

Universidade Federal de Juiz de Fora
Instituto de Ciências Exatas
Programa de Pós-Graduação em Matemática

Mariane dos Santos Bispo

**Estimadores de máxima verossimilhança L_q para todos os parâmetros da
distribuição Skew-t**

Juiz de Fora

2020

Mariane dos Santos Bispo

**Estimadores de máxima verossimilhança L_q para todos os parâmetros da
distribuição Skew-t**

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal de Juiz de Fora, na área de concentração em Matemática Aplicada, como requisito parcial para obtenção do título de Mestre em Matemática.

Orientadora: Camila Borelli Zeller

Juiz de Fora

2020

Ficha catalográfica elaborada através do Modelo Latex do CDC da
UFJF com os dados fornecidos pelo(a) autor(a)

Bispo, Mariane dos Santos.

Estimadores de máxima verossimilhança L_q para todos os parâmetros da distribuição Skew-t / Mariane dos Santos Bispo. – 2020.

103 f. : il.

Orientadora: Camila Borelli Zeller

Dissertação (Mestrado) – Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas. Programa de Pós-Graduação em Matemática, 2020.

1. Algoritmo EM. 2. Assimetria. 3. Distância de Mahalanobis. 4. Distribuição Skew-t. 5. Método de Máxima Verossimilhança L_q . 6. Robustez. I. Zeller, Camila Borelli, orient. II. Título.

Mariane dos Santos Bispo

Estimadores de Máxima Verossimilhança Lq para todos os parâmetros da distribuição Skew-t

Dissertação apresentada ao Programa de Pós-graduação em Matemática da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Matemática.
Área de concentração: Matemática Aplicada


Aprovada em 18 de dezembro de 2020

BANCA EXAMINADORA



Prof. Dr.ª Camila Borelli Zeller - Orientadora

Universidade Federal de Juiz de Fora



Prof. Dr. Rocío Paola Maehara Aliaga

Universidade del Pacífico



Prof. Dr. Tiago Maia Magalhães

Universidade Federal de Juiz de Fora

Dedico este trabalho a meu avô Gileno pedreiro e ao meu padrinho Messias (in memoriam), com todo amor e gratidão pelo tempo que passamos juntos, que nunca terá sido o suficiente.

AGRADECIMENTOS

Agradeço primeiramente a Deus, pois sem Ele não sou nada. Agradeço a CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pelo apoio financeiro, sem o qual não seria possível a realização deste trabalho. Agradeço a minha orientadora Camila, pela enorme paciência, apoio, compreensão e pelas risadas em meio as reuniões e conversas sobre a vida.

Agradeço ao corpo docente dos departamentos de Matemática e Estatística da UFJF que fez parte da minha jornada e que me deu ensinamentos valiosos, não esquecendo de agradecer a Paula, que mais que uma secretaria do departamento se tornou uma amiga. Agradeço também ao corpo docente do DMA da UFS, onde tive minha base. Em especial ao professor Hassan Sherafat por quem tenho um grande carinho e que me fez continuar na matemática. Agradeço aos meus “pais” da graduação, meu orientador Evilson e a sua esposa Ana, que sempre tiveram uma palavra amiga a me dizer, conselhos e incentivos para perseverar e alcançar novos voos.

Agradeço a todos os amigos que me ajudaram a enfrentar essa árdua jornada e que tornaram tudo sempre mais leve. Agradeço a todos os novos amigos que fiz no mestrado seja da salinha ou da sala grande, saibam que todos foram importantes de alguma forma, mas em especial gostaria de agradecer a Italo, Rodrigo, Julio, Sebastian e Wilian, pelas risadas e cafés. Agradeço também aos amigos mais antigos que a todo momento (mesmo longe) mandavam mensagens de apoio e coragem. Lenda, Mara, Marcos, Alice, Denis, Laura, Ian, Lueidsom, João, Igor... Amigos de fases diferentes mas que me completam de formas diversas.

Um agradecimento em especial a duas pessoas que são mais que amigos, são família. Italo e Jéssica, que desde os primórdios da graduação estavam ao meu lado, principalmente a Jéssica por toda a convivência nesses dois últimos anos morando juntas.

Um agradecimento a Welton e Sabrina, os melhores vizinhos que alguém poderia ter em uma cidade desconhecida.

Agradeço aos meus pais, Antônio e Maria, ao meu irmão Henrique, as minhas avós Dona Helena e Dona Nair, a minha madrinha Inês, todos os tios e tias, primos e primas que fizeram o possível para me ajudar a enfrentar essa jornada do mestrado.

Agradeço também ao meu computador que não me deixou na mão ao escrever essa dissertação. E por fim, Italo e Jéssica, nós sobrevivemos!

“Todos mudamos, se pensar nisso. Somos todos pessoas diferentes através de nossa vida.
E isso é bom, temos que continuar andando, desde que você se lembre de todas as
pessoas que já foi. ” Doctor Who

RESUMO

Neste trabalho, estudamos alguns aspectos de estimação dos parâmetros na classe de distribuições misturas de escala skew-normal, especificamente no modelo Skew-t (ST). A distribuição ST possui propriedades interessantes e importantes, como o fato de admitir uma representação estocástica que facilita a implementação do algoritmo EM, permite combinar assimetria com as caudas pesadas, possui como casos particulares as distribuições Skew-Cauchy, Skew-normal, t-Student e normal, além de atribuir pesos diferentes para cada observação e conseqüentemente controlar a influência da observação no processo de estimação. A estimação dos parâmetros será feita pelo método de estimação de máxima verossimilhança L_q (ML_q). Devido à representação de mistura de escala da distribuição ST, adaptamos o algoritmo EM para obter os estimadores de ML_q para todos os parâmetros do modelo ST, inclusive para os graus de liberdade. Além disso, discutimos a robustez deste método de estimação sob o modelo de interesse.

Palavras-chave: Algoritmo EM. Assimetria. Distância de Mahalanobis. Distribuição Skew-t. Método de Máxima Verossimilhança L_q . Robustez.

ABSTRACT

In this work, some aspects of estimation in the class of scale mixtures skew-normal distributions were taken into account, specifically in the Skew-t (ST) model. The ST distribution has interesting and important properties, such as the fact of admitting a stochastic representation that facilitates the implementation of the EM algorithm, allows combining asymmetry with heavy tails, it also has as particular cases of distributions the Skew-Cauchy, Skew-Normal, t-Student and normal, besides assigning different weights to each observation and consequently controlling the influence of observation in the estimation process. The parameters will be estimated using the maximum L_q -likelihood estimation method (ML_q). Due to the scale mixture representation of the ST distribution, we adapted the EM algorithm for the estimators ML_q for all ST model parameters, including the degrees of freedom. In addition, we discuss the robustness of this estimation method under the model of interest.

Keywords: EM Algorithm. Skewness. Mahalanobis Distance. Skew-t Distribution. Maximum L_q Likelihood Method. Robustness.

SUMÁRIO

1	INTRODUÇÃO	10
1.1	MOTIVAÇÃO	10
1.2	OUTLIERS E ROBUSTEZ	11
1.3	OBJETIVOS	13
1.4	ORGANIZAÇÃO DO TRABALHO	13
2	CONCEITOS PRELIMINARES	15
2.1	DISTRIBUIÇÕES MISTURA DE ESCALA SKEW-NORMAL (SMSN)	15
2.2	DISTRIBUIÇÃO SKEW-T (ST)	17
2.3	DISTÂNCIA DE MAHALANOBIS	20
2.4	ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA (ML)	21
2.4.1	Algoritmo EM	24
2.4.1.1	Algoritmo EM no modelo ST	25
3	ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA L_q VIA ALGORITMO EM: UMA ESTIMAÇÃO ROBUSTA DO MO- DELO SKEW-T	30
3.1	INTRODUÇÃO	30
3.2	ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA L_q	31
3.3	ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA L_q VIA ALGO- RITMO EM	32
3.3.1	Algoritmo EM_q no modelo ST	32
3.4	ESCOLHA DO PARÂMETRO DE DISTORÇÃO q	35
3.5	ROBUSTEZ DAS ESTIMATIVAS ML_q	35
3.5.1	Cenário 1: Grau de liberdade da ST conhecido	37
3.5.2	Cenário 2: Grau de liberdade da ST desconhecido	37
4	APLICAÇÕES NUMÉRICAS	47
4.1	ESTUDOS DE SIMULAÇÕES	47
4.1.1	Cenário 1: Caso 1	48
4.1.2	Cenário 1: Caso 2	54
4.1.3	Cenário 2: Caso 1	60
4.1.4	Cenário 2: Caso 2	71
4.2	APLICAÇÕES COM DADOS REAIS	83
4.2.1	Stack-Loss	83
4.2.1.1	Caso 1	84

4.2.1.2	Caso 2	86
4.2.2	Land Rent	88
4.2.3	 AIS	90
5	 CONCLUSÕES E TRABALHOS FUTUROS	94
5.1	CONCLUSÕES	94
5.2	TRABALHOS FUTUROS	94
	 REFERÊNCIAS	 95
	 APÊNDICE A – RESULTADOS ADICIONAIS DO CAPÍTULO	
	4	100

1 INTRODUÇÃO

1.1 MOTIVAÇÃO

Eventos raros possuem grande importância em diferentes domínios das ciências sociais ou naturais, conforme podemos ver em [38]. Como, por exemplo, as altas velocidades de vento e temperaturas extremas são de grande importância para a meteorologia; a ocorrência de ondas extremas, chuvas e inundações para a oceanografia e para a hidrologia. Do mesmo modo como a ocorrência de epidemias generalizadas e virulentas é de grande importância na epidemiologia; veja [32], [33], [60] e [59]. Podemos notar que apesar de serem raros, tais dados atípicos podem alterar substancialmente a inferência estatística quando retirados do conjunto de dados. Na literatura estatística tem havido portanto, grande interesse na detecção de tais observações com o intuito de avaliar o impacto das mesmas no modelo ajustado.

É bem conhecido que a inferência sob a distribuição normal é vulnerável a dados atípicos, para mais detalhes veja [64] e [40]. A distribuição t-Student fornece uma alternativa flexível e robusta para a distribuição normal. A distribuição t-Student possui o parâmetro graus de liberdade que é o responsável pela flexibilidade desta distribuição por controlar a curtose e a espessura da cauda, sendo por isso conhecido como parâmetro de forma. Quanto menor o valor atribuído aos graus de liberdade, temos caudas mais pesadas e podemos recuperar a distribuição normal no caso limite da distribuição t-Student, ao tender os graus de liberdade ao infinito. Além disso, a distribuição t-Student pertence à família das distribuições elípticas e é uma mistura de escala da distribuição normal; veja [23] e [40]. Porém, quando os dados demonstram um comportamento assimétrico, podemos fazer a modelagem dos mesmos através de membros da classe de distribuições assimétricas, também denominadas de skew-elípticas; veja [14].

Dessa forma, uma considerável pesquisa tem sido feita para introduzir famílias paramétricas flexíveis que acomodem desvios da suposição de normalidade/simetria. Nesta linha de pesquisa, destaca-se a classe de distribuições misturas de escala skew-normal, conforme discutida em [14], [36], [37] e [64], por exemplo. Esta classe de distribuições tem sido usada em diversas aplicações estatísticas para inferência robusta e compreende as distribuições misturas de escala normal (por exemplo, normal e t-Student), skew-normal, skew-t, entre outras.

Neste trabalho, desenvolvemos alguns resultados adicionais para a distribuição Skew-t (ST), introduzida por [14] e estudada com mais detalhes em [7]. As principais características relacionadas à família ST, tais como, assimetria e caudas pesadas, serão apresentadas no próximo capítulo. Aqui estaremos apenas evidenciando sua utilidade para

a modelagem de alguns problemas, tanto em sua forma multivariada quanto a univariada.

Em [11], podemos ver uma exploração numérica, em dados de várias origens e naturezas, enfatizando o potencial da distribuição ST como ferramenta da inferência robusta. Tais aspectos relacionados a robustez do processo de estimação nessa distribuição também foram discutidos em [9] e de forma mais extensiva em [10]. A importância da distribuição ST pode ser notada pelo grande número de artigos que vem sendo publicados em vários periódicos. Alguns exemplos de trabalhos são: [48] e [1] desenvolveram métodos para alocação de portfólios em um contexto financeiro, assumindo o modelo ST, onde as caudas pesadas e assimetria são características padrões neste tipo de problema. Além disso, [61] e [54] propuseram a distribuição ST para modelar produtos da indústria cinematográfica e musical como função de variáveis explicativas.

Para o modelo ST, a estimação via máxima verossimilhança (ML) com o uso do algoritmo EM (Esperança-Maximização), introduzido por [19], e suas extensões, é um dos procedimentos padrão. Adicionalmente, [25] introduziram, recentemente, o método de máxima verossimilhança L_q (ML_q), como uma generalização da estimação por ML. Este método de estimação por ML_q produz estimadores robustos para os parâmetros de interesse. O comportamento do estimador de ML_q é caracterizado pelo grau de distorção q aplicado ao modelo assumido. Mais informações sobre este método serão apresentadas no Capítulo 3.

Na literatura estatística existem poucos trabalhos aplicados sobre estimação por ML_q . Em particular, destaca-se o trabalho de [21], onde eles propõem a estimação por ML_q para todos os parâmetros da distribuição t-Student multivariada. Para o nosso conhecimento, até o momento, não há trabalhos sobre estimação por ML_q no contexto da distribuição ST univariada/multivariada. O objetivo principal deste trabalho é obter os estimadores de ML_q para todos os parâmetros da distribuição multivariada ST como alternativa aos estimadores de ML. Em nossa abordagem, o papel das observações é alterado com uma mudança no modelo original ao se adicionar um parâmetro de distorção q . Então, podemos considerar a estimação por ML_q como a minimização da discrepância de um modelo mais geral em relação aquele que diminui o papel das observações extremas gerando estimadores mais robustos que os de ML, em termos da redução dos pesos destas observações no processo de estimação.

