $\mathbf{x}_2 \in R_2$. Isto ocorre apenas se $\boldsymbol{\Sigma}_{12} = \mathbf{0}$ e a função g satisfaz a equação funcional $g(u + v) = g(u)g(v)$, com $u \geq 0$ e $v \geq 0$, que tem solução $g(u) = \exp(-ku)$, para algum $k \geq 0$ (Gupta et al. et al.[23, Seção 1.3]). A partir da condição (Fang et al. [9])

$$\int_0^\infty t^{p-1} g(t^2) dt = \frac{\Gamma(p/2)}{2\pi^{p/2}},$$

encontramos que $k = 1/2$. Portanto, $\mathbf{X} \sim \text{TN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; R; g)$.

(\Leftarrow) Se $\mathbf{X} \sim \text{TN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; R; g)$ e $\boldsymbol{\Sigma}_{12} = \mathbf{0}$, então pela identidade

$$(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) = (\mathbf{x}_1 - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) + (\mathbf{x}_2 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}_{22}^{-1} (\mathbf{x}_2 - \boldsymbol{\mu}_2)$$

e a propriedade $\exp(u + v) = \exp(u) \exp(v)$, $u \in \mathbb{R}$ e $v \in \mathbb{R}$, obtemos que $f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{X}_1}(\mathbf{x}_1) f_{\mathbf{X}_2}(\mathbf{x}_2)$, isto é, \mathbf{X}_1 e \mathbf{X}_2 são independentes.

A.11 Prova Teorema 3.2.1

Seja $f_{\mathbf{W}}$ a f.d.p de \mathbf{W} . A f.d.p de \mathbf{Y} é dada por

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{W}}(\mathbf{w}) \prod_{j=1}^p \frac{y_j^{\lambda_j - 1}}{\mu_j^{\lambda_j}}, \quad \mathbf{w} = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}(\mathbf{y}), \quad \mathbf{y} \in \mathbb{R}_+^p.$$

Portanto, a f.d.p marginal de Y_k , $k = 1, \dots, p$, é dada por

$$\begin{aligned} f_{Y_k}(y_k) &= \int_{\mathbb{R}_+^{p-1}} f_{\mathbf{W}}(\mathbf{w}) \prod_{j=1}^p \frac{y_j^{\lambda_j - 1}}{\mu_j^{\lambda_j}} d\mathbf{y}_{-k} \\ &= \frac{1}{\mu_k} \left\{ \int_{A(\boldsymbol{\lambda}_{-k})} f_{\mathbf{W}}(\mathbf{w}) d\mathbf{w}_{-k} \right\} \left(\frac{y_k}{\mu_k} \right)^{\lambda_k - 1} \end{aligned}$$

Como w_k , o k -ésimo componente de \mathbf{w} , depende de y_k/μ_k através de (3.5), então

$$f_{Y_k}(y_k) = \frac{1}{\mu_k} f\left(\frac{y_k}{\mu_k}\right),$$

em que

$$f(u) = \left\{ \int_{A(\boldsymbol{\lambda}_{-k})} f_{\mathbf{W}}(w_1, \dots, w_{k-1}, u, w_{k+1}, \dots, w_p) d\mathbf{w}_{-k} \right\} u^{\lambda_k - 1}.$$

Isto prova que μ_k é parâmetro de escala de Y_k , para $k = 1, \dots, p$.

A.12 Prova Teorema 3.2.2

1. Seja $\mathbf{Y} \in \mathbb{R}_+^p$. Então,

$$(T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})} \circ U)(\mathbf{Y}) = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}(U(\mathbf{Y})) = T_{(\boldsymbol{\lambda}, \boldsymbol{\mu})}(\mathbf{D}_a \mathbf{Y}) = \mathbf{Z}.$$

O k -ésimo componente Z é dado por

$$Z_k = \begin{cases} \frac{(Y_k/a_k^{-1}\mu_k)^{\lambda_k} - 1}{\lambda_k}, & \lambda_k \neq 0, \\ \log(Y_k/a_k^{-1}\mu_k), & \lambda_k = 0, \end{cases}$$

para $k = 1, \dots, p$. Portanto, $T_{(\lambda, \mu)} \circ U = T_{(\lambda, D_a^{-1}\mu)}$.

2. Seja $\mathbf{Y} \in \mathbb{R}_+^p$. Então,

$$(V \circ T_{(\lambda, \mu)})(\mathbf{Y}) = V(T_{(\lambda, \mu)}(\mathbf{Y})) = D(T_{(\lambda, \mu)}(\mathbf{Y}) - \mathbf{b}) = \mathbf{Z},$$

O k -ésimo componente de Z é dado por

$$\begin{aligned} Z_k &= \frac{1}{1 + \lambda_k b_k} \begin{cases} \frac{(Y_k/\mu_k)^{\lambda_k} - 1}{\lambda_k} - b_k, & \lambda_k \neq 0, \\ \log(Y_k/\mu_k) - b_k, & \lambda_k = 0, \end{cases} \\ &= \begin{cases} \frac{(Y_k/\delta_k)^{\lambda_k} - 1}{\lambda_k}, & \lambda_k \neq 0, \\ \log(Y_k/\delta_k), & \lambda_k = 0, \end{cases} \end{aligned}$$

em que

$$\delta_k = \begin{cases} \mu_k(1 + \lambda_k b_k)^{1/\lambda_k}, & \lambda_k \neq 0, \\ \mu_k \exp(b_k), & \lambda_k = 0, \end{cases}$$

para $k = 1, \dots, p$. Portanto, $V \circ T_{(\lambda, \mu)} = T_{(\lambda, \delta)}$.

A.13 Prova Teorema 3.3.1

Temos que $\mathbf{Y} \sim \text{BCE}\ell_p(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$ se $\mathbf{W} = T_{(\lambda, \mu)}(\mathbf{Y}) \sim \text{TE}\ell_p(\mathbf{0}, \boldsymbol{\Sigma}; A(\boldsymbol{\lambda}); g)$. A f.d.p de \mathbf{W} é dada por

$$f_{\mathbf{W}}(\mathbf{w}) = \frac{g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w})}{\int_{A(\boldsymbol{\lambda})} g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) d\mathbf{w}}, \quad \mathbf{w} \in A(\boldsymbol{\lambda}).$$

O Jacobiano da transformação $\mathbf{W} = T_{(\lambda, \mu)}(\mathbf{Y})$ é

$$J(\mathbf{w} \rightarrow \mathbf{y}) = \prod_{k=1}^p \frac{y_k^{\lambda_k - 1}}{\mu_k^{\lambda_k}}.$$

Portanto, a f.d.p de \mathbf{Y} é calculada como

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{g(\mathbf{w}'\Sigma^{-1}\mathbf{w}) \prod_{k=1}^p \frac{y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}}}{\int_{A(\lambda)} g(\mathbf{w}'\Sigma^{-1}\mathbf{w}) d\mathbf{w}}, \quad \mathbf{w} = T_{(\lambda,\mu)}(\mathbf{y}), \quad \mathbf{y} \in \mathbb{R}_+^p.$$

A.14 Prova Teorema 3.4.1

(\Rightarrow) Como $\mathbf{X} = T_{(\lambda,\mu)}(\mathbf{Y}) \sim \text{TEL}_p(\mathbf{b}, \Sigma; A(\lambda); g)$, então sua f.d.p é dada por

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{g((\mathbf{x} - \mathbf{b})'\Sigma^{-1}(\mathbf{x} - \mathbf{b}))}{\int_{A(\lambda)} g((\mathbf{x} - \mathbf{b})'\Sigma^{-1}(\mathbf{x} - \mathbf{b})) d\mathbf{x}}, \quad \mathbf{x} \in A(\lambda).$$

Seja $V : \mathbb{R}^p \rightarrow \mathbb{R}^p$ a transformação dada por $V(\mathbf{x}) = \mathbf{D}(\mathbf{x} - \mathbf{b})$, sendo $\mathbf{D} = \mathbf{D}_{1+\mathbf{D}\lambda\mathbf{b}}^{-1}$. Transformando $\mathbf{W} = V(\mathbf{X})$, temos que $\mathbf{W} \sim \text{TEL}_p(\mathbf{0}, \mathbf{D}\Sigma\mathbf{D}; A(\lambda); g)$, isto foi obtido em virtude do Teorema 2.4.2 e notando que $V(A(\lambda)) = A(\lambda)$. Portanto, a f.d.p de \mathbf{W} é

$$f_{\mathbf{W}}(\mathbf{w}) = \frac{g(\mathbf{w}'(\mathbf{D}\Sigma\mathbf{D})^{-1}\mathbf{w})}{\int_{A(\lambda)} g(\mathbf{w}'(\mathbf{D}\Sigma\mathbf{D})^{-1}\mathbf{w}) d\mathbf{w}}, \quad \mathbf{w} \in A(\lambda).$$