1.2 OUTLIERS E ROBUSTEZ

"Um outlier é uma observação que se diferencia tanto das demais observações que levanta suspeitas de que aquela observação foi gerada por um mecanismo distinto"; veja [31]. Em outras palavras, os outliers são dados que se distanciam radicalmente de todos os outros. Os outliers ou valores atípicos podem ser causados por vários fatores, podem ser naturais ou artificiais, no caso de outliers naturais pode-se pensar em alguém que

declarou sua renda, e se saiu muito melhor, ou pior do que o restante de pessoas de seu perfil econômico. Os valores atípicos artificiais podem ser devido a uma leitura ou erro de anotações e transcrição incorreta de dados, bem como mudanças não controláveis nas condições experimentais. Após a detecção de um outlier deve-se excluí-lo ou conservá-lo, dependendo do real motivo para o seu aparecimento. Caso haja motivos concretos como um erro na observação ou na execução do experimento o certo é eliminá-lo, mas se não houver explicação para o aparecimento deste dado então ele pode refletir uma característica do que está sendo estudado, neste caso a observação feita deve ser incluída na análise e deve ser tratada especialmente. Outliers e robustez são termos que serão considerados com muita importância ao longo desta dissertação.

O termo outlier, muitas vezes, é usado informalmente e [44] ressalta que outliers são definidos com respeito aos modelos. Dessa forma, observações podem ser outliers em um modelo e serem perfeitamente regulares para outro modelo. É bem conhecido que, no contexto da família de distribuições elípticas, a inferência sobre a distribuição normal é vulnerável a dados atípicos. Como já citado anteriormente, os dados atípicos podem ser encontrados em vários ramos incluindo comércio e indústria, epidemiologia, entre outras áreas. Além da presença de dados atípicos, a distribuição dos dados nestes campos geralmente é assimétrica. A distribuição Skew-normal, por exemplo, é um modelo assimétrico que vem sendo utilizado com bastante frequência, como podemos ver em [6], [8] e [13]. Entretanto, esta distribuição possui estimativas vulneráveis a dados atípicos, estimulando o desenvolvimento de metodologias robustas em relação a estas observações.

O termo robustez está relacionada com a sensibilidade dos resultados de uma análise estatística aos desvios nas suposições de que esta análise se baseia; veja [50], páginas 382–413, para mais detalhes. Neste contexto, a distribuição ST se destaca. O uso do modelo ST pode resolver o problema de observações atípicas para as distribuições assimétricas e/ou caudas pesadas da mesma forma que [39] e [44] fizeram ao substituir a distribuição normal pela distribuição t-Student. Isto lhes permitiram obter estimativas robustas para os parâmetros de interesse, o que foi comprovado por um primeiro conceito de robustez, nomeado de robustez qualitativa; para mais detalhes, veja [50].

Neste trabalho, consideramos métodos robustos na presença de outliers, propondo um novo algoritmo EM, para obter os estimadores de ML para todos os parâmetros da ST, denotado por EM_q (algoritmo da Esperança-Maximização com a verossimilhança L_q), isto é, que aborda ML_q dentro da estrutura EM. Ao fazer isso, herdamos a robustez dos estimadores obtidos por ML_q e o tornamos disponível no contexto do modelo ST. Assim, obtemos estimadores duplamente reponderados para todos os parâmetros da distribuição ST.

1.3 OBJETIVOS

O objetivo do presente trabalho é obter os estimadores de ML_q para todos os parâmetros da distribuição ST multivariada como alternativa aos estimadores de ML.

Em relação ao cálculo dos estimadores de ML para os parâmetros do modelo ST multivariado, ao contrário da distribuição normal, as equações de verossimilhança não podem ser resolvidas analiticamente. Portanto, algoritmos numéricos devem ser usados para obter as estimativas. Devido à representação de mistura de escala da distribuição ST multivariada, o algoritmo EM, e suas variantes, são frequentemente usadas para encontrar as estimativas de ML. Da mesma forma, as equações de estimativa ML_q também não podem ser resolvidas analiticamente para obter fórmulas explícitas para os estimadores de ML_q , nos fazendo recorrer ao uso de algoritmos numéricos.

Portanto, neste trabalho, iremos adaptar o algoritmo EM para obter os estimadores de ML_q para todos os parâmetros do modelo ST, semelhante aos algoritmos de [27] e [21], no contexto da distribuição t-Student, e [3], no contexto de misturas finitas de distribuições normais. É importante ressaltar que esta dissertação contribui positivamente para o desenvolvimento das áreas Estatística/Matemática Aplicada, estendendo e complementando alguns resultados recentes encontrados na literatura, por exemplo, [21], dentre outros.

1.4 ORGANIZAÇÃO DO TRABALHO

No Capítulo 2, antes de iniciar os estudos envolvendo o método de estimação de máxima verossimilhança L_q para todos os parâmetros da distribuição ST, destacamos alguns conceitos preliminares que serão úteis para um maior entendimento dos capítulos seguintes. Neste capítulo, apresentamos as distribuições misturas de escala skew-normal introduzida por [14] e posteriormente estudada por [14], [36], [37] e [64], por exemplo, com a finalidade de descrever suas propriedades, tais como, representação estocástica, momentos, formas quadráticas e casos particulares, em especial a distribuição ST. Além disso, apresentamos uma breve descrição do algoritmo EM e discutimos com mais detalhes o algoritmo EM para obtenção dos estimadores de máxima verossimilhança no contexto do modelo ST.

No Capítulo 3, propomos os estimadores de máxima verossimilhança L_q , via algoritmo EM, tomando como referência os trabalhos de [25] e [21], dando ênfase no problema dos outliers. Neste capítulo, apresentamos uma breve descrição dos estimadores de máxima verossimilhança L_q e mostramos com mais detalhes os estimadores de máxima verossimilhança L_q via algoritmo EM, onde se baseia o objetivo deste trabalho. Além disso, discussões sobre a escolha do grau de distorção q aplicado ao modelo assumido e sobre a robustez das estimativas de ML_q são também apresentadas.

No Capítulo 4, exemplos numéricos considerando dados reais e simulados são apresentados para ilustrar a utilidade da metodologia proposta.

O Capítulo 5 finaliza esta dissertação com conclusões e diretrizes para trabalhos futuros.

Note que todas as siglas que serão utilizadas, neste trabalho, estarão em inglês para uma maior facilidade do leitor em uma busca bibliográfica.

2 CONCEITOS PRELIMINARES

Neste capítulo, apresentamos uma revisão dos principais conceitos que são tratados nesta dissertação. Portanto, este capítulo está organizado como segue. Na próxima seção, definimos as distribuições misturas de escala skew-normal e destacamos algumas propriedades, tais como momentos, representação estocástica, formas quadráticas e casos particulares, em especial a distribuição ST. Na Seção 2.2, descrevemos a distribuição ST, discutida por [7], nas formas padrão e locação-escala, destacando suas propriedades, momentos e representação estocástica. Em seguida, na Seção 2.3, descrevemos algumas propriedades da forma quadrática distância de Mahalanobis. Além disso, na Seção 2.4, apresentamos uma breve descrição do algoritmo EM e discutimos com mais detalhes este algoritmo para obtenção dos estimadores de máxima verossimilhança no contexto do modelo ST.

2.1 DISTRIBUIÇÕES MISTURA DE ESCALA SKEW-NORMAL (SMSN)

Apresentamos as distribuições Mistura de Escala Skew-Normal (SMSN) introduzida por [14], destacando algumas propriedades que são compartilhadas por todos elementos desta classe. Inicialmente, apresentamos uma breve introdução da distribuição Skew-normal usada neste trabalho; veja [6] e [5] para mais detalhes.

Definição 2.1. Seja \mathbf{W} um vetor aleatório p -dimensional com vetor de locação $\boldsymbol{\mu} \in \mathbb{R}^p$, matriz de dispersão $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$ positiva definida e $\boldsymbol{\lambda} \in \mathbb{R}^p$ o vetor de assimetria, onde $\mathbb{R}^{p \times p}$ é o conjunto das matrizes quadradas $p \times p$ com entradas reais. Dizemos que \mathbf{W} segue uma distribuição Skew-normal se sua função de densidade de probabilidade (pdf) é dada por

$$f(\mathbf{w}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}) = 2\phi_p(\mathbf{w}|\boldsymbol{\mu}, \boldsymbol{\Sigma})\Phi(\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{w} - \boldsymbol{\mu})), \quad (2.1)$$

em que $\phi_p(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ é a pdf da distribuição normal p -variada com vetor de médias $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$ e $\Phi(\cdot)$ é a função de distribuição acumulada (cdf) da normal padrão, tal que $\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}} = \boldsymbol{\Sigma}^{-1}$. Portanto, quando \mathbf{W} segue uma distribuição com fdp dada em (2.1), denotaremos que $\mathbf{W} \sim SN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda})$. Observe que ao fazer o vetor de assimetrias $\boldsymbol{\lambda} = \mathbf{0}$, em (2.1) obtemos a pdf da distribuição normal p -variada.

Considerando a construção da classe de distribuições mistura de escala normal introduzida por [4] e discutida posteriormente por [40], temos que um vetor aleatório com distribuição SMSN é construído a partir da mistura de escala de um vetor aleatório Skew-normal e uma variável aleatória positiva, ou seja,

$$\mathbf{Y} = \boldsymbol{\mu} + \kappa^{\frac{1}{2}}(U)\mathbf{Z}, \quad (2.2)$$

onde $\boldsymbol{\mu}$ é o vetor de locação, $\mathbf{Z} \sim SN_p(0, \boldsymbol{\Sigma}, \boldsymbol{\lambda})$, $\kappa(\cdot)$ é a função peso e U é uma variável aleatória de mistura positiva independente de \mathbf{Z} , com função densidade de probabilidade $h(u|\boldsymbol{\nu})$ e função de distribuição acumulada $H(u|\boldsymbol{\nu})$, que estão indexadas por um vetor aleatório ou um escalar $\boldsymbol{\nu}$ que controla as caudas da distribuição, podendo ser conhecido (fixo) ou não. De (2.2), temos que dado $U = u$, \mathbf{Y} seguirá uma distribuição $SN_p(\boldsymbol{\mu}, \kappa(u)\boldsymbol{\Sigma}, \boldsymbol{\lambda})$. Deste modo, integrando em relação a u , a fdp de \mathbf{Y} será

$$f(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = 2 \int_0^\infty \phi_p(\mathbf{y}|\boldsymbol{\mu}, \kappa(u)\boldsymbol{\Sigma}) \Phi(\kappa^{-\frac{1}{2}}(u)\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{y} - \boldsymbol{\mu})) dH(u|\boldsymbol{\nu}), \mathbf{y} \in \mathbb{R}^p. \quad (2.3)$$

Denotaremos por $\mathbf{Y} \sim SMSN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}; H(u|\boldsymbol{\nu}))$, um vetor aleatório com pdf dada em (2.3). Note que ao assumir $\boldsymbol{\lambda} = \mathbf{0}$, estaremos trabalhando com a classe simétrica de distribuições mistura de escala normal, $SMN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; H(u|\boldsymbol{\nu}))$; veja [4] e [40] para mais informações.

De acordo com [36], um vetor aleatório \mathbf{Y} com fdp dada em (2.3) possui a seguinte representação estocástica:

$$\mathbf{Y} \stackrel{d}{=} \boldsymbol{\mu} + \kappa^{\frac{1}{2}}(U)\boldsymbol{\Sigma}^{\frac{1}{2}} \left[\delta |T_0| + (\mathbf{I}_p - \delta \delta^\top) T_1 \right], \text{ com } \delta = \frac{\boldsymbol{\lambda}}{\sqrt{1 + \boldsymbol{\lambda}^\top \boldsymbol{\lambda}}}, \quad (2.4)$$

em que o símbolo $\stackrel{d}{=}$ significa "equivalência em distribuição" e \mathbf{I}_p é a matriz identidade $p \times p$, $|T_0|$ denota o valor absoluto de $T_0 \sim N(0, 1)$, $T_1 \sim N_p(0, \mathbf{I}_p)$ e $U \sim H(\cdot|\boldsymbol{\nu})$ são variáveis independentes. Esta representação estocástica além de facilitar a implementação do algoritmo EM, pode ser usada para simular dados e também permite derivar muitas propriedades da distribuição de \mathbf{Y} , como, por exemplo, da forma quadrática distância de Mahalanobis $d_\lambda = (\mathbf{Y} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \boldsymbol{\mu})$, útil para detectar observações atípicas; veja a Seção 2.3, deste trabalho, para mais detalhes.

Além disso, segundo [37], apresentamos alguns resultados relacionados aos momentos condicionais de funções de U dado \mathbf{y} , denotados por u_r e η_r , essenciais para a implementação do algoritmo EM e encontram-se resumidos na proposição que segue.

Proposição 2.2. *Seja $\mathbf{Y} \sim SMSN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}; H(u|\boldsymbol{\nu}))$ e $U \sim H(u|\boldsymbol{\nu})$ a variável de mistura. Então*

$$u_r = E[\kappa^{-r}(U)|\mathbf{y}] = 2 \frac{f_0(\mathbf{y})}{f(\mathbf{y})} E[\kappa^{-r}(U_y) \Phi(\kappa^{-\frac{1}{2}}(U_y)\mathbf{A})] \quad (2.5)$$

$$\eta_r = E[\kappa^{-\frac{r}{2}}(U) W_\Phi(\kappa^{-\frac{1}{2}}(U)\mathbf{A})|\mathbf{y}] = 2 \frac{f_0(\mathbf{y})}{f(\mathbf{y})} E[\kappa^{-\frac{r}{2}}(U_y) \phi_1(\kappa^{-\frac{r}{2}}(U_y)\mathbf{A})] \quad (2.6)$$

onde $\mathbf{A} = \boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{y} - \boldsymbol{\mu})$, $W_\Phi(\cdot) = \frac{\phi_1(\cdot)}{\Phi(\cdot)}$, f_0 é a fdp de $Y_0 \sim SMN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; H(u|\boldsymbol{\nu}))$ e $U_y \stackrel{d}{=} U|\mathbf{Y}_0 = \mathbf{y}$.

Demonstração. Ver prova em [64]. □

Esta classe de distribuições contém como casos particulares as distribuições Skew-normal, Skew-t, Skew-slash, Skew-normal contaminada, entre outras distribuições que podem ser obtidas a partir da mudança de distribuição da variável aleatória positiva U . A seguir apresentamos a distribuição ST, onde se baseia o nosso objetivo.

2.2 DISTRIBUIÇÃO SKEW-T (ST)

A construção de famílias paramétricas de distribuições assimétricas que sejam analiticamente tratáveis e que possam acomodar valores práticos de assimetria e curtose, incluindo a distribuição Skew-normal (normal e t-Student, no contexto simétrico) como caso particular, pode ser útil para a modelagem de dados, análises estatísticas e estudos de métodos robustos. A distribuição ST, desde sua introdução à literatura, vem recebendo muita atenção, seja de modo teórico ao se estudar suas propriedades ou em aplicações numéricas como modelo para o ajuste de dados. Pode-se dizer que um dos grande motivos para gerar tanto interesse se deve ao seu elevado grau de flexibilidade, uma vez que seus parâmetros abrangem uma faixa admissível, com uma ampla variação das medidas associadas a assimetria e a curtose.