Pelo item 2 do Teorema 3.2.2 temos que $\mathbf{W} = V(T_{(\lambda,\mu)}(\mathbf{Y})) = T_{(\lambda,\delta)}(\mathbf{Y})$, em que $\delta = T_{(\lambda,\mu)}^{-1}(\mathbf{b})$. Assim, o Jacobiano desta transformação é dado por

$$J(\mathbf{w} \rightarrow \mathbf{y}) = \prod_{k=1}^p \frac{y_k^{\lambda_k-1}}{\delta_k^{\lambda_k}}.$$

Portanto, a f.d.p de \mathbf{Y} é

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{g(\mathbf{w}'(\mathbf{D}\Sigma\mathbf{D})^{-1}\mathbf{w}) \prod_{k=1}^p \frac{y_k^{\lambda_k-1}}{\delta_k^{\lambda_k}}}{\int_{A(\lambda)} g(\mathbf{w}'(\mathbf{D}\Sigma\mathbf{D})^{-1}\mathbf{w}) d\mathbf{w}}, \quad \mathbf{w} = T_{(\lambda,\delta)}(\mathbf{y}), \quad \mathbf{y} \in \mathbb{R}_+^p. \quad (\text{A.2})$$

Isto prova que $\mathbf{Y} \sim \text{BCEL}_p(\delta, \lambda, \mathbf{D}\Sigma\mathbf{D}; g)$.

(\Leftarrow) Como $\mathbf{Y} \sim \text{BCEL}_p(\delta, \lambda, \mathbf{D}\Sigma\mathbf{D}; g)$, então a f.d.p de \mathbf{Y} é dada por (A.2). Basta considerar a transformação $\mathbf{W} = T_{(\lambda,\mu)}(\mathbf{Y})$, com Jacobiano $J(\mathbf{y} \rightarrow \mathbf{w}) = \prod_{k=1}^p \mu_k (1 + \lambda_k w_k)^{1/\lambda_k-1}$, para obter o resultado.

A.15 Prova Teorema 3.4.2

1. Seja $T = D_a Y$, então $J(\mathbf{y} \rightarrow \mathbf{t}) = \prod_{k=1}^p a_k^{-1}$. A f.d.p de T é

$$f_T(\mathbf{t}) = \frac{g(\mathbf{w}'\Sigma^{-1}\mathbf{w}) \prod_{k=1}^p \frac{t_k^{\lambda_k-1}}{(a_k \mu_k)^{\lambda_k}}}{\int_{A(\lambda)} g(\mathbf{w}'\Sigma^{-1}\mathbf{w}) d\mathbf{w}}, \quad \mathbf{t} \in \mathbb{R}_+^p,$$

em que, pelo item 1 do Teorema 3.2.2, $\mathbf{w} = T_{(\lambda, \mu)}(D_a^{-1}\mathbf{t}) = T_{(\lambda, D_a \mu)}(\mathbf{t})$.

2. Fazendo a transformação $\mathbf{w} = D_r^{-1}\mathbf{v}$ na integral que aparece em (3.6), com Jacobiano $J(\mathbf{w} \rightarrow \mathbf{v}) = \prod_{k=1}^p r_k^{-1}$, temos que a f.d.p de Y pode ser expressa como

$$f_Y(\mathbf{y}) = \frac{g(\mathbf{v}'(D_r \Sigma D_r)^{-1}\mathbf{v}) \prod_{k=1}^p \frac{|r_k| y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}}}{\int_{A(D_r^{-1}\lambda)} g(\mathbf{v}'(D_r \Sigma D_r)^{-1}\mathbf{v}) d\mathbf{v}}, \quad \mathbf{y} \in \mathbb{R}_+^p,$$

em que $\mathbf{v} = D_r T_{(\lambda, \mu)}(\mathbf{y})$. Note que o k -ésimo componente de \mathbf{v} pode ser expresso como

$$v_k = \begin{cases} \frac{[(y_k/\mu_k)^{r_k}]^{\lambda_k/r_k} - 1}{\lambda_k/r_k}, & \text{se } \lambda_k \neq 0, \\ \log(y_k/\mu_k)^{r_k}, & \text{se } \lambda_k = 0, \end{cases} \quad (\text{A.3})$$

para $k = 1, \dots, p$. Consideramos a transformação $U_k = (Y_k/\mu_k)^{r_k}$, $k = 1, \dots, p$, com Jacobiano

$$J(\mathbf{y} \rightarrow \mathbf{u}) = \prod_{k=1}^p \frac{\mu_k u_k^{1/r_k-1}}{r_k}.$$

Portanto, a f.d.p de $U = (U_1, \dots, U_p)$ é

$$f_U(\mathbf{u}) = \frac{g(\mathbf{v}'(D_r \Sigma D_r)^{-1}\mathbf{v}) \prod_{k=1}^p u_k^{\lambda_k/r_k-1}}{\int_{A(D_r^{-1}\lambda)} g(\mathbf{v}'(D_r \Sigma D_r)^{-1}\mathbf{v}) d\mathbf{v}}, \quad \mathbf{v} = T_{(D_r^{-1}\lambda, 1)}(\mathbf{u}), \quad \mathbf{u} \in \mathbb{R}_+^p.$$

Isto prova o resultado.

3. Fazendo $\lambda = 1$ em (3.6) temos que a f.d.p de Y toma a forma

$$f_Y(\mathbf{y}) = \frac{g(\mathbf{w}'\Sigma^{-1}\mathbf{w}) \prod_{k=1}^p \frac{1}{\mu_k}}{\int_{A(1)} g(\mathbf{w}'\Sigma^{-1}\mathbf{w}) d\mathbf{w}}, \quad \mathbf{y} \in \mathbb{R}_+^p,$$

em que $\mathbf{w} = D_\mu^{-1}(\mathbf{y} - \mu)$. Assim

$$f_Y(\mathbf{y}) = \frac{g((\mathbf{y} - \mu)'(D_\mu \Sigma D_\mu)^{-1}(\mathbf{y} - \mu))}{\int_{\mathbb{R}_+^p} g((\mathbf{y} - \mu)'(D_\mu \Sigma D_\mu)^{-1}(\mathbf{y} - \mu)) d\mathbf{y}}, \quad \mathbf{y} \in \mathbb{R}_+^p,$$

onde foi feita a mudança de variáveis $\mathbf{w} = D_\mu^{-1}(\mathbf{y} - \mu)$ na integral do denomina-

dor, com Jacobiano $J(\mathbf{w} \rightarrow \mathbf{y}) = \prod_{k=1}^p \mu_k^{-1}$. Isto completa a prova.

A.16 Prova Teorema 3.4.3

A partir de (2.12) e (3.10) temos que

$$\begin{aligned} \int_0^\infty f_{\mathbf{Y}|v}(\mathbf{y}) f_V(v) dv &= \frac{\det(\boldsymbol{\Sigma})^{-1/2} (\tau/2)^{\tau/2}}{(2\pi)^{p/2} \Gamma(\tau/2)} \\ &\quad \times \int_0^\infty v^{-\frac{\tau+p}{2}-1} \exp\left[-\frac{1}{2v}(\tau + \mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w})\right] dv \prod_{k=1}^p \frac{1}{y_k} \\ &= \frac{\Gamma((\tau+p)/2) \det(\boldsymbol{\Sigma})^{-1/2}}{(\tau\pi)^{p/2} \Gamma(\tau/2)} \left(1 + \frac{1}{\tau} \mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}\right)^{-\frac{\tau+p}{2}} \prod_{k=1}^p \frac{1}{y_k} \\ &= f_{\mathbf{Y}}(\mathbf{y}), \end{aligned}$$

em que $\mathbf{w} = T_{(0,\mu)}(\mathbf{y})$.

A.17 Prova Teorema 3.4.4

Como $\boldsymbol{\Sigma}_{12} = \mathbf{0}$, então (3.15) assume a forma

$$f_{\mathbf{Y}_1}(\mathbf{y}_1) = \frac{\left\{ \int_{A(\lambda_2)} g(\mathbf{w}'_1 \boldsymbol{\Sigma}_{11}^{-1} \mathbf{w}_1 + \mathbf{w}'_2 \boldsymbol{\Sigma}_{22}^{-1} \mathbf{w}_2) d\mathbf{w}_2 \right\} \prod_{k=1}^r \frac{y_k^{\lambda_k - 1}}{\mu_k^{\lambda_k}}}{\int_{A(\lambda_1)} \left\{ \int_{A(\lambda_2)} g(\mathbf{w}'_1 \boldsymbol{\Sigma}_{11}^{-1} \mathbf{w}_1 + \mathbf{w}'_2 \boldsymbol{\Sigma}_{22}^{-1} \mathbf{w}_2) d\mathbf{w}_2 \right\} d\mathbf{w}_1}, \quad \mathbf{y}_1 \in \mathbb{R}_+^r,$$

em que $\mathbf{w}_1 = T_{(\lambda_1, \mu_1)}(\mathbf{y}_1)$. Fazendo $\mathbf{s}_2 = T(\mathbf{w}_2) = \boldsymbol{\Sigma}_{22}^{-1/2} \mathbf{w}_2$ nas integrais do numerador e denominador, com Jacobiano, $J(\mathbf{w}_2 \rightarrow \mathbf{s}_2) = \det(\boldsymbol{\Sigma}_{22})^{1/2}$, temos que

$$\begin{aligned} f_{\mathbf{Y}_1}(\mathbf{y}_1) &= \frac{\left\{ \int_{T(A(\lambda_2))} g(\mathbf{w}'_1 \boldsymbol{\Sigma}_{11}^{-1} \mathbf{w}_1 + \mathbf{s}'_2 \mathbf{s}_2) d\mathbf{s}_2 \right\} \prod_{k=1}^r \frac{y_k^{\lambda_k - 1}}{\mu_k^{\lambda_k}}}{\int_{A(\lambda_1)} \left\{ \int_{T(A(\lambda_2))} g(\mathbf{w}'_1 \boldsymbol{\Sigma}_{11}^{-1} \mathbf{w}_1 + \mathbf{s}'_2 \mathbf{s}_2) d\mathbf{s}_2 \right\} d\mathbf{w}_1} \\ &= \frac{g_1(\mathbf{w}'_1 \boldsymbol{\Sigma}_{11}^{-1} \mathbf{w}_1) \prod_{k=1}^r \frac{y_k^{\lambda_k - 1}}{\mu_k^{\lambda_k}}}{\int_{A(\lambda_1)} g_1(\mathbf{w}'_1 \boldsymbol{\Sigma}_{11}^{-1} \mathbf{w}_1) d\mathbf{w}_1}, \quad \mathbf{w}_1 = T_{(\lambda_1, \mu_1)}(\mathbf{y}_1), \quad \mathbf{y}_1 \in \mathbb{R}_+^r. \end{aligned}$$

Note que

$$g_1(u) = \int_{T(A(\lambda_2))} g(u + \mathbf{s}'_2 \mathbf{s}_2) d\mathbf{s}_2 \leq \int_{\mathbb{R}^{p-r}} g(u + \mathbf{s}'_2 \mathbf{s}_2) d\mathbf{s}_2 = h_1(u), \quad u \geq 0,$$

sendo h_1 uma função tal que $\int_0^\infty t^{r-1} h_1(t^2) dt < \infty$ (Fang et al. [9]). Isto completa a prova do resultado.

A.18 Prova Teorema 3.4.5

1. Como $Y \sim \text{LN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, então $\mathbf{X} = T_{(0, \boldsymbol{\mu})}(Y) \sim \text{N}_p(\mathbf{0}, \boldsymbol{\Sigma})$. Assim, $\mathbf{X}_1 = T_{(0, \boldsymbol{\mu}_1)}(Y_1) \sim \text{N}_r(\mathbf{0}, \boldsymbol{\Sigma}_{11})$. Portanto, $Y_1 \sim \text{LN}_r(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$.
2. Similar à prova do item 1.

A.19 Prova Teorema 3.4.6

A partir de (3.6) e (3.16) temos que a f.d.p condicional de $Y_1|Y_2$ é

$$f_{Y_1|Y_2}(\mathbf{y}_1) = \frac{g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) \prod_{k=1}^r \frac{y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}}}{\int_{A(\boldsymbol{\lambda}_1)} g(\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w}) d\mathbf{w}_1}, \quad \mathbf{w}_1 = T_{(\boldsymbol{\lambda}_1, \boldsymbol{\mu}_1)}(\mathbf{y}_1), \quad \mathbf{y}_1 \in \mathbb{R}_+^r.$$

Pela identidade (Seber [41])

$$\mathbf{w}'\boldsymbol{\Sigma}^{-1}\mathbf{w} = (\mathbf{w}_1 - \boldsymbol{\mu}_1(\mathbf{w}_2))'\boldsymbol{\Sigma}_{11.2}^{-1}(\mathbf{w}_1 - \boldsymbol{\mu}_1(\mathbf{w}_2)) + q(\mathbf{w}_2),$$

com $\boldsymbol{\Sigma}_{11.2}$ e $q(\mathbf{w}_2)$ dados em (3.18), temos que

$$\begin{aligned} f_{Y_1|Y_2}(\mathbf{y}_1) &= \frac{g_{q(\mathbf{w}_2)}((\mathbf{w}_1 - \boldsymbol{\mu}_1(\mathbf{w}_2))'\boldsymbol{\Sigma}_{11.2}^{-1}(\mathbf{w}_1 - \boldsymbol{\mu}_1(\mathbf{w}_2))) \prod_{k=1}^r \frac{y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}}}{\int_{A(\boldsymbol{\lambda}_1)} g_{q(\mathbf{w}_2)}((\mathbf{w}_1 - \boldsymbol{\mu}_1(\mathbf{w}_2))'\boldsymbol{\Sigma}_{11.2}^{-1}(\mathbf{w}_1 - \boldsymbol{\mu}_1(\mathbf{w}_2))) d\mathbf{w}_1} \\ &= \frac{g_{q(\mathbf{w}_2)}(\mathbf{u}_1'(\mathbf{D}_1\boldsymbol{\Sigma}_{11.2}\mathbf{D}_1)^{-1}\mathbf{u}_1) \prod_{k=1}^r \frac{y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}}}{\int_{A(\boldsymbol{\lambda}_1)} g_{q(\mathbf{w}_2)}(\mathbf{u}_1'(\mathbf{D}_1\boldsymbol{\Sigma}_{11.2}\mathbf{D}_1)^{-1}\mathbf{u}_1) d\mathbf{u}_1}, \quad \mathbf{y}_1 \in \mathbb{R}_+^r, \end{aligned}$$

em que $\mathbf{u}_1 = T_{(\boldsymbol{\lambda}_1, \boldsymbol{\delta}_1)}(\mathbf{y}_1)$. Fazendo $\mathbf{u}_1 = \mathbf{D}_1(\mathbf{w}_1 - \boldsymbol{\mu}_1(\mathbf{w}_2))$ na integral do denominador, com Jacobiano $J(\mathbf{w}_1 \rightarrow \mathbf{u}_1) = \prod_{k=1}^r (1 + \lambda_k \mu_{1k}(\mathbf{w}_2))$, e levando em consideração que se $\mathbf{w}_1 \in A(\boldsymbol{\lambda}_1)$, então $\mathbf{u}_1 \in A(\boldsymbol{\lambda}_1)$, temos que a f.d.p de $Y_1|Y_2$ é dada por

$$f_{Y_1|Y_2}(\mathbf{y}_1) = \frac{g_{q(\mathbf{w}_2)}(\mathbf{u}_1'(\mathbf{D}_1\boldsymbol{\Sigma}_{11.2}\mathbf{D}_1)^{-1}\mathbf{u}_1) \prod_{k=1}^r \frac{y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}(1+\lambda_k\mu_{1k}(\mathbf{w}_2))}}{\int_{A(\boldsymbol{\lambda}_1)} g_{q(\mathbf{w}_2)}(\mathbf{u}_1'(\mathbf{D}_1\boldsymbol{\Sigma}_{11.2}\mathbf{D}_1)^{-1}\mathbf{u}_1) d\mathbf{u}_1}, \quad \mathbf{y}_1 \in \mathbb{R}_+^p.$$