Do mesmo modo que em [12], para um melhor entendimento, inicialmente, introduzimos a distribuição ST padrão univariada, isto é, considerando o parâmetro de locação μ nulo e seu parâmetro de escala $\sigma^2 = 1$. Neste caso, a fdp é dada por

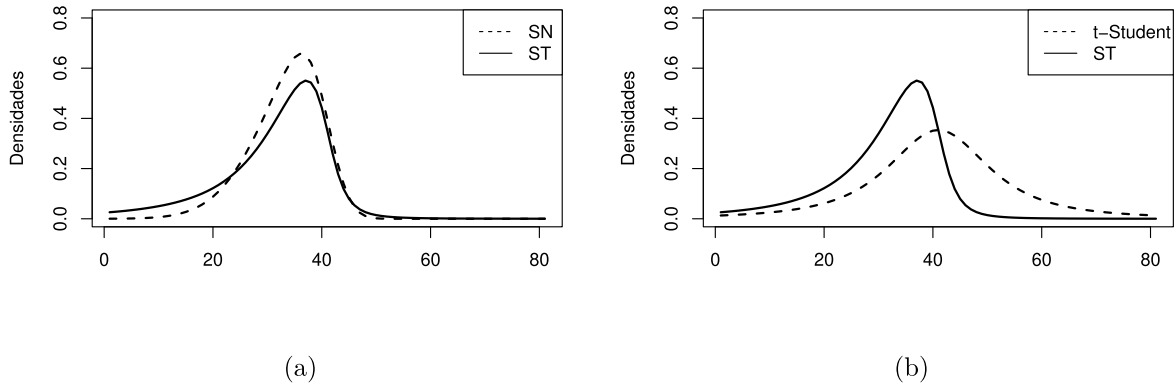
$$f(z|0, 1, \lambda, \nu) = 2t(z|\nu)T\left(z\lambda\sqrt{\frac{\nu+1}{\nu+z^2}}|\nu+1\right), \quad z \in \mathbb{R}, \quad (2.7)$$

em que $t(\cdot|\nu)$ e $T(\cdot|\nu+1)$ são, respectivamente, a fdp e a fda da distribuição t-Student padrão univariada com ν graus de liberdade.

Segundo [12], se Z é uma variável aleatória com função de densidade dada em (2.7), então $Y = \mu + \sigma Z$ segue uma distribuição $ST(\mu, \sigma^2, \lambda, \nu)$. Um caso particular da distribuição ST é a distribuição Skew-cauchy quando $\nu = 1$. Se $\lambda = 0$, obtemos a distribuição t-Student com os parâmetros μ, σ^2 e ν . Quando $\nu \rightarrow \infty$, obtemos a distribuição Skew-normal com os parâmetros μ, σ^2 e λ . Além disso, quando $\nu \rightarrow \infty$ e $\lambda = 0$, obtemos a distribuição Normal com os parâmetros μ e σ^2 .

Na Figura 2.1 (a), podemos ver as densidades da distribuição SN(0,1,-3) e ST(0,1,-3,2). Note que ambas distribuições são assimétricas negativas e a distribuição ST têm caudas mais pesadas que a distribuição SN. Na Figura 2.1 (b), podemos ver as densidades da distribuição ST(0,1,-3,2) e t-Student univariada com os parâmetros $\mu = 0, \sigma^2 = 1$ e $\nu = 2$. Podemos concluir, sem perda de generalidade, que a distribuição ST é obtida como uma perturbação da distribuição t-Student clássica, sendo λ , o parâmetro de forma, representando o desvio da simetria.

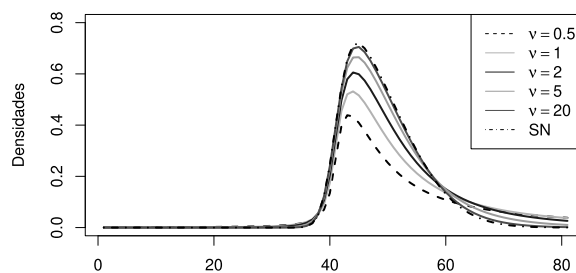
Figura 2.1 – Dados simulados - (a) Densidades da SN(0,1,-3) e ST(0,1,-3,2). (b) Densidades da ST(0,1,-3,2) e t-Student univariada com os parâmetros $\mu = 0$, $\sigma^2 = 1$ e $\nu = 2$.



Fonte: Elaborado pela autora em 2020.

Na Figura 2.2, podemos observar o comportamento da distribuição ST a medida que aumentamos o valor de ν , o que corrobora que a medida que ν tende ao infinito a distribuição ST tende a uma distribuição SN.

Figura 2.2 – Dados simulados - Densidades da SN e ST padrões univariadas, ambas com assimetria $\lambda = 5$, com os graus de liberdade ν variando em uma faixa de valores.



Fonte: Elaborado pela autora em 2020.

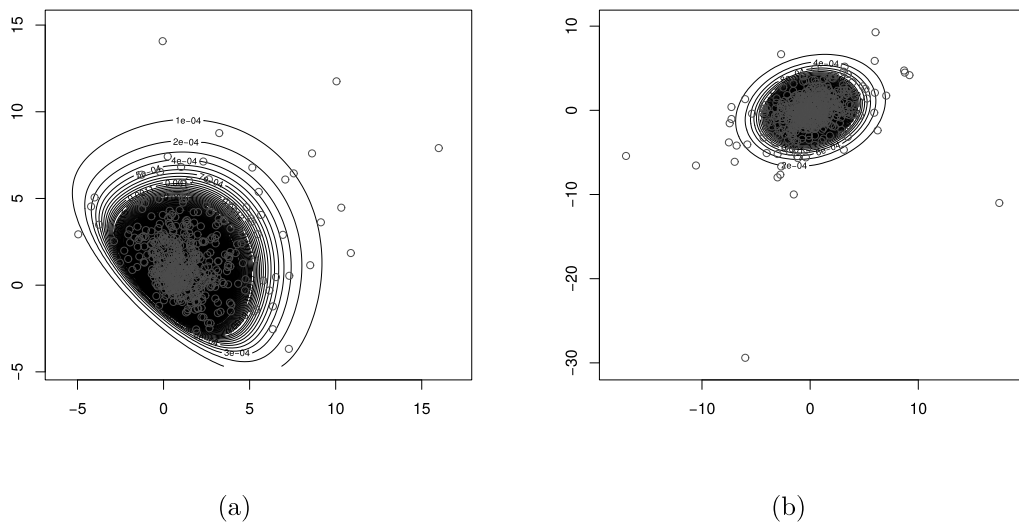
A distribuição ST multivariada com ν graus de liberdade, apresentada em [7] e [12], denotada por $ST_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \nu)$, pode ser obtida de (2.3) ao se considerar $\kappa(u) = 1/u$, com $U \sim \text{Gamma}(\nu/2, \nu/2)$, $u > 0$ e $\nu > 0$. Dessa forma, a fdp de \mathbf{Y} é dada por

$$f(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \nu) = 2t_p(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)T\left(\sqrt{\frac{p+\nu}{d_\lambda+\nu}}\boldsymbol{\lambda}^\top\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{y}-\boldsymbol{\mu})|\nu+p\right), \quad \mathbf{y} \in \mathbb{R}^p, \quad (2.8)$$

em que d_λ é a distancia de Mahalanobis, $t_p(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ é a fdp da distribuição t-Student p variada e $T(\cdot|\nu)$ é a fda da distribuição t-Student padrão univariada. Nas Figuras 2.3

(a) e (b), podemos ver os contornos das densidades da ST e t-Student com $\boldsymbol{\mu} = (0, 0)^\top$, $\boldsymbol{\Sigma} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$, $\boldsymbol{\lambda} = (4, 4)^\top$ e $\nu = 4$, para fins ilustrativos. Note que os contornos da densidade da ST (Figura 2.3 (a)) não são elípticos e podem ser bastante assimétricos dependendo da escolha do parâmetro $\boldsymbol{\lambda}$.

Figura 2.3 – Dados simulados - (a) e (b) Contornos das densidades da ST e t-Student bivariadas, respectivamente.



Fonte: Elaborado pela autora em 2020.

O vetor de médias e a matriz de covariâncias de $\mathbf{Y} \sim ST_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \nu)$ são dados por

$$E[\mathbf{Y}] = \boldsymbol{\mu} + \sqrt{\frac{\nu}{\pi}} \frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \boldsymbol{\Delta}, \text{ com } \nu > 1, \quad (2.9)$$

$$Var[\mathbf{Y}] = \frac{\nu}{\nu-2} \boldsymbol{\Sigma} - \frac{\nu}{\pi} \left(\frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \right)^2 \boldsymbol{\Delta} \boldsymbol{\Delta}^\top, \text{ com } \nu > 2, \quad (2.10)$$

em que $\Gamma(\cdot)$ é a função Gamma e $\boldsymbol{\Delta} = \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{\delta}$. Além disso, segue da Proposição 2.2 que $\mathbf{Y}_0 \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, ou seja, $Y_0|U = u \sim N_p(\boldsymbol{\mu}, u^{-1}\boldsymbol{\Sigma})$, com $U \sim \text{Gamma}(\nu/2, \nu/2)$. Assim, usando o fato de que $U_y \stackrel{d}{=} U|\mathbf{Y}_0 = \mathbf{y} \sim \text{Gamma}((\nu+p)/2, (\nu+d_\lambda)/2)$, após algumas manipulações algébricas, temos os seguintes momentos condicionais u_r e η_r , dadas por

$$u_r = \frac{f_0(\mathbf{y})}{f(\mathbf{y})} \frac{2^{r+1} \Gamma\left(\frac{\nu+p+2r}{2}\right) (\nu+d_\lambda)^{-r}}{\Gamma\left(\frac{\nu+p}{2}\right)} T\left(\sqrt{\frac{\nu+p+2r}{\nu+d_\lambda}} A | \nu+p+2r\right), \quad (2.11)$$

$$\eta_r = \frac{f_0(\mathbf{y})}{f(\mathbf{y})} \frac{2^{\frac{r+1}{2}} \Gamma\left(\frac{\nu+p+r}{2}\right)}{\pi^{\frac{1}{2}} \Gamma\left(\frac{\nu+p}{2}\right)} \frac{(\nu+d_\lambda)^{\frac{\nu+p}{2}}}{(\nu+d_\lambda + A^2)^{\frac{\nu+p+r}{2}}}, \quad (2.12)$$

na qual $f_0(y)$ é a função de densidade de probabilidade de Y_0 e $A = \boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{y} - \boldsymbol{\mu})$. Para mais detalhes, veja [14], [40], [7] e [64], por exemplo.

Portanto, note que a distribuição ST possui propriedades interessantes e importantes, como o fato de admitir uma representação estocástica que facilita a implementação do algoritmo EM, permite combinar assimetria com as caudas pesadas, possui como casos particulares as distribuições Skew-Cauchy, Skew-normal, t-Student e normal, além de atribuir pesos diferentes para cada observação e conseqüentemente controlar a influência da observação no processo de estimação, por exemplo, como veremos na Seção 2.4.1.1. Aplicações da distribuição ST em procedimentos de estimação robustos também podem ser encontradas em [11] e [64]. Nesta linha de pesquisa, na próxima seção, discutimos sobre a forma quadrática distância de Mahalanobis, útil para identificar observações atípicas, candidatas às observações influentes. Os elementos do conjunto de dados que efetivamente controlam aspectos da análise são ditos influentes. Em particular, uma observação, é dita influente se ela produzir alterações relevantes no resultado de análise quando for excluída, ou submetida a uma pequena perturbação; veja [17] para maiores informações. Numa análise estatística é importante levar em consideração a estabilidade dos resultados obtidos e das inferências de um modelo geral.

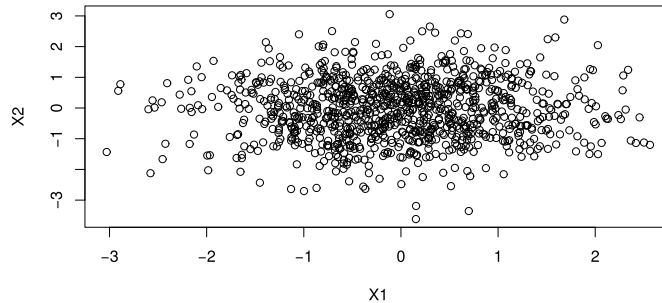
2.3 DISTÂNCIA DE MAHALANOBIS

Segundo [42], para detectar outliers, a partir de um pensamento matemático ou estatístico, devemos levar em consideração a forma/estrutura do conjunto de dados. Por exemplo, imagine um conjunto de dados no \mathbb{R}^2 , tendo uma forma elíptica proveniente de uma normal bivariada padrão. Na Figura 2.4, podemos observar que os dados possuem um comportamento elíptico, onde alguns pontos se aproximam mais do centro, enquanto outros estão mais afastados. Porém, não podemos tirar conclusões precipitadas em relação aos pontos mais distantes, em termos da distância Euclidiana, concluindo que pertencem menos a amostra que os pontos mais próximos, pois isso faz parte do padrão subjacente da distribuição normal. Isso ocorre porque a distância Euclidiana é uma distância entre dois pontos apenas. Ela não considera como o resto dos pontos no conjunto de dados variam. Portanto, não pode ser usada para julgar o quão próximo um ponto está de uma distribuição de pontos.

Sendo assim, ao invés de considerarmos a distância clássica, devemos considerar uma distância que leve em consideração a forma das observações sob uma análise mais minuciosa. Tal distância é a distância de Mahalanobis, introduzida na década de 30 por [45], que estaremos denotando, em sua forma quadrática, da seguinte maneira:

$$d = (\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{X} - \boldsymbol{\mu}), \quad (2.13)$$

Figura 2.4 – Gráfico de dispersão de duas variáveis X_1 e X_2 provenientes de uma $N(0,1)$, assumindo correlação $\rho = 0,5$.