Por fim, notando que

$$\prod_{k=1}^r \frac{y_k^{\lambda_k-1}}{\mu_k^{\lambda_k}(1+\lambda_k\mu_{1k}(\mathbf{w}_2))} = \prod_{k=1}^r \frac{y_k^{\lambda_k-1}}{\delta_k^{\lambda_k}},$$

obtemos a f.d.p de $Y_1|Y_2$ como

$$f_{Y_1|Y_2}(\mathbf{y}_1) = \frac{g_{q(\mathbf{w}_2)}(\mathbf{u}'_1(\mathbf{D}_1\Sigma_{11.2}\mathbf{D}_1)^{-1}\mathbf{u}_1) \prod_{k=1}^r \frac{y_k^{\lambda_k-1}}{\delta_k^{\lambda_k}}}{\int_{A(\lambda_1)} g_{q(\mathbf{w}_2)}(\mathbf{u}'_1(\mathbf{D}_1\Sigma_{11.2}\mathbf{D}_1)^{-1}\mathbf{u}_1) d\mathbf{u}_1}, \quad \mathbf{u}_1 = T_{(\lambda_1, \delta_1)}(\mathbf{y}_1), \quad \mathbf{y}_1 \in \mathbb{R}_+^p.$$

Isto prova o teorema.

A.20 Prova Teorema 3.4.8

Pela definição de valor esperado e (3.6) temos que

$$E\left(\prod_{k=1}^p Y_k^{h_k}\right) = \frac{\int_{\mathbb{R}_+^p} g(\mathbf{w}'\Sigma^{-1}\mathbf{w}) \prod_{k=1}^p \frac{y_k^{\lambda_k+h_k-1}}{\mu_k^{\lambda_k}} d\mathbf{y}}{\int_{A(\lambda)} g(\mathbf{w}'\Sigma^{-1}\mathbf{w}) d\mathbf{w}},$$

em que $\mathbf{w} = T_{(\lambda, \mu)}(\mathbf{y})$. Transformando $\mathbf{u} = \mathbf{D}_\mu^{-1}\mathbf{y}$, com Jacobiano $J(\mathbf{y} \rightarrow \mathbf{u}) = \prod_{k=1}^p \mu_k$, na integral do numerador obtemos que

$$E\left(\prod_{k=1}^p Y_k^{h_k}\right) = \frac{(\prod_{k=1}^p \mu_k^{h_k}) \int_{\mathbb{R}_+^p} g(\mathbf{w}'\Sigma^{-1}\mathbf{w}) \prod_{k=1}^p u_k^{\lambda_k+h_k-1} d\mathbf{u}}{\int_{A(\lambda)} g(\mathbf{u}'\Sigma^{-1}\mathbf{u}) d\mathbf{u}},$$

em que, pelo item 1 do Teorema 3.2.2, $\mathbf{w} = T_{(\lambda, 1)}(\mathbf{u})$. Portanto,

$$E\left(\prod_{k=1}^p Y_k^{h_k}\right) = \left(\prod_{k=1}^p \mu_k^{h_k}\right) E\left(\prod_{k=1}^p U_k^{h_k}\right).$$

A.21 Prova Corolário 3.4.8.2

Usando (3.19) temos que, para $i \neq j$, $\text{Cov}(Y_i, Y_j) = \mu_i \mu_j \text{Cov}(U_i, U_j)$; e, para $i = j$, $\text{Var}(Y_i) = \mu_i^2 \text{Var}(U_i)$, $i = 1, \dots, p$, $j = 1, \dots, p$. Portanto,

$$\Sigma_Y = \mathbf{D}_\mu \Sigma_U \mathbf{D}_\mu,$$

Para provar a segunda parte do corolário note que para $i \neq j$

$$\begin{aligned} \text{Corr}(Y_i, Y_j) &= \frac{\mu_i \mu_j \text{Cov}(U_i, U_j)}{\sqrt{\mu_i^2 \text{Var}(U_i) \mu_j^2 \text{Var}(U_j)}} \\ &= \text{Corr}(U_i, U_j), \end{aligned}$$

e para $i = j$, $\text{Corr}(Y_i, Y_j) = 1$. Portanto,

$$\rho_Y = \rho_Z.$$

A.22 Prova Teorema 3.5.1

Como $Y \sim \text{BCE}\ell_p(\mu, \lambda, \Sigma; g)$, então a f.d.p de Y_k , $k = 1, \dots, p$, é dada por (3.24), em que $S_k = \sigma_{kk}^{-1/2} T_{(\lambda_k, \mu_k)}(Y_k)$ tem f.d.p dada por (3.28). A f.d.a de S_k é calculada como

$$F_{S_k}(s_k) = \begin{cases} \frac{F_{U_k}(s_k) - F_{U_k}(-1/\lambda_k \sqrt{\sigma_{kk}})}{1 - F_{U_k}(-1/\lambda_k \sqrt{\sigma_{kk}})}, & \text{se } \lambda_k > 0, \\ \frac{1 - F_{U_k}(-1/\lambda_k \sqrt{\sigma_{kk}}) + F_{U_k}(s_k)}{F_{U_k}(-1/\lambda_k \sqrt{\sigma_{kk}})}, & \text{se } \lambda_k < 0, \\ F_{U_k}(s_k), & \text{se } \lambda_k = 0. \end{cases} \quad (\text{A.4})$$

Assim, usando a relação entre Y_k e S_k , através da transformação de Box-Cox estendida univariada, temos que

$$\begin{aligned} \text{P}(Y_k \leq y_{k,\alpha}) = \alpha &\Leftrightarrow \text{P}[T_{(\lambda_k, \mu_k)}(Y_k) \leq T_{(\lambda_k, \mu_k)}(y_{k,\alpha})] = \alpha \\ &\Leftrightarrow \text{P}[S_k \leq \sigma_{kk}^{-1/2} T_{(\lambda_k, \mu_k)}(y_{k,\alpha})] = \alpha \\ &\Leftrightarrow y_{k,\alpha} = T_{(\lambda_k, \mu_k)}^{-1}(\sqrt{\sigma_{kk}} s_{k,\alpha}) \\ &\Leftrightarrow y_{k,\alpha} = \begin{cases} \mu_k (1 + \lambda_k \sqrt{\sigma_{kk}} s_{k,\alpha})^{1/\lambda_k}, & \text{se } \lambda_k \neq 0, \\ \mu_k \exp(\sqrt{\sigma_{kk}} s_{k,\alpha}), & \text{se } \lambda_k = 0, \end{cases} \end{aligned}$$

em que $s_{k,\alpha}$ é tal que $F_{S_k}(s_{k,\alpha}) = \alpha$, com F_{S_k} dada em (A.4). Portanto, $s_{k,\alpha}$ é obtido como

$$s_{k,\alpha} = \begin{cases} F_{U_k}^{-1}(\alpha + (1 - \alpha)F_{U_k}(-1/\lambda_k \sqrt{\sigma_{kk}})), & \text{se } \lambda_k > 0, \\ F_{U_k}^{-1}((1 + \alpha)F_{U_k}(-1/\lambda_k \sqrt{\sigma_{kk}}) - 1), & \text{se } \lambda_k < 0, \\ F_{U_k}^{-1}(\alpha), & \text{se } \lambda_k = 0, \end{cases}$$

em que F_{U_k} é a f.d.a dada em (3.27).

A.23 Prova Corolário 3.5.1.1

1. Considere $\lambda = 0$ no Teorema 3.5.1. Portanto, o α -quantil $y_{k,\alpha}$ de Y_k , $k = 1, \dots, p$, é dado por $y_{k,\alpha} = \mu_k \exp(\sqrt{\sigma_{kk}} s_{k,\alpha})$, em que $s_{k,\alpha} = F_{U_k}^{-1}(\alpha)$, sendo U_k a variável aleatória com f.d.p dada por (3.31). Isto vem do fato de que se $\lambda = 0$, então a

região de integração que envolve (3.25) é tal que $A(\boldsymbol{\lambda}_{-k}) = \mathbb{R}^{p-1}$. Portanto, a f.d.p (3.26) e dada por (3.31). Assim, $s_{k,\alpha}$ é o α -quantil de uma distribuição simétrica padrão com f.g.d dada por (3.30).