Fonte: Elaborado pela autora em 2020.

em que \mathbf{X} é um vetor de variáveis (X_1, \dots, X_p) , $\boldsymbol{\mu}$ é um vetor de dimensão p e $\boldsymbol{\Sigma}$ é uma matriz simétrica de dimensão $p \times p$. A distância de Mahalanobis mede a distância de um ponto \mathbf{X} ao centro $\boldsymbol{\mu}$ na métrica $\boldsymbol{\Sigma}$, o que implica que ela depende da estrutura de $\boldsymbol{\Sigma}$. Na prática, os valores de $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$ são desconhecidos e portanto, precisam ser estimados. Neste trabalho, a obtenção do estimador de ML dos parâmetros de interesse será baseada no algoritmo EM. Dessa forma, na próxima seção, apresentamos uma breve descrição deste algoritmo e algumas de suas extensões, inclusive os detalhes do algoritmo EM no contexto do modelo ST.

De acordo com [14], [40]; e [64], se $\mathbf{Y} \sim ST_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \nu)$, a forma quadrática $d_\lambda = (\mathbf{Y} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-\frac{1}{2}} (\mathbf{Y} - \boldsymbol{\mu})$ tem a mesma distribuição de $d = (\mathbf{Y}_0 - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-\frac{1}{2}} (\mathbf{Y}_0 - \boldsymbol{\mu})$, onde $\mathbf{Y}_0 \sim t\text{-Student}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$. Dessa forma, podemos concluir de [40] que $d_\lambda \sim pF(p, \nu)$. Ao substituir as estimativas de ML de $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$ em d_λ , podemos, através de gráficos da distância de Mahalanobis e considerando o quantil II como marca de referência da distribuição da forma quadrática d_λ , identificar outliers; veja [53], por exemplo.

2.4 ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA (ML)

O método de estimação por máxima verossimilhança (ML) é uma das técnicas mais populares para encontrar os estimadores para os parâmetros do modelo de interesse. No modelo ST, quando o grau de liberdade é assumido como fixo (conhecido) e considerado como um valor pequeno, os estimadores de ML para os parâmetros de locação, de escala e de assimetria são muito populares em análises estatísticas robustas, uma vez que fornecem estimadores robustos alternativos aos estimadores clássicos obtidos a partir da distribuição skew-normal (ou normal). Esses estimadores são considerados estimadores M com a função objetivo $-\log f$. De forma geral, um estimador M é o zero de uma função que esta sendo estimada; veja [29] e [28] para mais detalhes.

Definição 2.3. As variáveis aleatórias Y_1, \dots, Y_n são chamadas de amostra aleatória de tamanho n da população Y , se Y_1, \dots, Y_n forem variáveis aleatórias mutuamente independentes e a função de densidade de probabilidade (pdf) ou a função de probabilidade (pf) marginal de cada Y_i for a mesma função $f(y|\boldsymbol{\theta})$. De modo alternativo Y_1, \dots, Y_n são chamadas variáveis aleatórias *independentes e identicamente distribuídas* com pdf ou pf $f(y|\boldsymbol{\theta})$, comumente abreviado como variáveis aleatórias i.i.d.

Definição 2.4. Seja $\mathbf{Y} = (Y_1, \dots, Y_n)$ um vetor cujas coordenadas constituem uma amostra aleatória de uma população com pdf (ou pf) f dependente do vetor de parâmetros $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p) \in \Theta$. A função de verossimilhança é definida por

$$L(\boldsymbol{\theta} | \mathbf{y}) = L(\theta_1, \dots, \theta_p | y_1, \dots, y_n) = \prod_{i=1}^n f(y_i | \theta_1, \dots, \theta_p).$$

Se a função de verossimilhança é diferenciável em θ_i , os possíveis candidatos para estimadores são os valores que resolvem

$$\frac{\partial L(\boldsymbol{\theta} | \mathbf{y})}{\partial \theta_i} = 0. \quad (2.14)$$

Observe que a solução da equação (2.14) nos fornece apenas possíveis candidatos uma vez que ao obter a primeira derivada nula é somente condição necessária para máximo e não suficiente; veja [16], página 282, para mais detalhes.

Como estaremos trabalhando com a diferenciabilidade da função de verossimilhança, podemos trabalhar com a *log-verossimilhança* que é definida da seguinte forma

$$\ell(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta} | \mathbf{y}) = \log \prod_{i=1}^n f(y_i | \boldsymbol{\theta}) = \sum_{i=1}^n \log f(y_i | \boldsymbol{\theta}). \quad (2.15)$$

Uma vez que a função logarítmica apresenta um comportamento crescente no intervalo $(0, \infty)$, temos que o máximo de $\ell(\boldsymbol{\theta})$ coincidirá com o máximo de $L(\boldsymbol{\theta})$, ou seja, a estimativa de máxima verossimilhança $\hat{\boldsymbol{\theta}}$ é o valor de $\boldsymbol{\theta} \in \Theta$ que maximiza $\ell(\boldsymbol{\theta})$, isto é,

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta} \in \Theta} \ell(\boldsymbol{\theta}). \quad (2.16)$$

Note que não existem soluções explícitas para o problema de maximização da função de log-verossimilhança do modelo ST, onde a fdp está definida em (2.8) e $\boldsymbol{\theta} = (\boldsymbol{\mu}^\top, \boldsymbol{\alpha}^\top, \boldsymbol{\lambda}^\top, \nu)^\top$, com $\boldsymbol{\alpha}$ sendo o conjunto mínimo de parâmetros da matriz $\boldsymbol{\Sigma}$, de modo que $\boldsymbol{\Sigma}$ esteja bem definida. Neste caso, podemos maximizar numericamente usando, por exemplo, o Matlab, ver [46], e o R, ver [58]. Estes softwares contêm rotinas prontas para tratar problemas de maximização (minimização) de qualquer função. Uma vez que a função de log-verossimilhança envolve expressões complexas, é muito difícil maximizar diretamente $\ell(\boldsymbol{\theta})$ para estimativa de ML dos parâmetros do modelo ST. Para superar este

problema, neste trabalho, usamos o Algoritmo EM - do inglês Expectation-Maximization - para o cálculo dos estimadores de máxima verossimilhança de modo a obtermos soluções analíticas para os estimadores dos parâmetros de locação, de escala e de assimetria do modelo ST.

Os estimadores de ML, por serem da classe dos estimadores M, como podemos ver em [47], é considerado um bom método de estimação por causa de sua robustez, sob certas condições. Por exemplo, em [44], encontramos uma discussão sobre a robustez dos estimadores de ML dos parâmetros de locação e de escala do modelo t-Student, quando o parâmetro ν é considerado fixo (conhecido). A robustez local dos estimadores M pode ser discutida investigando suas funções de influência e os estimadores M com funções de influência limitada são considerados como os estimadores com boas propriedades de robustez local; veja [29] e [28] para mais detalhes.

Os estimadores de máxima verossimilhança apresentam, sob certas condições de regularidade (ver [16], página 516), uma série de propriedades assintóticas, como a consistência (conforme a Definição 10.1.1 encontrada em [16], página 468), que são responsáveis por justificar seu uso frente a outros estimadores. Veja [57] e [52], por exemplo, para mais detalhes sobre tais propriedades assintóticas.

Em alguns casos, a consistência é verificada computacionalmente (i.e., de forma empírica) por meio de medidas de viés e desvios para os conjuntos de dados simulados com diferentes tamanhos amostrais como faremos no Capítulo 4. A seguir listamos tais medidas usadas, onde N é o número de simulações realizadas para um determinado vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_p)$, o índice i se refere ao parâmetro de interesse e o índice j se refere à amostra.

1. Medidas de Viés

- Viés Mediano: $MB_i = \underset{i \leq j \leq N}{med}(\hat{\boldsymbol{\theta}}_{ij} - \boldsymbol{\theta}_i)$, para cada $i = 1, \dots, p$.
- Viés Médio: $Bias_i = \frac{\sum_{j=1}^N (\hat{\boldsymbol{\theta}}_{ij} - \boldsymbol{\theta}_i)}{N}$, para cada $i = 1, \dots, p$.

2. Desvio Puro

- Erro Quadrático Médio: $MSE_i = \sqrt{\frac{\sum_{j=1}^N (\hat{\boldsymbol{\theta}}_{ij} - \boldsymbol{\theta}_i)^2}{N}}$, para cada $i = 1, \dots, p$.

Na próxima seção, apresentamos uma breve descrição do algoritmo EM e discutimos com mais detalhes este algoritmo para obtenção dos estimadores de máxima verossimilhança no contexto do modelo ST.

2.4.1 Algoritmo EM

O algoritmo EM foi desenvolvido na década de 1970 como uma alternativa aos métodos tradicionais de otimização, usualmente considerados na estimação por ML, particularmente aqueles dos tipos Newton ou Quasi-Newton. [19] destacou-se, neste contexto, com a publicação do primeiro artigo completo sobre o tema.

De um modo geral, o algoritmo EM se baseia no tratamento dos dados observados como “incompletos”. Notadamente casos em que realmente faltam observações (missing, em inglês) podem ter a estimação por ML feita a partir do algoritmo EM, conforme discutido em [52] e [16]. No entanto, esse algoritmo vem sendo amplamente usado em casos nos quais a função log-verossimilhança é considerada complicada demais para os algoritmos convencionais. Dessa forma, tem como base a ideia de substituir uma difícil maximização da log-verossimilhança por uma seqüência de maximizações mais fáceis, cujo limite é a resposta para o problema original.

Muitos problemas em Estatística podem ser considerados utilizando uma formulação de dados aumentados permitindo assim simplificar a obtenção de estimativas de ML. Os dados aumentados, também chamados dados completos, correspondem aos dados observados, em conjunto com os dados perdidos ou não observáveis. Neste contexto, as funções de verossimilhança, baseadas nos dados completos e observados, são denominadas verossimilhança de dados completos e dados incompletos, respectivamente. É importante salientar que a parte aumentada dos dados, referente aos dados não observados, não requer que eles sejam “perdidos” no sentido estrito da palavra, pois somente representam um mecanismo técnico.

Considere que \mathbf{y}_{obs} denota os dados observados e \mathbf{y}_{mis} denota os dados não observáveis, sendo os dados completos $\mathbf{y}_c = (\mathbf{y}_{obs}, \mathbf{y}_{mis})$. Denotamos por $f(\mathbf{y}_c|\boldsymbol{\theta})$ a função de verossimilhança dos dados completos e $\ell_c(\boldsymbol{\theta}|\mathbf{y}_c) = \log(f(\mathbf{y}_c|\boldsymbol{\theta}))$ a função de log-verossimilhança dos dados completos. O algoritmo EM aborda problemas com dados incompletos indiretamente mediante a substituição da parte não observável em $\ell_c(\boldsymbol{\theta}|\mathbf{y}_c)$ por suas esperanças condicionadas a \mathbf{y}_{obs} , usando o ajuste atual para $\boldsymbol{\theta}$. Dessa forma, definimos $Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})$ (função-Q) como a esperança da função de log-verossimilhança dos dados completos, ou seja, $Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}) = E\{\ell_c(\boldsymbol{\theta}|\mathbf{y}_c)|\mathbf{y}_{obs}, \hat{\boldsymbol{\theta}}\}$, $\boldsymbol{\theta} \in \Theta$.

Cada iteração do algoritmo EM consiste de dois passos: esperança (passo E) e maximização (passo M). Definimos $\hat{\boldsymbol{\theta}}^{(k)}$ como a estimativa do vetor $\boldsymbol{\theta}$ na k -ésima iteração e assim, a $(k + 1)$ -ésima iteração do algoritmo EM é definida como

Passo E: Para $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}^{(k)}$, calcular $Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(k)})$ como

$$Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(k)}) = E\{\ell_c(\boldsymbol{\theta}|\mathbf{y}_c)|\mathbf{y}_{obs}, \hat{\boldsymbol{\theta}}^{(k)}\};$$

Passo M: Obter $\widehat{\boldsymbol{\theta}}^{(k+1)}$ que maximize $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)})$, tal que

$$Q(\widehat{\boldsymbol{\theta}}^{(k+1)}|\widehat{\boldsymbol{\theta}}^{(k)}) > Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}), \quad \forall \boldsymbol{\theta} \in \Theta.$$

Deve-se alternar os passos E e M repetidamente até atingir a convergência. Cada iteração do algoritmo EM incrementa o logaritmo da função de verossimilhança observada $\ell(\widehat{\boldsymbol{\theta}}|\mathbf{y}_{obs})$ de modo que $\ell(\widehat{\boldsymbol{\theta}}^{(k)}|\mathbf{y}_{obs}) \leq \ell(\widehat{\boldsymbol{\theta}}^{(k+1)}|\mathbf{y}_{obs})$ e o algoritmo tipicamente converge. Como critério de convergência, usamos $\left| \frac{\ell(\widehat{\boldsymbol{\theta}}^{(k+1)}|\mathbf{y}_{obs})}{\ell(\widehat{\boldsymbol{\theta}}^{(k)}|\mathbf{y}_{obs})} - 1 \right| < \varepsilon$, onde $\varepsilon = 10^{-4}$.

Em [49] e [43] foram desenvolvidas extensões do algoritmo EM denominadas, respectivamente, por ECM e ECME. Ambas variantes do algoritmo EM particionam o vetor $\boldsymbol{\theta}$ em dois subvetores $\boldsymbol{\theta}_1 \in \Theta_1$ e $\boldsymbol{\theta}_2 \in \Theta_2$, onde $\Theta = \Theta_1 \cup \Theta_2$, e substituem o passo M da seguinte forma:

$$\begin{aligned} \bullet \text{ Algoritmo ECM} & \left\{ \begin{array}{l} \text{Passo CM-1: Escolha } \widehat{\boldsymbol{\theta}}_1^{(k+1)} \in \arg \max_{\boldsymbol{\theta}_1 \in \Theta_1} Q^{(k)}(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(k)}); \\ \text{Passo CM-2: Escolha } \widehat{\boldsymbol{\theta}}_2^{(k+1)} \in \arg \max_{\boldsymbol{\theta}_2 \in \Theta_2} Q^{(k)}(\widehat{\boldsymbol{\theta}}_1^{(k+1)}, \boldsymbol{\theta}_2). \end{array} \right. \\ \bullet \text{ Algoritmo ECME} & \left\{ \begin{array}{l} \text{Passo CM: Escolha } \widehat{\boldsymbol{\theta}}_1^{(k+1)} \in \arg \max_{\boldsymbol{\theta}_1 \in \Theta_1} Q^{(k)}(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2^{(k)}); \\ \text{Passo CME: Escolha } \widehat{\boldsymbol{\theta}}_2^{(k+1)} \in \arg \max_{\boldsymbol{\theta}_2 \in \Theta_2} \ell^{(k)}(\widehat{\boldsymbol{\theta}}_1^{(k+1)}, \boldsymbol{\theta}_2). \end{array} \right. \end{aligned}$$

É interessante comentar que no programa R, há os pacotes *mixsmsn* e *sn*, que permitem a obtenção das estimativas de máxima verossimilhança dos parâmetros da distribuição ST, sendo que o pacote *mixsmsn*, as obtém via algoritmo EM. Na próxima seção, apresentamos com maiores detalhes o algoritmo EM para obtenção dos estimadores de máxima verossimilhança no contexto do modelo ST.

2.4.1.1 Algoritmo EM no modelo ST

Antes de explanarmos sobre algoritmo EM para a distribuição ST, definimos o vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\mu}^\top, \boldsymbol{\alpha}^\top, \boldsymbol{\lambda}^\top, \nu)^\top$, onde $\boldsymbol{\alpha}$ é o conjunto mínimo de parâmetros da matriz $\boldsymbol{\Sigma}$, de modo que $\boldsymbol{\Sigma}$ esteja bem definida. Para explorar o algoritmo EM, apresentamos o modelo ST no contexto de dados incompletos, usando os resultados descritos na Seção 2.4.1. Dessa forma, de (2.4), o modelo ST, descrito na Seção 2.2, pode ser reescrito

hierarquicamente como

$$\mathbf{Y}|T = t, U = u \stackrel{ind}{\sim} N_p(\boldsymbol{\mu} + \boldsymbol{\Delta}t, u^{-1}\boldsymbol{\Gamma}), \quad (2.17)$$

$$T|U_i = u_i \stackrel{ind}{\sim} HN_1(0, u^{-1}), \quad (2.18)$$

$$U \stackrel{iid}{\sim} \text{Gamma}(\nu/2, \nu/2), \quad (2.19)$$

onde $\boldsymbol{\Delta} = \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\delta}$, $\boldsymbol{\Gamma} = \boldsymbol{\Sigma} - \boldsymbol{\Delta}\boldsymbol{\Delta}^\top$ e $HN_1(0, \sigma^2)$ denota a distribuição half- $N(0, \sigma^2)$; veja [34]. Considere $\mathbf{y} = (\mathbf{y}_1^\top, \dots, \mathbf{y}_n^\top)^\top$, o vetor de respostas observáveis para as n unidades amostrais, $\mathbf{u} = (u_1, \dots, u_n)^\top$ e $\mathbf{t} = (t_1, \dots, t_n)^\top$. Então, sob a representação hierárquica (2.17)-(2.19), temos que a função de log-verossimilhança dos dados completos, $\mathbf{y}_c = (\mathbf{y}^\top, \mathbf{u}^\top, \mathbf{t}^\top)^\top$, é definida como segue

$$\begin{aligned} \ell_c(\boldsymbol{\theta}|\mathbf{y}_c) = & c - \frac{n}{2} \log |\boldsymbol{\Gamma}| - \frac{1}{2} \sum_{i=1}^n u_i [\nu + (\mathbf{y}_i - \boldsymbol{\mu} - \boldsymbol{\Delta}t_i)^\top \boldsymbol{\Gamma}^{-1} (\mathbf{y}_i - \boldsymbol{\mu} - \boldsymbol{\Delta}t_i)] \\ & + \frac{n\nu}{2} \log \left(\frac{\nu}{2} \right) - n \log \left(\Gamma \left(\frac{\nu}{2} \right) \right) + \left(\frac{\nu + p}{2} - 1 \right) \sum_{i=1}^n \log(u_i), \end{aligned} \quad (2.20)$$

onde c é uma constante que independe do vetor de parâmetros $\boldsymbol{\theta}$. Partindo da equação (2.20), podemos definir a função $Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(k)})$, tal que

$$\begin{aligned} Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}^{(k)}) = & c - \frac{n}{2} \log |\boldsymbol{\Gamma}| - \frac{1}{2} \sum_{i=1}^n \widehat{u}_i^{(k)} (\mathbf{y}_i - \boldsymbol{\mu})^\top \boldsymbol{\Gamma}^{-1} (\mathbf{y}_i - \boldsymbol{\mu}) + \sum_{i=1}^n \widehat{ut}_i^{(k)} (\mathbf{y}_i - \boldsymbol{\mu})^\top \boldsymbol{\Gamma}^{-1} \boldsymbol{\Delta} \\ & - \frac{1}{2} \sum_{i=1}^n \widehat{ut}_i^{2(k)} \boldsymbol{\Delta}^\top \boldsymbol{\Gamma}^{-1} \boldsymbol{\Delta} - \frac{1}{2} \sum_{i=1}^n \widehat{u}_i^{(k)} \nu + \frac{n\nu}{2} \log \left(\frac{\nu}{2} \right) - n \log \left(\Gamma \left(\frac{\nu}{2} \right) \right) \\ & + \left(\frac{\nu + p}{2} - 1 \right) \sum_{i=1}^n \widehat{\log}(u_i)^{(k)}, \end{aligned} \quad (2.21)$$

onde $\widehat{u}_i = E[U_i|\hat{\boldsymbol{\theta}}^{(k)}, \mathbf{y}_i]$, $\widehat{ut}_i = E[U_i T_i|\hat{\boldsymbol{\theta}}^{(k)}, \mathbf{y}_i]$, $\widehat{ut}_i^2 = E[U_i T_i^2|\hat{\boldsymbol{\theta}}^{(k)}, \mathbf{y}_i]$ e $\widehat{\log}(u_i) = E[\log U_i|\hat{\boldsymbol{\theta}}^{(k)}, \mathbf{y}_i]$. Usando as propriedades conhecidas da esperança condicional, as quantidades \widehat{ut}_i e \widehat{ut}_i^2 são obtidas a partir das equações (2.11)-(2.12). Para maiores detalhes, veja [36].

[35] apresenta duas propostas para a estimação dos graus de liberdade da distribuição ST. Na primeira proposta de estimação, [35] considera um valor fixo para ν . Porém, no contexto de dados reais, assume que ν estará em uma malha de valores a serem avaliados de forma a escolher o valor de ν que maximize a log-verossimilhança dos dados observados. Já a segunda proposta de estimação está relacionada ao passo M, onde particionamos o vetor de parâmetros $\boldsymbol{\theta}$ da seguinte forma: $\boldsymbol{\theta}_1 = (\boldsymbol{\mu}, \boldsymbol{\alpha}, \boldsymbol{\lambda})$ e $\boldsymbol{\theta}_2 = \nu$, decidindo qual variação do algoritmo EM, **ECM** ou **ECME**, ambas apresentadas na Seção 2.4.1, podemos usar. Neste trabalho, para o modelo ST, usamos o algoritmo **ECME**.

Passos do algoritmo ECME para a distribuição ST

1. Forneça o chute inicial $\hat{\boldsymbol{\theta}}^{(0)} = (\hat{\boldsymbol{\mu}}^{(0)}, \hat{\boldsymbol{\alpha}}^{(0)}, \hat{\boldsymbol{\lambda}}^{(0)}, \hat{\nu}^{(0)})$ e o valor de ε para o critério de parada.

2. Passo E: Dado $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}^{(k)}$, calcule

$$\widehat{u}_i^{(k)} = \frac{\widehat{\nu}^{(k)} + p}{\widehat{\nu}^{(k)} + \widehat{d}_\lambda^{(k)}}, \quad \widehat{ut}_i^{(k)} = \widehat{u}_i^{(k)} \widehat{\mu}_{T_i}^{(k)} + \widehat{M}_T^{(k)} \widehat{\eta}_{1i}^{(k)},$$

$$\widehat{ut}_i^2{}^{(k)} = \widehat{u}_i^{(k)} \widehat{\mu}_{T_i}^2{}^{(k)} + \widehat{M}_T^2{}^{(k)} + \widehat{\mu}_{T_i}^{(k)} \widehat{M}_T^{(k)} \widehat{\eta}_{1i}^{(k)},$$

$$\log(\widehat{u}_i)^{(k)} = \psi\left(\frac{\widehat{\nu}^{(k)} + p}{2}\right) + \log\left(\frac{\widehat{\nu}^{(k)} + \widehat{d}_\lambda^{(k)}}{2}\right),$$

onde $\widehat{M}_T^2 = [1 + \widehat{\boldsymbol{\Delta}}^\top \widehat{\boldsymbol{\Gamma}}^{-1} \widehat{\boldsymbol{\Delta}}]^{-1}$, $\widehat{\mu}_T = \widehat{M}_T^2 \widehat{\boldsymbol{\Delta}}^\top \widehat{\boldsymbol{\Gamma}}^{-1} (\mathbf{y} - \widehat{\boldsymbol{\mu}})$, \widehat{d}_λ é a distância de Mahalanobis avaliada nas estimativas de máxima verossimilhança de $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$, $\psi(\cdot)$ é a função Digamma e $\widehat{\eta}_{1i}$ pode ser facilmente obtida de (2.12).

3. Passo-CM: Obtenha $\widehat{\boldsymbol{\mu}}^{(k+1)}$, $\widehat{\boldsymbol{\Sigma}}^{(k+1)}$ e $\widehat{\boldsymbol{\lambda}}^{(k+1)}$ ao maximizar $Q(\boldsymbol{\theta} | \widehat{\boldsymbol{\theta}}^{(k)})$ em relação aos parâmetros $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ e $\boldsymbol{\lambda}$, respectivamente, e considerando $\nu = \widehat{\nu}^{(k)}$. Essa maximização resulta nas seguintes expressões fechadas

$$\widehat{\boldsymbol{\mu}}^{(k+1)} = \frac{\sum_{i=1}^n \widehat{u}_i^{(k)} \mathbf{y}_i - \sum_{i=1}^n \widehat{ut}_i^{(k)} \widehat{\boldsymbol{\Delta}}^{(k)}}{\sum_{i=1}^n \widehat{u}_i^{(k)}},$$

$$\widehat{\boldsymbol{\Gamma}}^{(k+1)} = \frac{1}{n} \sum_{i=1}^n \left[\widehat{u}_i^{(k)} (\mathbf{y}_i - \widehat{\boldsymbol{\mu}}^{(k)})^\top (\mathbf{y}_i - \widehat{\boldsymbol{\mu}}^{(k)}) - \widehat{ut}_i^{(k)} (\mathbf{y}_i - \widehat{\boldsymbol{\mu}}^{(k)})^\top \widehat{\boldsymbol{\Delta}}^{(k)} - \widehat{ut}_i^{(k)} \widehat{\boldsymbol{\Delta}}^{(k)\top} (\mathbf{y}_i - \widehat{\boldsymbol{\mu}}^{(k)}) + \widehat{ut}_i^2{}^{(k)} \widehat{\boldsymbol{\Delta}}^{(k)\top} \widehat{\boldsymbol{\Delta}}^{(k)} \right],$$

e

$$\widehat{\boldsymbol{\Delta}}^{(k+1)} = \frac{\sum_{i=1}^n \widehat{ut}_i^{(k)} (\mathbf{y}_i - \widehat{\boldsymbol{\mu}}^{(k)})}{\sum_{i=1}^n \widehat{ut}_i^2{}^{(k)}}.$$

Assim, pela propriedade de invariância dos estimadores de ML, obtemos

$$\widehat{\boldsymbol{\Sigma}}^{(k+1)} = \widehat{\boldsymbol{\Gamma}}^{(k+1)} - \widehat{\boldsymbol{\Delta}}^{(k+1)} \widehat{\boldsymbol{\Delta}}^{(k+1)\top}$$

e

$$\widehat{\boldsymbol{\lambda}}^{(k+1)} = \frac{(\widehat{\boldsymbol{\Sigma}}^{(k+1)})^{-1/2} \widehat{\boldsymbol{\Delta}}^{(k+1)}}{(1 - \widehat{\boldsymbol{\Delta}}^{(k+1)\top} (\widehat{\boldsymbol{\Sigma}}^{(k+1)})^{-1} \widehat{\boldsymbol{\Delta}}^{(k+1)})^{1/2}}.$$

3. Passo-CME: Atualize $\widehat{\nu}^{(k+1)}$, tal que $\widehat{\nu}^{(k+1)} = \arg \max_{\nu} \ell^{(k)}(\widehat{\boldsymbol{\theta}}_1^{(k+1)}, \nu)$, onde $\ell(\cdot)$ é função de log-verossimilhança dos dados observados, obtida a partir de (2.8).

4. Repita os passos anteriores até o critério de parada, definido na Seção 2.4.1 ser alcançado.

Os valores iniciais são necessários para a implementação do algoritmo EM. Então, propomos como chute inicial, $\hat{\boldsymbol{\theta}}^{(0)} = (\hat{\boldsymbol{\mu}}^{(0)}, \hat{\boldsymbol{\alpha}}^{(0)}, \hat{\boldsymbol{\lambda}}^{(0)}, \hat{\nu}^{(0)})$, a média amostral para $\boldsymbol{\mu}$, a matriz de covariâncias amostrais para $\boldsymbol{\Sigma}$, o sinal da assimetria amostral para $\boldsymbol{\lambda}$ e $\hat{\nu}^{(0)} = 1$ (indicativo de caudas pesadas). Ou seja,

$$\hat{\boldsymbol{\mu}}^{(0)} = \frac{1}{n} \sum_{i=0}^n \mathbf{y}_i, \quad \hat{\boldsymbol{\Sigma}}^{(0)} = \sqrt{\frac{1}{n-1} \sum_{i=0}^n (\mathbf{y}_i - \boldsymbol{\mu}^{(0)})^2}, \quad (2.22)$$

e

$$\hat{\boldsymbol{\lambda}}^{(0)} = \text{sign} \left(\frac{\frac{1}{n} \sum_{i=0}^n (\mathbf{y}_i - \boldsymbol{\mu}^{(0)})^3}{\left[\frac{1}{n-1} \sum_{i=0}^n (\mathbf{y}_i - \boldsymbol{\mu}^{(0)})^2 \right]^{\frac{3}{2}}} \right). \quad (2.23)$$

Uma outra proposta sobre chute inicial no contexto da estimação por máxima verossimilhança dos parâmetros da distribuição ST pode ser encontrada em [12].

Podemos chamar o algoritmo EM de algoritmo de reponderação iterativa, uma vez que seus pesos $\hat{u}_i, i = 1, \dots, n$, obtidos na etapa E , são inversamente proporcionais a distância de Mahalanobis, isto é, este processo de estimação tende a atribuir menor peso às observações atípicas de acordo com a distância de Mahalanobis. No próximo capítulo, propomos o algoritmo EM_q , que podemos chamar de algoritmo de dupla reponderação iterativa, pois irá fornecer estimadores duplamente reponderados para os todos os parâmetros da distribuição ST.