2. Seja $\mathbf{Y} \sim \text{BCE}_{\ell_p}(\boldsymbol{\mu}, \boldsymbol{\lambda}, \boldsymbol{\Sigma}; g)$, com $\boldsymbol{\lambda} \neq \mathbf{0}$. Dado que $\lambda_j \sqrt{\sigma_{jj}} \rightarrow 0$, $j = 1, \dots, p$, então das equações (3.32) e (3.33) dadas no Teorema 3.5.1, segue que o α -quantil de Y_k , $\alpha \in (0, 1)$, $k = 1, \dots, p$, é dado por $y_{k,\alpha} = \mu_k \exp(\sqrt{\sigma_{kk}} s_{k,\alpha})$, com $s_{k,\alpha} = F_{U_k}^{-1}(\alpha)$, sendo U_k a variável aleatória com f.d.p dada por (3.26), em que a função $g_{\boldsymbol{\Upsilon}_k}$ pode ser expressa como

$$g_{\boldsymbol{\Upsilon}_k}(u_k) = \int_{A(\boldsymbol{\Delta}_{-k,-k} \boldsymbol{\lambda}_{-k})} g((1 + \boldsymbol{\Upsilon}_k \boldsymbol{\Upsilon}_k') u_k^2 - 2\boldsymbol{\Upsilon}_k \boldsymbol{\Omega}_k \mathbf{w}_{-k} u_k + \mathbf{w}_{-k}' \boldsymbol{\Omega}_k' \boldsymbol{\Omega}_k \mathbf{w}_{-k}) d\mathbf{w}_{-k},$$

onde $u_k \in \mathbb{R}$, $\boldsymbol{\Delta} = \text{diag}\{\sqrt{\sigma_{11}}, \dots, \sqrt{\sigma_{pp}}\}$, $\boldsymbol{\Omega}_k = (\boldsymbol{\Sigma}_{-k,-k} - \sigma_{kk}^{-1} \boldsymbol{\Sigma}_{-k,k} \boldsymbol{\Sigma}_{k,-k})^{-1/2} \boldsymbol{\Delta}_{-k,-k}$ e $\boldsymbol{\Upsilon}_k = \sigma_{kk}^{-1/2} \boldsymbol{\Sigma}_{k,-k} \boldsymbol{\Omega}_k \boldsymbol{\Delta}_{-k,-k}^{-1}$. Como $\lambda_j \sqrt{\sigma_{jj}} \rightarrow 0$, $j = 1, \dots, p$, então $\boldsymbol{\Delta}_{-k,-k} \boldsymbol{\lambda}_{-k} \rightarrow \mathbf{0}$, $k = 1, \dots, p$. Notando que

$$\lim_{\boldsymbol{\Delta}_{-k,-k} \boldsymbol{\lambda}_{-k} \rightarrow \mathbf{0}} A(\boldsymbol{\Delta}_{-k,-k} \boldsymbol{\lambda}_{-k}) = \mathbb{R}^{p-1},$$

temos que a condição $\lambda_j \sqrt{\sigma_{jj}} \rightarrow 0$, $j = 1, \dots, p$, implica que

$$g_{\boldsymbol{\Upsilon}_k}(u_k) = \int_{\mathbb{R}^{p-1}} g((1 + \boldsymbol{\Upsilon}_k \boldsymbol{\Upsilon}_k') u_k^2 - 2\boldsymbol{\Upsilon}_k \boldsymbol{\Omega}_k \mathbf{w}_{-k} u_k + \mathbf{w}_{-k}' \boldsymbol{\Omega}_k' \boldsymbol{\Omega}_k \mathbf{w}_{-k}) d\mathbf{w}_{-k}.$$

Portanto, a f.d.p (3.26) e dada por (3.31). Isto prova que $s_{k,\alpha}$ é o α -quantil de uma distribuição simétrica padrão com f.g.d dada por (3.30).

A.24 Prova Teorema 4.4.1

1. Como $\mathbf{X} \sim \text{LN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, então pela Definição 3.3.1 segue que $\mathbf{W} = T_{(\mathbf{0}, \boldsymbol{\mu})}(\mathbf{X}) \sim \text{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Portanto, $\mathbf{W}' \boldsymbol{\Sigma}^{-1} \mathbf{W} \sim \chi_p^2$ (ver Tong [26]).
2. Similar à prova do item 1.

Apêndice B

Programação em R

```
#####  
## Esta função calcula a f.d.a sobre retângulos (determinados ##  
## pelos vértices l e u) da distribuição t multivariada de ##  
## parâmetros mu(p*1), sigma(p*p) e tau>0 graus de liberdade. ##  
#####  
library(sn)  
library(gtools)  
pmvt1 <- function(l,u,mu,sigma,tau){  
  p <- length(l)  
  subs <- function(x,y,j){  
    x[j] <- y[j]  
    x  
  }  
  points1 <- function(j){subs1 <- list();  
  coord <- function(i){coord1 <- matrix(NA,nr=choose(p,j),nc=p);  
  coord1 <- subs(u,l,combinations(p,j)[i,]);coord1}  
  i <- 1:choose(p,j)  
  subs1 <- t(sapply(i,coord))}  
  j <- 1:p  
  points <- sapply(j,points1)  
  vertices <- rbind(u,do.call(rbind,points),deparse.level = 0)  
  probs <- apply(vertices,1,pmst,xi=mu,Omega=sigma,  
                 alpha=rep(0,p),nu=tau)  
  sig1 <- function(i){sig <- c(NA);  
  sig <- -rep((-1)^i,choose(p,0:p)[i])}  
  i<-2:(p+1)  
  sig <- c(1,unlist(sapply(i,sig1)))  
  return(sum(probs*sig))  
}
```

```
#####
## Esta função gera um vetor coluna a partir dos elementos da ##
## diagonal principal de uma matriz X e abaixo dela.      ##
#####
vech <- function(X){
  rows <- nrow(X)
  columns <- ncol(X);
  pos <- function(i){v = c(); v = c(v, X[i:rows,i]); v}
  k <- 1:columns;
  u <- unlist(sapply(k,pos));
  u;
}

#####
## Esta função gera um vetor coluna a partir dos elementos da ##
## diagonal principal de uma matriz X e acima dela.      ##
#####
vech2 <- function(X){
  rows <- nrow(X)
  columns <- ncol(X);
  pos <- function(i){v = c(); v = c(v, X[i,i:columns]); v}
  k <- 1:rows;
  u <- unlist(sapply(k,pos));
  u;
}

#####
## Esta função gera uma matriz simétrica a partir dos compo- ##
## nentes do vetor v.                                     ##
#####
inv_vech <- function(v){
  p <- -0.5+sqrt(0.5^2+2*length(v))
  x <- matrix(NA,nrow=p,ncol=p)
  for (i in 1:p){
    x[i:p,i] <- v[((i-1)*(p-(i-2)*0.5)+1) : (i*(p-(i-1)*0.5))]
    x[i,i:p] <- v[((i-1)*(p-(i-2)*0.5)+1) : (i*(p-(i-1)*0.5))]
  }
  x
}

#####
## Esta função determina, a partir do vetor t, os limites in- ##
## feriores da região de integração da integral que envolve a ##
```

```

## f.d.p da distribuição Box-Cox elíptica. ##
#####
low <- function(t){
  linf <- function(i){l <- NA; ifelse(t[i] > 0, l <- -1/t[i],
                                     l <- -Inf); l}
  k <- 1:length(t);
  li <- sapply(k,linf);
  li;
}

#####
## Esta função determina, a partir do vetor t, os limites su- ##
## periores da região de integração da integral que envolve a ##
## f.d.p da distribuição Box-Cox elíptica. ##
#####
up <- function(t){
  lsup <- function(i){u <- NA; ifelse(t[i] < 0, u <- -1/t[i],
                                     u <- Inf); u}
  k <- 1:length(t);
  ls <- sapply(k,lsup);
  ls;
}

#####
## Esta função gera amostras aleatórias de tamanho n da dis- ##
## tribuição Box-Cox normal multivariada de parâmetros ##
## mu(p*1), lambda(p*1) e sigma(p*p). ##
#####
library(tmvtnorm)
rmvbcnorm <- function(n,mu,lambda,sigma){
  p <- length(mu);
  am1 <- rtmvnorm(n,mean=rep(0,length(mu)),sigma=sigma,
                 lower=low(lambda),upper=up(lambda),
                 algorithm="gibbs")
  tr <- function(j){y <- matrix(NA,nr=nrow(am1),nc=ncol(am1));
  ifelse(lambda[j] != 0,
         y <- (mu[j])*(1+(lambda[j])*(am1[,j]))^{1/lambda[j]},
         y <- (mu[j])*exp(am1[,j])); y}
  s <- 1:p
  am <- sapply(s,tr)
  return(am);
}