Em [28], podemos ver que a robustez local dos estimadores M pode ser investigada ao se discutir sua função influência, isto é, os estimadores M com funções de influência limitada são considerados como os estimadores com boas propriedades de robustez local. Conforme apresentado em [28], a função influência dos estimadores do tipo M é uma transformação linear do vetor escore (vetor das primeiras derivadas da log-verossimilhança dos dados observados). Portanto, para os estimadores de ML, o comportamento da função de influência é determinado pelas funções de escore de cada parâmetro. Se as funções de escore são limitadas, a função influência para o estimador de ML correspondente será limitada. Em [44], podemos ver essa relação para o caso da distribuição t-Student, as funções de escore para os parâmetros de locação e escala são limitadas quando o parâmetro de graus de liberdade é conhecido (fixo). Assim, a função de influência correspondente para os estimadores de ML é limitada e, portanto, os estimadores ML para os parâmetros de locação e escala são localmente robustos, quando ν é conhecido. Entretanto, em alguns casos é desejável estimar o parâmetro de graus de liberdade junto com os demais parâmetros. Neste contexto, [44] ao fazer a estimação simultânea dos graus de liberdade com os outros parâmetros da distribuição t-Student, observou que diferentemente dos outros parâmetros, a função de escore para ν é ilimitada. Assim, os estimadores de ML não são mais localmente robustos devido ao caráter ilimitado da função de escore relacionada ao parâmetro de graus de liberdade. Portanto, em uma análise estatística para obter

estimadores robustos, o parâmetro de graus de liberdade é geralmente assumido como conhecido; veja [44], [39], [21] e [51], por exemplo.

Motivados por estes trabalhos, no próximo capítulo, discutiremos sobre os estimadores ML_q para todos os parâmetros do modelo ST.

3 ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA L_q VIA ALGORITMO EM: UMA ESTIMAÇÃO ROBUSTA DO MODELO SKEW-T

3.1 INTRODUÇÃO

Neste capítulo, estudamos alguns aspectos de estimação na classe de distribuições misturas de escala skew-normal, especificamente no modelo Skew-t. A distribuição ST possui propriedades interessantes e importantes, como o fato de admitir uma representação estocástica que facilita a implementação do algoritmo EM e suas extensões. A estimação por máxima verossimilhança é um dos procedimentos mais usados devido às boas propriedades desses estimadores, tais como invariância e consistência. Neste contexto, a estimação de todos os parâmetros do modelo ST, inclusive para os graus de liberdade, será pelo método de estimação de máxima verossimilhança L_q (ML_q) introduzido por [25]. Este procedimento se baseia na função q -entropia, que será definida na próxima seção, como uma generalização da estimação por ML, que pode acarretar a robustez dos estimadores, sob determinado modelo de interesse. Em sua abordagem, somos introduzidos à um parâmetro de distorção q , tal que $0 < q < 1$, que modifica ligeiramente o papel das observações no modelo durante o processo de estimação. Sendo assim, o método de estimação por máxima verossimilhança L_q pode ser considerado como um procedimento que modifica a distribuição do modelo para diminuir o papel das observações extremas. Se $q = 1$, todas as observações recebem o mesmo peso.

Mais recentemente, uma série de aplicações baseadas na q -entropia apareceram nas áreas de biomedicina, finanças, ciências ambientais; veja [26] e [24], por exemplo. Conforme [25], apesar do amplo sucesso, até o momento poucos esforços foram feitos para abordar as implicações inferenciais do uso de entropias de uma perspectiva estatística. [21] considerou o método de máxima verossimilhança L_q na distribuição t-Student multivariada. Eles modificaram o algoritmo EM para obter os estimadores de ML_q para todos os parâmetros do modelo de interesse e discutiram a robustez deste método de estimação sob a distribuição t-Student multivariada. Seguindo essa linha de pesquisa, neste capítulo, consideramos uma estimação robusta para todos os parâmetros do modelo ST. Dessa forma, propomos os estimadores de máxima verossimilhança L_q , via algoritmo EM modificado, tomando como base o trabalho de [21], dando ênfase no problema dos outliers.

Este capítulo está organizado como segue. Na seção 3.2 são definidos os estimadores de máxima verossimilhança L_q . Em seguida na Seção 3.3, apresentamos o algoritmo EM adaptado para o cálculo dos estimadores de ML_q para todos os parâmetros da distribuição ST multivariada. Na Seção 3.4, discutimos a escolha do parâmetro de distorção q . E concluindo, na Seção 3.5 abordamos aspectos sobre a robustez dos estimadores ML_q no modelo ST quando ν é desconhecido e portanto, é estimado.

3.2 ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA L_q

Seja Y uma variável aleatória com fdp $f(y|\boldsymbol{\theta})$, com o parâmetro $\boldsymbol{\theta} \in \Theta^m$ em \mathbb{R}^m , para $m \geq 1$. Então, a entropia de Shannon, mais conhecida no contexto da teoria da informação, para uma variável aleatória Y é definida como $H(Y) = -E[\log(f(y|\boldsymbol{\theta}))]$. Conforme [2] apontou, temos que maximizar a função de log-verossimilhança de um modelo paramétrico é equivalente à minimizar a versão empírica da entropia de Shannon, definida como $-\sum_{i=1}^n \log(f(y_i|\boldsymbol{\theta}))$.

O conceito de q -entropia, que iremos usar, foi inicialmente proposto por [30] e [60], definido da seguinte forma

$$H_q(Y) = -E[L_q(f(Y|\boldsymbol{\theta}))], \quad (3.1)$$

na qual L_q é a função q -logarítmica dada por

$$L_q(u) = \begin{cases} \log(u), & u \geq 0, q = 1, \\ \frac{u^{1-q} - 1}{1 - q}, & u \geq 0, q \neq 1. \end{cases} \quad (3.2)$$

No contexto estatístico, a função L_q representa uma transformação Box-Cox e observe que a q -entropia pode ser considerada uma generalização da entropia de Shannon. Quando $q = 1$ em (3.1), obtemos a entropia de Shannon.

Usando a versão empírica da q -entropia, dada nas equações (3.1)-(3.2), [25] introduziram o método de estimação por ML_q . Considere $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ uma amostra aleatória proveniente de $f(\mathbf{y}|\boldsymbol{\theta})$. O estimador de ML_q para $\boldsymbol{\theta}$ é dada por

$$\widehat{\boldsymbol{\theta}}_q = \arg \max_{\boldsymbol{\theta}} \sum_{i=1}^n L_q(f(\mathbf{y}_i|\boldsymbol{\theta})), \quad (3.3)$$

com $0 < q < 1$. Note que quando $q = 1$, obtemos os estimadores de ML usuais. Além disso, após algumas manipulações algébricas, [25] mostram que o estimador de ML_q para $\boldsymbol{\theta}$ é solução da seguinte equação de estimação

$$\sum_{i=1}^n \mathbf{U}_q(\mathbf{y}_i|\boldsymbol{\theta}) = \sum_{i=1}^n \mathbf{U}(\mathbf{y}_i|\boldsymbol{\theta}) f(\mathbf{y}_i|\boldsymbol{\theta})^{1-q} = 0, \quad (3.4)$$

onde $\mathbf{U}(\mathbf{y}_i|\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \log(f(\mathbf{y}_i|\boldsymbol{\theta}))$ é o vetor escore. Podemos observar que a equação (3.4) oferece uma interpretação natural do ML_q , como uma solução de uma verossimilhança ponderada. Se $q = 1$, todas as observações recebem o mesmo peso. Porém, quando $q \neq 1$, o ML_q fornece uma reponderação relativa ao modelo e assim, as observações atípicas

recebem menores pesos durante o processo de estimação, como veremos na próxima seção no contexto do modelo ST.

3.3 ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA L_q VIA ALGORITMO EM

Seja $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ uma amostra aleatória proveniente da distribuição ST multivariada com fdp $f(\mathbf{y}|\theta)$ definida em (2.8), onde $\theta = (\boldsymbol{\mu}, \boldsymbol{\alpha}, \boldsymbol{\lambda}, \nu)^\top$, com $\boldsymbol{\alpha}$ sendo o conjunto mínimo de parâmetros da matriz $\boldsymbol{\Sigma}$, de modo que esteja bem definida. Seguindo [25], os estimadores ML_q para os parâmetros da distribuição ST podem ser obtidos como solução da equação de estimação dada em (3.4). Observe que, semelhante aos estimadores de ML, discutidos na Seção 2.4, não existem soluções explícitas para o problema de maximização da função

$$\ell_q(\theta) = \sum_{i=1}^n L_q(f(\mathbf{y}_i|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \nu)). \quad (3.5)$$

Assim, as estimativas de ML_q também podem ser obtidas usando métodos de maximização numérica que podem ser instáveis dependendo da complexidade do modelo. Uma vez que a função (3.5) envolve expressões complexas, é muito difícil maximizar diretamente para obtenção das estimativas de ML_q dos parâmetros do modelo ST.

Dessa forma, um algoritmo do tipo EM, semelhante aos algoritmos propostos por [56], [27], [21] e [63], será proposto para obter as estimativas de ML_q para os parâmetros da distribuição ST. A descrição do algoritmo EM proposto neste trabalho, denominado EM_q , será dada a seguir.

3.3.1 Algoritmo EM_q no modelo ST

Seja $(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \hat{\boldsymbol{\lambda}}, \hat{\nu})$ as estimativas atuais dos parâmetros e definimos $W_{qi}(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \hat{\boldsymbol{\lambda}}, \hat{\nu}) = f(\mathbf{y}_i|\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \hat{\boldsymbol{\lambda}}, \hat{\nu})^{1-q}$. Assim, as novas estimativas ML_q serão obtidas ao maximizar a seguinte função

$$\sum_{i=1}^n W_{qi}(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \hat{\boldsymbol{\lambda}}, \hat{\nu}) \log(f(\mathbf{y}_i|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \nu)), \quad (3.6)$$

com fdp $f(\mathbf{y}|\theta)$, definida em (2.8), da distribuição ST. Para implementar o algoritmo EM_q , consideramos a log-verossimilhança dos dados completos dada em (2.20). No contexto dos dados completos $\mathbf{y}_c = (\mathbf{y}^\top, \mathbf{u}^\top, \mathbf{t}^\top)^\top$, discutido na Seção 2.4.1.1, a equação de estimação (3.4) pode ser reescrita como

$$\sum_{i=1}^n W_{qi}(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \hat{\boldsymbol{\lambda}}, \hat{\nu}) U(\mathbf{y}_i, u_i, t_i|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \nu) = 0, \quad (3.7)$$

onde $U(\mathbf{y}_i, u_i, t_i|\theta) = \frac{\partial}{\partial \theta} \log(f(\mathbf{y}_i, u_i, t_i|\theta))$ é a derivada da log-verossimilhança dos dados completos com respeito aos parâmetros. Então, escrevendo as derivadas de (3.7), obtemos

as seguintes equações para os parâmetros $\boldsymbol{\mu}$, $\boldsymbol{\Gamma}$ e $\boldsymbol{\Delta}$, respectivamente,

$$\sum_{i=1}^n W_{qi}(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}, \widehat{\boldsymbol{\lambda}}, \widehat{\nu}) [u_i \boldsymbol{\Gamma}^{-1}(\mathbf{y}_i - \boldsymbol{\mu}) - u_i t_i \boldsymbol{\Gamma}^{-1} \boldsymbol{\Delta}] = 0, \quad (3.8)$$

$$\sum_{i=1}^n W_{qi}(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}, \widehat{\boldsymbol{\lambda}}, \widehat{\nu}) \left(-\frac{n \boldsymbol{\Gamma}^{-1}}{2} + \frac{1}{2} u_i (\mathbf{y}_i - \boldsymbol{\mu})^\top \boldsymbol{\Gamma}^{-2} (\mathbf{y}_i - \boldsymbol{\mu}) - u_i t_i (\mathbf{y}_i - \boldsymbol{\mu})^\top \boldsymbol{\Gamma}^{-2} \boldsymbol{\Delta} + \frac{1}{2} u_i t_i^2 \boldsymbol{\Delta}^\top \boldsymbol{\Gamma}^{-2} \boldsymbol{\Delta} \right) = 0 \quad (3.9)$$

e

$$\sum_{i=1}^n W_{qi}(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}, \widehat{\boldsymbol{\lambda}}, \widehat{\nu}) (-u_i t_i^2 \boldsymbol{\Gamma}^{-1} \boldsymbol{\Delta} + u_i t_i (\mathbf{y}_i - \boldsymbol{\mu})^\top \boldsymbol{\Gamma}^{-1}) = 0, \quad (3.10)$$

com $\boldsymbol{\Delta} = \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{\delta}$, $\boldsymbol{\Gamma} = \boldsymbol{\Sigma} - \boldsymbol{\Delta} \boldsymbol{\Delta}^\top$ e $\boldsymbol{\delta} = \frac{\boldsymbol{\lambda}}{\sqrt{1 + \boldsymbol{\lambda}^\top \boldsymbol{\lambda}}}$. Para solucionar o efeito das variáveis latentes U_i e T_i nessas equações, temos que considerar as esperanças condicionais dessas equações dado y_i . Então, obtemos

$$\sum_{i=1}^n W_{qi}(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}, \widehat{\boldsymbol{\lambda}}, \widehat{\nu}) [\widehat{u}_{qi} \boldsymbol{\Gamma}^{-1}(\mathbf{y}_i - \boldsymbol{\mu}) - \widehat{u}_{qi} t_i \boldsymbol{\Gamma}^{-1} \boldsymbol{\Delta}] = 0, \quad (3.11)$$

$$\sum_{i=1}^n W_{qi}(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}, \widehat{\boldsymbol{\lambda}}, \widehat{\nu}) \left(-\frac{n \boldsymbol{\Gamma}^{-1}}{2} + \frac{1}{2} \widehat{u}_{qi} (\mathbf{y}_i - \boldsymbol{\mu})^\top \boldsymbol{\Gamma}^{-2} (\mathbf{y}_i - \boldsymbol{\mu}) - \widehat{u}_{qi} t_i (\mathbf{y}_i - \boldsymbol{\mu})^\top \boldsymbol{\Gamma}^{-2} \boldsymbol{\Delta} + \frac{1}{2} \widehat{u}_{qi} t_i^2 \boldsymbol{\Delta}^\top \boldsymbol{\Gamma}^{-2} \boldsymbol{\Delta} \right) = 0 \quad (3.12)$$

e

$$\sum_{i=1}^n W_{qi}(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}, \widehat{\boldsymbol{\lambda}}, \widehat{\nu}) (-\widehat{u}_{qi} t_i^2 \boldsymbol{\Gamma}^{-1} \boldsymbol{\Delta} + \widehat{u}_{qi} t_i (\mathbf{y}_i - \boldsymbol{\mu})^\top \boldsymbol{\Gamma}^{-1}) = 0, \quad (3.13)$$

$\widehat{u}_{qi} = E[U_i | \widehat{\boldsymbol{\theta}}_q, \mathbf{y}_i]$, $\widehat{u}_{qi} t_i = E[U_i T_i | \widehat{\boldsymbol{\theta}}_q, \mathbf{y}_i]$ e $\widehat{u}_{qi} t_i^2 = E[U_i T_i^2 | \widehat{\boldsymbol{\theta}}_q, \mathbf{y}_i]$. Como foi comentado na Seção 2.4.1.1, usando as propriedades conhecidas da esperança condicional, as quantidades \widehat{u}_{qi} e $\widehat{u}_{qi} t_i^2$ são obtidas a partir das equações (2.11)-(2.12). Para maiores detalhes, veja [36]. A seguir, apresentamos as etapas do algoritmo EM_q proposto.