```

```
#####
## Esta função gera amostras aleatórias de tamanho n da dis- ##
## tribuição Box-Cox t multivariada de parâmetros mu(p*1), ##
## lambda(p*1), sigma(p*p) e tau>0 graus de liberdade. ##
#####
library(tmvtnorm)
rmvbct <- function(n,mu,sigma,lambda,tau){
  p <- length(mu);
  am1 <- rtmvt(n,mean=rep(0,length(mu)),sigma=sigma,df=tau,
              lower=low(lambda),upper=up(lambda),
              algorithm="gibbs");
  tr <- function(j){y <- matrix(NA,nr=nrow(am1),nc=ncol(am1));
  ifelse(lambda[j] != 0,
          y <- (mu[j])*(1+(lambda[j])*(am1[,j]))^{1/lambda[j]},
          y <- (mu[j])*exp(am1[,j])); y}
  s <- 1:p
  am <- sapply(s,tr)
  return(am);
}

#####
## Esta função estima os parâmetros mu(p*1), lambda(p*1) e ##
## sigma(p*p) da distribuição Box-Cox normal multivariada a ##
## partir da matriz de dados Y(n*p). ##
#####
library(mvtnorm)
mlemvbcnorm <- function(Y,
                        start=list(mu=rep(1,ncol(Y)),
                                   lambda=rep(0,ncol(Y)),
                                   sigma=diag(ncol(Y))),
                        fixed=list(), method="BFGS",
                        cholesky=FALSE,
                        lower.bounds=-Inf,
                        upper.bounds=+Inf,...)
{
  p <- ncol(Y)
  n <- nrow(Y)
  ifelse(cholesky, theta <- c(log(start$mu), start$lambda,
                              vech2(t(chol(start$sigma)))),
        theta <- c(log(start$mu), start$lambda,
                    vech2(start$sigma)))
  nmmu <- paste("mu_",1:p,sep="")

```



```

nmlambda <- paste("lambda_", 1:p, sep="")
nmsigma <- paste("sigma_", vech2(outer(1:p, 1:p, paste, sep="."))
                , sep="")
names(theta) <- c(nmmu, nmlambda, nmsigma)
negloglik <- function(x) {
  nf <- names(formals())
  theta <- sapply(nf, function(x) {eval(parse(text=x))})
  mu <- exp(theta[1:p])
  lambda <- theta[(p+1):(2*p)]
  if(cholesky) {
    L <- inv_vech(theta[-(1:(2*p))])
    L[lower.tri(L, diag=FALSE)] <- 0
    sigma <- t(L) %*% L
  }
  else {
    sigma <- inv_vech(theta[-(1:(2*p))])
  }
  g <- function(x, m, t, s) {
    tr <- function(j) {y <- matrix(NA, nr=n, nc=p);
      ifelse(t[j] != 0, y <- (1/t[j]) * ((x[,j]/m[j])^(t[j]-1),
        y <- log(x[,j]/m[j])); y}
    d <- 1:p;
    w <- sapply(d, tr)
    g1 <- dmvnorm(w, mean=rep(0, p), sigma=s, log=T)
    return(sum(g1))
  }
  h <- function(x, m, t) {
    te <- function(j) {h1 <- c(NA);
      h1 <- (t[j]-1) * sum(log(x[,j])) - n * (t[j]) * log(m[j])}
    e <- 1:p;
    h2 <- sapply(e, te)
    return(sum(h2))
  }
  f <- -(- n*log(pmvnorm(lower=low(lambda), upper=up(lambda),
                        mean=rep(0, p), sigma=sigma)) +
        g(Y, mu, lambda, sigma) + h(Y, mu, lambda))
  if(is.infinite(f) || is.na(f)) {return(.Machine$integer.max)}
  f
}
formals(negloglik) <- theta
if((length(lower.bounds) > 1L || length(upper.bounds) > 1L ||

```

```

    lower.bounds[1L] != -Inf || upper.bounds[1L] != Inf) &&
method == "L-BFGS-B"){mle.fit <-
  eval.parent(substitute(mle(negloglik, start=as.list(theta),
                            fixed=fixed, method = method,
                            lower=lower.bounds,
                            upper=upper.bounds, ...)))

  return (mle.fit)
}
else{
  mle.fit <- eval.parent(substitute(
    mle(negloglik, start=as.list(theta), fixed=fixed,
        method = method, ...)))
  return (mle.fit)
}
}

#####
## Esta função estima os parâmetros  $\mu(p \times 1)$ ,  $\lambda(p \times 1)$ , ##
##  $\sigma(p \times p)$  e  $\tau > 0$  da distribuição Box-Cox  $t$  multivariada ##
## a partir da matriz de dados  $Y(n \times p)$ . ##
#####
library(mvtnorm)
mlemvbct <- function(Y,
                    start=list(mu=rep(1, ncol(Y)),
                               lambda=rep(0, ncol(Y)),
                               sigma=diag(ncol(Y)),
                               tau=1),
                    fixed=list(), method="BFGS",
                    cholesky=FALSE,
                    lower.bounds=-Inf,
                    upper.bounds=+Inf,
                    ...)
{
  p <- ncol(Y)
  n <- nrow(Y)
  ifelse(cholesky, theta <- c(log(start$mu), start$lambda,
                              vech2(t(chol(start$sigma)))),
        log(start$tau)),
        theta <- c(log(start$mu), start$lambda,
                  vech2(start$sigma), log(start$tau)))
  nmmu <- paste("mu_", 1:p, sep="")
  nmlambda <- paste("lambda_", 1:p, sep="")

```

```

nmsigma <- paste("sigma_",
                vech2(outer(1:p,1:p, paste, sep=".")),sep="")
nmtau <- paste("tau")
names(theta) <- c(nmmu, nmlambda, nmsigma,nmtau)
negloglik <- function(x){
  nf <- names(formals())
  theta <- sapply(nf,function(x){eval(parse(text=x))})
  mean <- exp(theta[1:p])
  lambda <- theta[(p+1):(2*p)]
  tau <- exp(theta[1 + (p*(p+5)/2)])
  if(cholesky){
    L <- inv_vech(theta[(2*p+1):(p*(p+5)/2)])
    L[lower.tri(L, diag=FALSE)] <- 0
    sigma <- t(L) %*% L
  }
  else{
    sigma <- inv_vech(theta[(2*p+1):(p*(p+5)/2)])
  }
  g <- function(x,m,t,s,r){
    tr <- function(j){y <- matrix(NA,nr=n ,nc=p);
    ifelse(t[j] != 0, y <- (1/t[j])*((x[,j]/m[j])^(t[j]-1),
    y <- log(x[,j]/m[j])); y}
    d <- 1:p;
    w <- sapply(d,tr)
    g1 <- dmvt(x=w, delta=rep(0,p), sigma=s, df=r, log=TRUE)
    return(sum(g1))
  }
  h <- function(x,m,t){
    te <- function(j){h1<-c(NA);
    h1 <- (t[j]-1)*sum(log(x[,j]))-n*(t[j])*log(m[j])}
    e <- 1:p;
    h2 <- sapply(e,te)
    return(sum(h2))
  }
  f <- -(- n*log(pmvtn1(l=low(lambda), u=up(lambda),
    mu=rep(0,p), sigma=sigma, tau=tau)) +
    g(Y,mean,lambda,sigma,tau) + h(Y,mean,lambda))
  if(is.infinite(f) || is.na(f)){return(.Machine$integer.max)}
  f
}
formals(negloglik) <- theta