Passos do algoritmo EM_q para a distribuição ST

1. Dado $0 < q < 1$ fixo, forneça o chute inicial $\widehat{\boldsymbol{\theta}}_q^{(0)} = (\widehat{\boldsymbol{\mu}}_q^{(0)}, \widehat{\boldsymbol{\alpha}}_q^{(0)}, \widehat{\boldsymbol{\lambda}}_q^{(0)}, \widehat{\nu}_q^{(0)})$ e o valor de ε para o critério de parada.
2. **Passo E:** Dado $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}_q^{(k)}$, calcule

$$\widehat{u}_{qi}^{(k)} = \frac{\widehat{\nu}_q^{(k)} + p}{\widehat{\nu}_q^{(k)} + \widehat{d}_{q\lambda}^{(k)}}, \quad \widehat{u}_{qi} t_i^{(k)} = \widehat{u}_{qi}^{(k)} \widehat{\mu}_{Tqi}^{(k)} + \widehat{M}_{Tq}^{(k)} \widehat{\eta}_{1qi}^{(k)},$$

$$\widehat{u}_{qi} t_i^2^{(k)} = \widehat{u}_{qi}^{(k)} \widehat{\mu}_{Tqi}^2^{(k)} + \widehat{M}_{Tq}^2^{(k)} + \widehat{\mu}_{Tqi}^{(k)} \widehat{M}_{Tq}^{(k)} \widehat{\eta}_{1qi}^{(k)},$$

onde $\widehat{M}_{Tq}^2 = [1 + \widehat{\boldsymbol{\Delta}}_q^\top \widehat{\boldsymbol{\Gamma}}_q^{-1} \widehat{\boldsymbol{\Delta}}_q]^{-1}$, $\widehat{\mu}_{Tq} = \widehat{M}_{Tq}^2 \widehat{\boldsymbol{\Delta}}_q^\top \widehat{\boldsymbol{\Gamma}}_q^{-1} (\mathbf{y} - \widehat{\boldsymbol{\mu}}_q)$, $\widehat{d}_{q\lambda}$ é a distância de Mahalanobis avaliada nas estimativas de ML_q de $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$, $\psi(\cdot)$ é a função Digamma e $\widehat{\eta}_{1qi}$ pode ser facilmente obtida de (2.12).

3. Passo-CM: Obtenha $\widehat{\boldsymbol{\mu}}_q^{(k+1)}$, $\widehat{\boldsymbol{\Sigma}}_q^{(k+1)}$ e $\widehat{\boldsymbol{\lambda}}_q^{(k+1)}$ ao resolver as equações (3.11)-(3.13) em relação aos parâmetros $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ e $\boldsymbol{\lambda}$, respectivamente, e considerando $\nu = \widehat{\nu}_q^{(k)}$. Essa maximização resulta nas seguintes expressões fechadas

$$\widehat{\boldsymbol{\mu}}_q^{(k+1)} = \frac{\sum_{i=1}^n W_{qi}^{(k)} \widehat{u}_{qi}^{(k)} \mathbf{y}_i - \sum_{i=1}^n W_{qi}^{(k)} \widehat{ut}_{qi}^{(k)} \widehat{\boldsymbol{\Delta}}_q^{(k)}}{\sum_{i=1}^n W_{qi}^{(k)} \widehat{u}_{qi}^{(k)}},$$

$$\begin{aligned} \widehat{\boldsymbol{\Gamma}}_q^{(k+1)} = & \frac{1}{n \sum_{i=1}^n W_{qi}^{(k)}} \sum_{i=1}^n \left[W_{qi}^{(k)} \widehat{u}_{qi}^{(k)} (\mathbf{y}_i - \widehat{\boldsymbol{\mu}}_q^{(k)})^\top (\mathbf{y}_i - \widehat{\boldsymbol{\mu}}_q^{(k)}) \right. \\ & - W_{qi}^{(k)} \widehat{ut}_{qi}^{(k)} (\mathbf{y}_i - \widehat{\boldsymbol{\mu}}_q^{(k)})^\top \widehat{\boldsymbol{\Delta}}_q^{(k)} - W_{qi}^{(k)} \widehat{ut}_{qi}^{(k)} \widehat{\boldsymbol{\Delta}}_q^{\top(k)} (\mathbf{y}_i - \widehat{\boldsymbol{\mu}}_q^{(k)}) \\ & \left. + W_{qi}^{(k)} \widehat{ut}_{qi}^2 \widehat{\boldsymbol{\Delta}}_q^{\top(k)} \widehat{\boldsymbol{\Delta}}_q^{(k)} \right], \end{aligned}$$

e

$$\widehat{\boldsymbol{\Delta}}_q^{(k+1)} = \frac{\sum_{i=1}^n W_{qi}^{(k)} \widehat{ut}_{qi}^{(k)} (\mathbf{y}_i - \widehat{\boldsymbol{\mu}}_q^{(k)})}{\sum_{i=1}^n W_{qi}^{(k)} \widehat{ut}_{qi}^2},$$

onde $W_{qi}^{(k)} = W_{qi}^{(k)}(\widehat{\boldsymbol{\mu}}_q^{(k)}, \widehat{\boldsymbol{\Sigma}}_q^{(k)}, \widehat{\boldsymbol{\lambda}}_q^{(k)}, \widehat{\nu}_q^{(k)})$. Assim, pela propriedade de invariância dos estimadores de máxima verossimilhança, obtemos

$$\widehat{\boldsymbol{\Sigma}}_q^{(k+1)} = \widehat{\boldsymbol{\Gamma}}_q^{(k+1)} - \widehat{\boldsymbol{\Delta}}_q^{(k+1)} \widehat{\boldsymbol{\Delta}}_q^{\top(k+1)}$$

e

$$\widehat{\boldsymbol{\lambda}}_q^{(k+1)} = \frac{(\widehat{\boldsymbol{\Sigma}}_q^{(k+1)})^{-1/2} \widehat{\boldsymbol{\Delta}}_q^{(k+1)}}{(1 - \widehat{\boldsymbol{\Delta}}_q^{\top(k+1)} (\widehat{\boldsymbol{\Sigma}}_q^{(k+1)})^{-1} \widehat{\boldsymbol{\Delta}}_q^{(k+1)})^{1/2}}.$$

3. Passo-CME: Atualize $\widehat{\nu}_q^{(k+1)}$, tal que

$$\widehat{\nu}_q^{(k+1)} = \arg \max_{\nu} \sum_{i=1}^n W_{qi}^{(k)} (\widehat{\boldsymbol{\mu}}_q^{(k)}, \widehat{\boldsymbol{\Sigma}}_q^{(k)}, \widehat{\boldsymbol{\lambda}}_q^{(k)}, \widehat{\nu}_q^{(k)}) \log(f(\mathbf{y}_i | \widehat{\boldsymbol{\mu}}_q^{(k)}, \widehat{\boldsymbol{\Sigma}}_q^{(k)}, \widehat{\boldsymbol{\lambda}}_q^{(k)}, \nu)), \quad (3.14)$$

com fdp $f(\mathbf{y}|\theta)$, definida em (2.8), da distribuição ST.

4. Repita os passos anteriores até o critério de parada, definido na Seção 2.4.1 ser alcançado.

Os valores iniciais são necessários para a implementação do algoritmo EM_q e usamos a mesma abordagem apresentada na Seção 2.4.1.1. Entretanto, com a finalidade de verificar se a estimativa de ML_q foi encontrada, recomenda-se rodar o algoritmo EM_q usando uma amplitude de diferentes valores iniciais. Note que quando $\boldsymbol{\lambda} = \mathbf{0}$ (ou $\boldsymbol{\Delta} = \mathbf{0}$), as equações do passo M se reduzem às equações obtidas quando assumimos a distribuição t-Student multivariada. Assim, este algoritmo proposto claramente generaliza os resultados encontrados em [21].

É importante ressaltar que ao considerar $W_q(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \hat{\boldsymbol{\lambda}}, \hat{\nu})$ nas equações acima, podemos chamar os estimadores resultantes de estimadores duplamente ponderados. [56] comentam que, quando $0 < q < 1$, o estimador de ML_q tende a colocar pesos mais baixos em observações com baixa verossimilhança e pesos mais altos em observações com maior verossimilhança. Concluindo, o método de estimação por máxima verossimilhança L_q pode ser considerado como um procedimento que modifica a distribuição do modelo para diminuir o papel das observações extremas. Dessa forma, seguindo esta ideia, na próxima seção discutimos a escolha do parâmetro de distorção q .

3.4 ESCOLHA DO PARÂMETRO DE DISTORÇÃO q

Quanto às aplicações de entropia presentes na literatura, embora algumas discussões sobre a escolha do q tenham sido feitas frequentemente a partir de considerações físicas, não está claro como fazer esta escolha partindo de uma perspectiva estatística. Em [25] há uma discussão sobre a escolha de q .

Conforme discutido em [25], uma possível abordagem para a escolha do valor de q , por exemplo, é minimizar a estimativa do erro quadrático médio assintótico do estimador quando é matematicamente tratável. Eles apresentam um exemplo no contexto da distribuição exponencial. Em geral, ao contrário do exemplo acima, as expressões de forma fechada do erro quadrático médio assintótico não estão disponíveis, o que exige mais trabalho/pesquisa neste assunto.

Segundo [63], geralmente é muito difícil escolher q de forma analítica. Atualmente, não existe uma maneira universal de fazer essa escolha. Diversos trabalhos escolhem q através de uma abordagem determinística; veja [25], [21] e [63], por exemplo.

Nesta linha de pesquisa, propomos fazer a escolha do parâmetro de distorção q através do procedimento descrito abaixo. Inicialmente, consideramos uma malha para q de 0 a 1 com incrementos de 0,01. Então, para cada q (fixo e conhecido), obtemos as estimativas ML_q , via algoritmo EM_q , dos parâmetros do modelo ST. Neste trabalho, consideramos métodos robustos na presença de outliers, e assim, escolhemos o valor de q que fornece a maior estimativa de probabilidade nas caudas da distribuição skew-t. Em outras palavras, escolhemos o valor de q que fornece o modelo de caudas mais pesadas, ou seja, com menor estimativa para o parâmetro grau de liberdade da ST.

3.5 ROBUSTEZ DAS ESTIMATIVAS ML_q

Uma das vantagens do modelo ST é a sua capacidade de fornecer estimadores robustos ML_q contra outliers, no cenário em que ν é considerado parâmetro desconhecido, e portanto é estimado.

Uma maneira formal de justificar a robustez de um processo de estimação, sob determinado modelo de interesse, é por meio da função de influência (IF), definida como segue:

$$IF(\mathbf{y}, \boldsymbol{\theta}, f) = -E[\boldsymbol{\Psi}'(\mathbf{y}_i, \boldsymbol{\theta}, f)]^{-1} \boldsymbol{\Psi}(\mathbf{y}_i, \boldsymbol{\theta}, f), \quad (3.15)$$

onde $\boldsymbol{\Psi}(\mathbf{y}_i, \boldsymbol{\theta}, f) = U(\boldsymbol{\theta})$ é a função de escore, tal que $\sum_{i=1}^n \boldsymbol{\Psi}(\mathbf{y}_i, \boldsymbol{\theta}, f) = 0$ produz os estimadores de ML de $\boldsymbol{\theta}$, $\boldsymbol{\Psi}'$ é sua derivada e f é a fdp associada à distribuição do modelo de interesse, ST, onde se baseia nosso objetivo.

Em [28], podemos ver que a robustez local dos estimadores M pode ser investigada ao se discutir sua função influência, isto é, os estimadores M com funções de influência limitada são considerados como os estimadores com boas propriedades de robustez local. Observe que a IF é, em certo sentido, proporcional à função de escore. Portanto, para os estimadores, o comportamento da função de influência é determinado pelas funções de escore de cada parâmetro. Se as funções de escore são limitadas, a função influência para o estimador correspondente será limitada. Neste contexto, após uma revisão bibliográfica na literatura Estatística, podemos citar alguns aspectos importantes:

- i) Em [44], na distribuição t-Student, no contexto univariado, as funções de escore para os parâmetros de locação e escala são limitadas quando ν é conhecido (fixo). Assim, a função de influência correspondente para os estimadores de ML é limitada e, portanto, os estimadores de ML para os parâmetros de locação e de escala são localmente robustos, quando ν é conhecido.
- ii) No modelo t-Student multivariado, [21] observou que, ao contrário do caso ML, além das funções de escore para os parâmetros de locação e de escala, a função de escore para o parâmetro ν , obtida a partir do método de estimação ML_q , também é limitada. Neste trabalho, eles mostram a robustez dos estimadores ML_q no modelo t-Student multivariado, quando ν é desconhecido. Esses resultados são uma extensão das conclusões do trabalho de [44].

Além disso, voltando nossa atenção para além dos modelos que pertencem à classe de distribuições SMSN, podemos citar o trabalho de [51] que discute a robustez dos estimadores ML no modelo log-Birnbaum-Saunders-t-Student, quando ν é fixo.