```



```

theta <- c(c(start$B), start$lambda, vech2(start$sigma))
nmbeta <- paste("beta_", outer(1:r, 1:p, paste, sep="."),
              sep="")
nmlambda <- paste("lambda_", 1:p, sep="")
nmsigma <- paste("sigma_", vech2(outer(1:p, 1:p, paste,
                                      sep=".")), sep="")
names(theta) <- c(nmbeta, nmlambda, nmsigma)
negloglik <- function(x){
  nf <- names(formals())
  theta <- sapply(nf, function(x){eval(parse(text=x))})
  beta <- matrix(theta[1:(p*r)], nr=r, nc=p, byrow=F)
  lambda <- theta[(p*r)+1:(p*(r+1))]
  if(cholesky){
    L <- inv_vech(theta[(1+p*(r+1)):(p*(r+((p+3)/2))]))
    L[lower.tri(L, diag=FALSE)] <- 0
    sigma <- t(L) %*% L
  }
  else{
    sigma <- inv_vech(theta[(1+p*(r+1)):(p*(r+((p+3)/2))]))
  }
  g <- function(y, x, b, t, s){
    tr <- function(j){z <- matrix(NA, nr=n, nc=p);
    ifelse(t[j] != 0,
           z <- (1/t[j])*
             ((y[,j]/(exp((x%*%b)[,j])))^(t[j]-1)),
           z <- log(y[,j]) - (x%*%b)[,j]); z}
    d = 1:p;
    w <- sapply(d, tr)
    g1 <- dmvnorm(x=w, mean=rep(0, p), sigma=s, log=T)
    return(sum(g1))
  }
  h <- function(y, x, b, t){
    te = function(j){h1<-c(NA);
    h1 <- (t[j]-1)*sum(log(y[,j])) - (t[j])*sum((x%*%b)[,j])}
    e = 1:p;
    h2 <- sapply(e, te)
    return(sum(h2))
  }
  f <- -(- n*log(pmvnorm(lower=low(nu), upper=up(nu),
                        mean=rep(0, p), sigma=sigma)) +
        g(Y, X, beta, lambda, sigma) + h(Y, X, beta, lambda))

```

```

    if(is.infinite(f) || is.na(f)){
      return(.Machine$integer.max)
    }
  }
  formals(negloglik) <- theta
  if((length(lower.bounds) > 1L || length(upper.bounds) > 1L
    || lower.bounds[1L] != -Inf || upper.bounds[1L] != Inf)
    && method == "L-BFGS-B"){mle.fit <-
    eval.parent(substitute(mle(negloglik,
                              start=as.list(theta),
                              fixed=fixed, method = method,
                              lower=lower.bounds,
                              upper=upper.bounds, ...)))

    return (mle.fit)
  }
  else{
    mle.fit <- eval.parent(substitute(
      mle(negloglik, start=as.list(theta), fixed=fixed,
          method = method, ...)))
    return(mle.fit)
  }
}

#####
## Esta função estima os parâmetros B(r*p), lambda(p*1), ##
## sigma(p*p) e tau>0 do modelo de regressão linear Box-Cox t ##
## multivariado, a partir da matriz de vetores resposta ##
## Y(n*p) e a matriz de variáveis explicativas X(n*r). ##
#####
library(mvtnorm)
mleregmbvbt <- function(Y,X,
                        start=list(B=matrix(0,nr=ncol(X),
                                           nc=ncol(Y)),
                                   lambda=rep(0,ncol(Y)),
                                   sigma=diag(ncol(Y)),
                                   tau=1),
                        fixed=list(),
                        method="BFGS",
                        cholesky=FALSE,
                        lower.bounds=-Inf,
                        upper.bounds=+Inf,
                        ...)

```

```

{
  X=as.matrix(X)
  p = ncol(Y)
  n = nrow(Y)
  r = ncol(X)
  ifelse(cholesky, theta <- c(c(start$B), start$lambda,
                             vech2(t(chol(start$sigma))),
                             log(start$tau)),
         theta <- c(c(start$B), start$lambda, vech2(start$sigma),
                    log(start$tau)))
  nmbeta <- paste("beta_", outer(1:r, 1:p, paste, sep="."), sep="")
  nmlambda <- paste("lambda_", 1:p, sep="")
  nmsigma <- paste("sigma_", vech2(outer(1:p, 1:p, paste, sep=".")),
                  sep="")
  nmtau <- paste("tau")
  names(theta) <- c(nmbeta, nmlambda, nmsigma, nmtau)
  negloglik <- function(x) {
    nf <- names(formals())
    theta <- sapply(nf, function(x) {eval(parse(text=x))})
    beta <- matrix(theta[1:(p*r)], nr=r, nc=p, byrow=F)
    lambda <- theta[((p*r)+1):(p*(r+1))]
    tau = exp(theta[1 + (p*(r+((p+3)/2))])
    if(cholesky) {
      L <- inv_vech(theta[(1+p*(r+1)):(p*(r+((p+3)/2))])
      L[lower.tri(L, diag=FALSE)] <- 0
      sigma <- t(L) %*% L
    }
    else{
      sigma <- inv_vech(theta[(1+p*(r+1)):(p*(r+((p+3)/2))])
    }
    g <- function(y, x, b, t, s, r) {
      tr <- function(j) {z <- matrix(NA, nr=n, nc=p);
        ifelse(t[j] != 0,
              z <- (1/t[j]) * ((y[, j] / (exp((x%*%b)[, j])))^(t[j]-1),
              z <- log(y[, j]) - (x%*%b)[, j]); z}
      d = 1:p;
      w <- sapply(d, tr)
      g1 <- dmvtn(x=w, delta=rep(0, p), sigma=s, df=tau, log=T)
      return(sum(g1))
    }
    h <- function(y, x, b, t) {

```

```

te = function(j) {h1<-c(NA);
h1 <- (t[j]-1)*sum(log(y[,j]))-(t[j])*sum((x%*%b)[,j])}
e = 1:p;
h2 <- sapply(e,te)
return(sum(h2))
}
f <- -(- n*log(pmvtn(l=low(lambda), u=up(lambda), mu=rep(0,p),
sigma=sigma, tau=tau)) +
g(Y,X,beta,lambda,sigma,tau) + h(Y,X,beta,lambda))
if(is.infinite(f) || is.na(f)){return(.Machine$integer.max)}
f
}
formals(negloglik) <- theta
if((length(lower.bounds) > 1L || length(upper.bounds) > 1L ||
lower.bounds[1L] != -Inf || upper.bounds[1L] != Inf)
&& method == "L-BFGS-B"){mle.fit <-
eval.parent(substitute(mle(negloglik, start=as.list(theta),
fixed=fixed, method = method,
lower=lower.bounds,
upper=upper.bounds, ...)))

return (mle.fit)
}
else{
mle.fit <- eval.parent(substitute(
mle(negloglik, start=as.list(theta), fixed=fixed,
method = method, ...)))
return(mle.fit)
}
}

```


Referências Bibliográficas

- [1] J. A. Tooze, V. Kipnis, D. W. Buckman, R. J. Carroll, L. S. Freedman, P. M. Guenther, S. M. Krebs-Smith, A. F. Subar, e K. W. Dodd, “A mixed-effects model approach for estimating the distribution of usual intake of nutrients: the NCI method,” *Statistics in Medicine*, vol. 29, no. 27, pags. 2857–2868, (2010). 9
- [2] G. D. Pimentel, K. C. Portero-McLellan, N. Maestá, J. E. Corrente, e R. C. Burini, “Accuracy of sagittal abdominal diameter as predictor of abdominal fat among brazilian adults: a comparison with waist circumference,” *Nutrición Hospitalaria*, vol. 25, no. 4, pags. 656–661, (2010). 9
- [3] K. L. Lange, R. J. A. Little, e J. M. G. Taylor, “Robust statistical modeling using the t distribution,” *Journal of the American Statistical Association*, vol. 84, no. 408, pags. 881–896, (1989). 9, 31, 63
- [4] S. L. P. Ferrari e G. Fumes, “Box-Cox symmetric distributions and applications to nutritional data,” *Advances in Statistical Analysis*, DOI: 10.1007/s10182-017-0291-6. 9, 33, 36, 40, 65
- [5] L. H. Vanegas e G. Paula, “Log-symmetric distributions: statistical properties and parameter estimation,” *Brazilian Journal of Probability and Statistics*, vol. 30, no. 2, pags. 196–220, (2016). 9, 33, 36
- [6] T. J. Cole e P. J. Green, “Smoothing reference centile curves: the LMS method and penalized likelihood,” *Statistics in Medicine*, vol. 11, no. 10, pags. 1305–1319, (1992). 9, 32, 65
- [7] R. A. Rigby e D. M. Stasinopoulos, “Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution,” *Statistics in Medicine*, vol. 23, no. 19, pags. 3053–3076, (2004). 9, 32, 65
- [8] ———, “Using the Box-Cox t distribution in GAMLSS to model skewness and kurtosis,” *Statistical Modelling*, vol. 6, no. 3, pags. 209–229, (2006). 9, 32, 47, 54, 65
- [9] K. T. Fang, S. Kotz, e K. W. NG, *Symmetric Multivariate and Related Distributions*. London: Chapman and Hall, (1990). 10, 13, 33, 35, 41, 43, 80, 85
- [10] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, (2016). Disponível em <https://www.R-project.org/> 11, 14, 28, 49
- [11] D. Kelker, “Distribution theory of spherical distributions and a location-scale parameter generalization,” *Sankhyā: The Indian Journal of Statistics*, vol. 32, no. 4, pags. 419–430, (1970). 13

- [12] Z. Birnbaum e P. L. Meyer, "On the effect of truncation in some or all coordinates of a multi-normal population," *Journal of the Indian Society of Agricultural Statistics*, vol. 5, no. 1, p. 17–28, (1953). 13
- [13] G. M. Tallis, "The moment generating function of the truncated multi-normal distribution," *Journal of the Royal Statistical Society*, vol. 23, no. 1, pags. 223–229, (1961). 14
- [14] —, "Elliptical and radial truncation in normal populations," *The Annals of Mathematical Statistics*, vol. 34, no. 3, pags. 940–944, (1963). 14
- [15] —, "Plane truncation in normal populations," *Journal of the Royal Statistical Society*, vol. 27, no. 2, pags. 301–307, (1965). 14
- [16] W. C. Horrace, "On ranking and selection from independent truncated normal distributions," *Journal of Econometrics*, vol. 126, no. 2, pags. 335–354, (2005). 14
- [17] B. G. Manjunath e S. Wilhelm, "Moments calculation for the double truncated multivariate normal density," *Working paper*, (2012). Disponível em arXiv:1206.5387v1 14
- [18] H. J. Ho, T. I. Lin, H. Y. Chen, e W. L. Wang, "Some results on the truncated multivariate t distribution," *Journal of Statistical Planning and Inference*, vol. 142, no. 1, pags. 25–40, (2012). 14
- [19] J. Geweke, *Efficient simulation from the multivariate normal and Student- t distributions subject to linear constraints and the evaluation of constraint probabilities*, (1991), Relatório técnico, University of Minnesota, Department of Economics. 14
- [20] S. Wilhelm e B. G. Manjunath, *tmvtnorm: truncated multivariate normal and Student t distribution*, (2015), R package version 1.4-10. Disponível em <http://CRAN.R-project.org/package=tmvtnorm> 14, 28
- [21] A. Genz and F. Bretz and T. Miwa and X. Mi and F. Leisch and F. Scheipl and T. Hothorn, *mvtnorm: multivariate normal and t distributions*, (2014), R package version 1.0-2. Disponível em <http://CRAN.R-project.org/package=mvtnorm> 14, 49
- [22] A. Genz e F. Bretz, *Computation of Multivariate Normal and t Probabilities*. Heidelberg: Springer-Verlag, (2009). 14, 49
- [23] A. K. Gupta, T. Varga, e T. Bodnar, *Elliptically Contoured Models in Statistics and Portfolio Theory*. New York: Springer-Verlag, (2013). 16, 17, 80
- [24] E. Gómez, M. A. Gomez-Villegas, e J. M. Marín, "A multivariate generalization of the power exponential family of distributions," *Communications in Statistics - Theory and Methods*, vol. 27, no. 3, pags. 589–600, (1998). 16, 41
- [25] J. Wang e M. G. Genton, "The multivariate skew-slash distribution," *Journal of Statistical Planning and Inference*, vol. 136, pags. 209–220, (2006). 17
- [26] Y. L. Tong, *The Multivariate Normal Distribution*. New York: Springer-Verlag, (1990). 25, 79, 89
- [27] A. Prékopa, "Logarithmic concave measures with application to stochastic programming," *Acta Scientiarum Mathematicarum*, vol. 32, no. 1, p. 301–316, (1971). 25
- [28] J. von Neumann, "Various techniques used in connection with random digits," *Journal of Research of the National Bureau of Standards*, vol. 12, pags. 36–38, (1951). 28

- [29] G. E. P. Box e D. R. Cox, "An analysis of transformations," *Journal of the Royal Statistical Society*, vol. 26, no. 2, pags. 211–252, (1964). 31
- [30] R. M. Sakia, "Box-Cox transformation technique: A review," *Journal of the Royal Statistical Society*, vol. 41, no. 2, pags. 169–178, (1992). 31
- [31] D. F. Andrews, R. Gnanadesikan, e J. L. Warner, "Transformations of multivariate data," *Biometrika*, vol. 27, no. 4, pags. 825–840, (1971). 32
- [32] S. Velilla, "A note on the multivariate Box-Cox transformation to normality," *Statistics & Probability Letters*, vol. 17, no. 4, pags. 259–263, (1993). 32
- [33] A. J. Quiroz, M. Nakamura, e F. J. Perez, "Estimation of a multivariate Box-Cox transformation to elliptical symmetry via the empirical characteristic function," *Annals of the Institute of Statistical Mathematics*, vol. 48, no. 4, pags. 687–709, (1996). 32
- [34] M. Goto e T. Hamasaki, "The bivariate power-normal distribution," *Bulletin of Informatics and Cybernetics*, vol. 34, no. 1, pags. 29–49, (2002). 32, 34
- [35] T. I. Lin, "Robust mixture modeling using multivariate skew t distributions," *Statistics and Computing*, vol. 20, no. 3, pags. 343–356, (2010). 32
- [36] K. Lo e R. Gottardo, "Flexible mixture modeling via the multivariate t distribution with the Box-Cox transformation: an alternative to the skew- t distribution," *Statistics and Computing*, vol. 22, no. 1, pags. 33–52, (2012). 32
- [37] T. J. Cole, "Fitting smoothed centile curves to reference data," *Journal of the Royal Statistical Society*, vol. 151, no. 3, pags. 385–418, (1988). 32
- [38] J. Nocedal e S. J. Wright, *Numerical Optimization*. New York: Springer, (2006). 49, 67
- [39] R. A. Rigby e D. M. Stasinopoulos, "Generalized additive models for location, scale and shape," *Journal of the Royal Statistical Society*, vol. 54, no. 3, pags. 507–554, (2005). 50, 67
- [40] A. Azzalini, *The R package sn: The Skew-Normal and Skew-t distributions*, R package version 1.4-0. 50, 63, 64, 67, 69
- [41] G. A. F. Seber, *A Matrix Handbook for Statisticians*. New Jersey: John Wiley & Sons, Inc., (2007). 50, 80, 86
- [42] P. J. Rousseeuw, I. Ruts, e J. W. Tukey, "The bagplot: A bivariate boxplot," *The American Statistician*, vol. 53, no. 4, pags. 382–387, (1999). 51
- [43] M. D. Branco e D. K. Dey, "A general class of multivariate skew-elliptical distributions," *Journal of Multivariate Analysis*, vol. 79, no. 1, pags. 99–113, (2001). 55
- [44] A. Azzalini, *The Skew-Normal and Related Families*. Cambridge: Cambridge University Press, (2013). 55
- [45] P. K. Sen, J. M. Singer, e A. C. P. de Lima, *From Finite Sample to Asymptotic Methods in Statistics*. Cambridge: Cambridge University Press, (2009). 55, 70
- [46] G. Heinz, L. J. Peterson, R. W. Johnson, e C. J. Kerk, "Exploring relationships in body dimensions," *Journal of Statistics Education*, vol. 11, no. 2, (2003). Disponível em <https://ww2.amstat.org/publications/jse/v11n2/datasets.heinz.html> 57

- [47] M. Iwasaki e H. Tsubaki, "A bivariate generalized linear model with an application to meteorological data analysis," *Statistical Methodology*, vol. 2, no. 3, pags. 175–190, (2005). 64
- [48] W. H. Bonat e B. Jørgensen, "Multivariate covariance generalized linear models," *Journal of the Royal Statistical Society*, vol. 65, no. 5, pags. 649–675, (2016). 64
- [49] M. J. R. Healy, "Multivariate normal plotting," *Journal of the Royal Statistical Society*, vol. 17, no. 2, pags. 157–161. 64, 68
- [50] L. H. Vanegas e G. A. Paula, "An extension of log-symmetric regression models: R codes and applications," *Journal of Statistical Computation and Simulation*, vol. 86, no. 9, pags. 1709–1735, (2016). 65