Inspirados por estes trabalhos, a seguir, avaliamos a robustez dos estimadores ML e ML_q no modelo ST, no contexto multivariado, quando ν é conhecido ou não. No que diz respeito às propriedades de robustez local dos estimadores, as funções de escore devem ser limitadas quando $|\mathbf{y}|$ tende ao ∞ . Sem perda de generalidade, as Figuras 3.1-3.10 exibem as funções de escore avaliadas na ST padrão bivariada com $\boldsymbol{\lambda} = (0,5, 1)^\top$ e $\nu = 3$, variando apenas o valor relacionado ao parâmetro de distorção q , no caso ML_q . Para

o parâmetro q , consideramos os seguintes valores 0,5, 0,6 e 0,85, pois as duas primeiras medidas são aproximações dos valores que foram obtidos nas aplicações em dados reais, que poderão ser vistas no próximo capítulo, enquanto que o último valor foi escolhido por ter sido usado em [21].

3.5.1 Cenário 1: Grau de liberdade da ST conhecido

Na Figura 3.1, exibimos as funções de escore para os parâmetros do modelo ST, no caso ML, quando ν é fixo. Observa-se que quando d_λ tende ao ∞ , a função de escore para $\boldsymbol{\mu}$ (letra (a)) tende ao zero (e é limitada) e as funções de escore para $\boldsymbol{\lambda}$ (letra (b)) e $\boldsymbol{\alpha} = (\sigma_{11}, \sigma_{12}, \sigma_{22})^\top$ (letra (c)) são limitadas. Isso significa que outliers, observações que se afastam da massa de dados e por isso apresentam valores grandes de d_λ , têm uma pequena influência nos estimadores de ML para $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ e $\boldsymbol{\lambda}$. Assim, a função de influência correspondente para os estimadores de ML é limitada e, portanto, os estimadores de ML para os parâmetros de localização, de dispersão e de assimetria são localmente robustos, quando ν é conhecido. Portanto, em uma análise estatística para obter estimadores robustos, o parâmetro de graus de liberdade é geralmente assumido como conhecido sob diversos modelos; veja [44], [39], [21], [51] e [64].

Entretanto, pode ser desejável estimar ν simultaneamente com os demais parâmetros, ou seja, adaptar o parâmetro de graus de liberdade para os dados reais ao invés de fixá-lo.

3.5.2 Cenário 2: Grau de liberdade da ST desconhecido

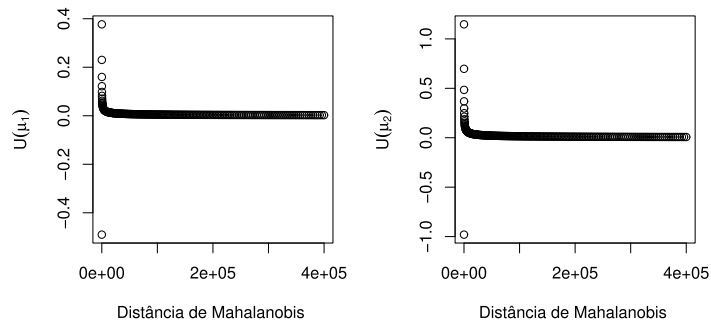
Nas Figuras 3.2-3.10, exibimos as funções de escore para os parâmetros do modelo ST, quando ν é desconhecido, nos casos ML e ML_q , com $q = 0,5, 0,6$ e $0,85$. Nas Figuras 3.2, 3.4 e 3.6, nas letras (a)-(d), note que quando d_λ tende ao ∞ , a função de escore para $\boldsymbol{\mu}$ tende ao zero (e é limitada) sob ambos os métodos de estimação. Nas Figuras 3.2 e 3.4, nas letras (e)-(h), observe que quando d_λ tende ao ∞ , a função de escore para $\boldsymbol{\lambda}$ tende ao zero (e é limitada) no caso ML_q , com $q = 0,5$ e $0,6$, mas no caso ML, as funções são limitadas. O mesmo comportamento é observado nas funções de escore para $\boldsymbol{\alpha}$, sob ambos os métodos de estimação com $q = 0,5$ e $0,6$; veja as Figuras 3.3 e 3.5. Adicionalmente, com $q = 0,85$, ou seja, quando o valor do parâmetro de distorção q vai se aproximando de 1, percebe-se que as funções de escore para $\boldsymbol{\lambda}$ e $\boldsymbol{\alpha}$ sob o método de estimação ML_q são limitadas quando d_λ tende ao ∞ , assim como no caso ML; veja as Figuras 3.6, nas letras (e)-(h), e Figura 3.7.

É importante ressaltar que nas Figuras 3.8-3.10, quando d_λ tende ao ∞ , nota-se que a função de escore para ν sob o método de estimação ML_q tende ao zero (e é limitada) enquanto que no caso ML é ilimitada. Como esperado, para o parâmetro de graus de liberdade ν , também podemos observar que a medida que o valor de q tende

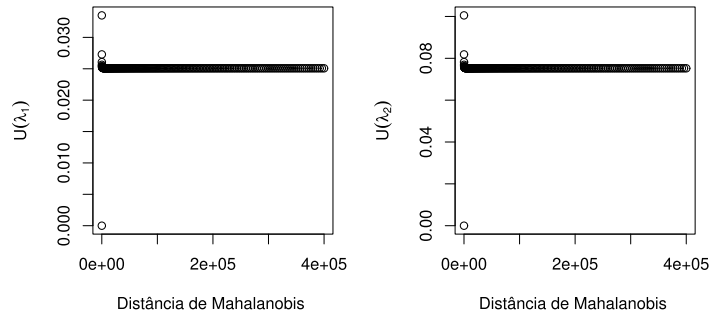
a 1, as funções de escore sob os métodos de estimação ML e ML_q começam a ter um comportamento similar.

Portanto, ao fazer a estimação simultânea dos graus de liberdade com os outros parâmetros da distribuição ST, observou-se que diferentemente dos outros parâmetros, a função de escore para ν é ilimitada sob o método de estimação ML. Assim, os estimadores de ML não são mais localmente robustos devido ao caráter ilimitado da função de escore relacionada ao parâmetro de graus de liberdade. Entretanto, ressaltamos que quando estimamos os graus de liberdade simultaneamente com $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ e $\boldsymbol{\lambda}$ usando o método de estimação de ML_q , os estimadores resultantes terão função de influência limitada que não é o caso para os estimadores de ML. Finalmente, podemos concluir que os estimadores de ML_q no modelo ST são robustos contra outliers, no cenário em que ν é considerado parâmetro desconhecido, e portanto é estimado.

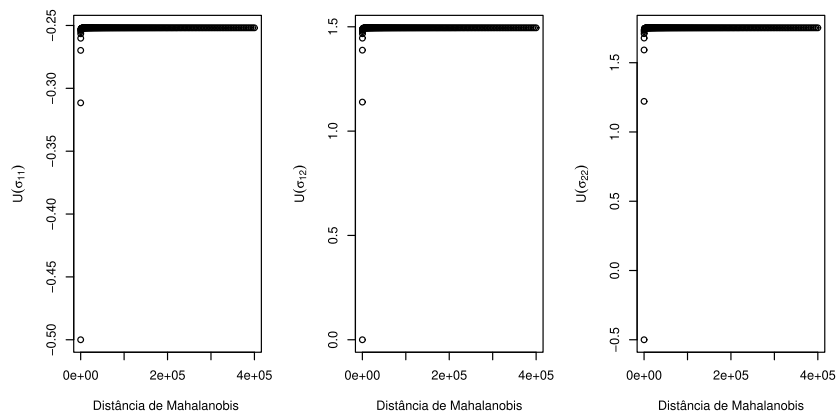
Figura 3.1 – Funções de escore para os parâmetros do modelo ST, no caso ML, quando ν é fixo.



(a)



(b)



(c)

Fonte: Elaborado pela autora em 2020.

Figura 3.2 – Funções de escore para os parâmetros μ e λ do modelo ST, quando ν é desconhecido. Primeira coluna: caso ML e Segunda Coluna: caso ML_q , com $q = 0,5$.

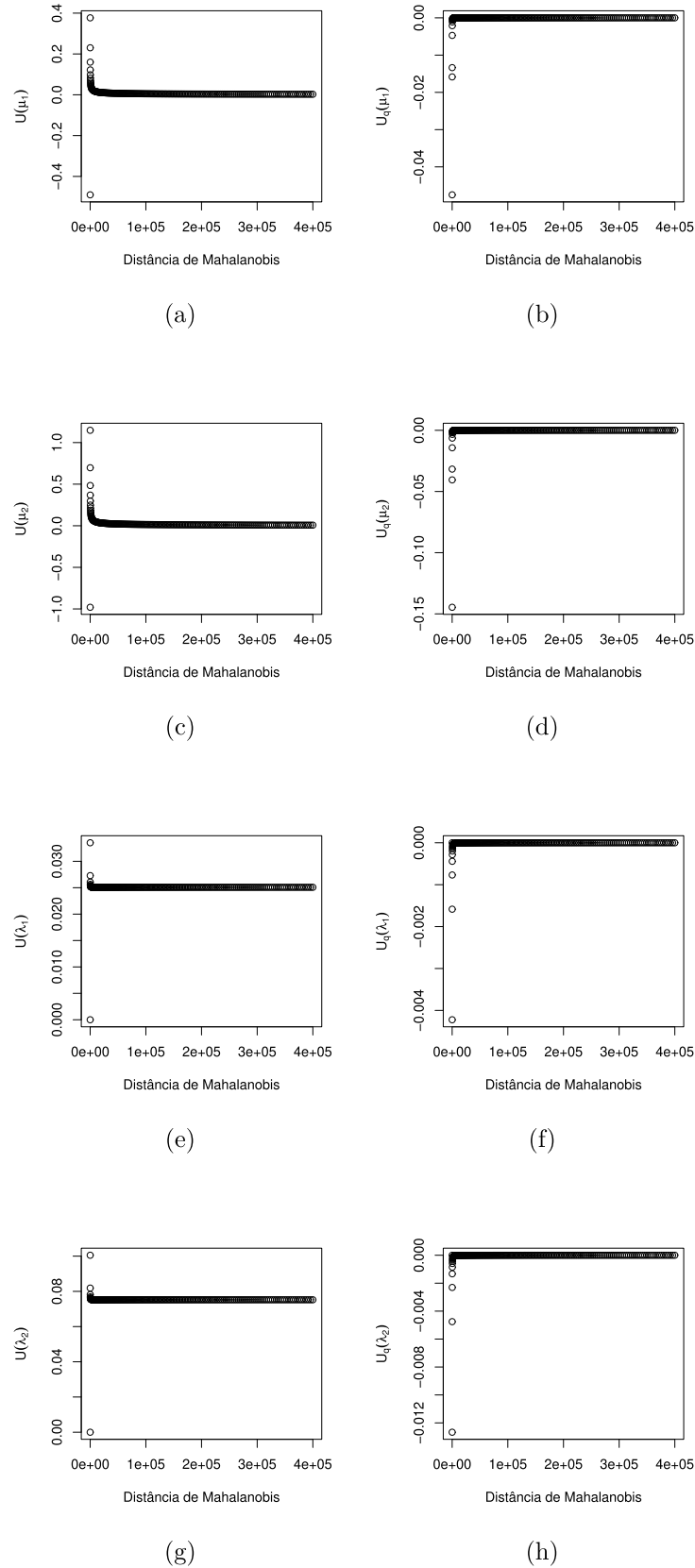
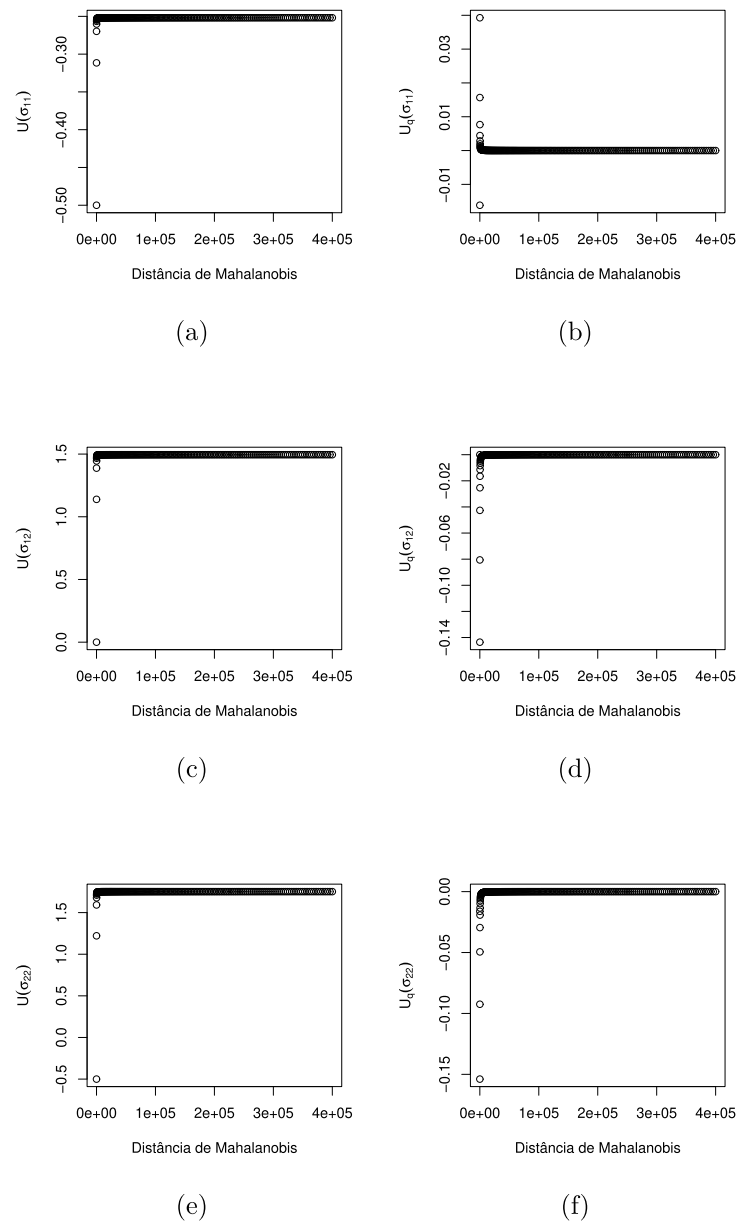


Figura 3.3 – Funções de escore para o parâmetro Σ do modelo ST, quando ν é desconhecido. Primeira coluna: caso ML e Segunda Coluna: caso ML_q , com $q = 0, 5$.



Fonte: Elaborado pela autora em 2020.

Figura 3.4 – Funções de escore para os parâmetros μ e λ do modelo ST, quando ν é desconhecido. Primeira coluna: caso ML e Segunda Coluna: caso ML_q , com $q = 0,6$.

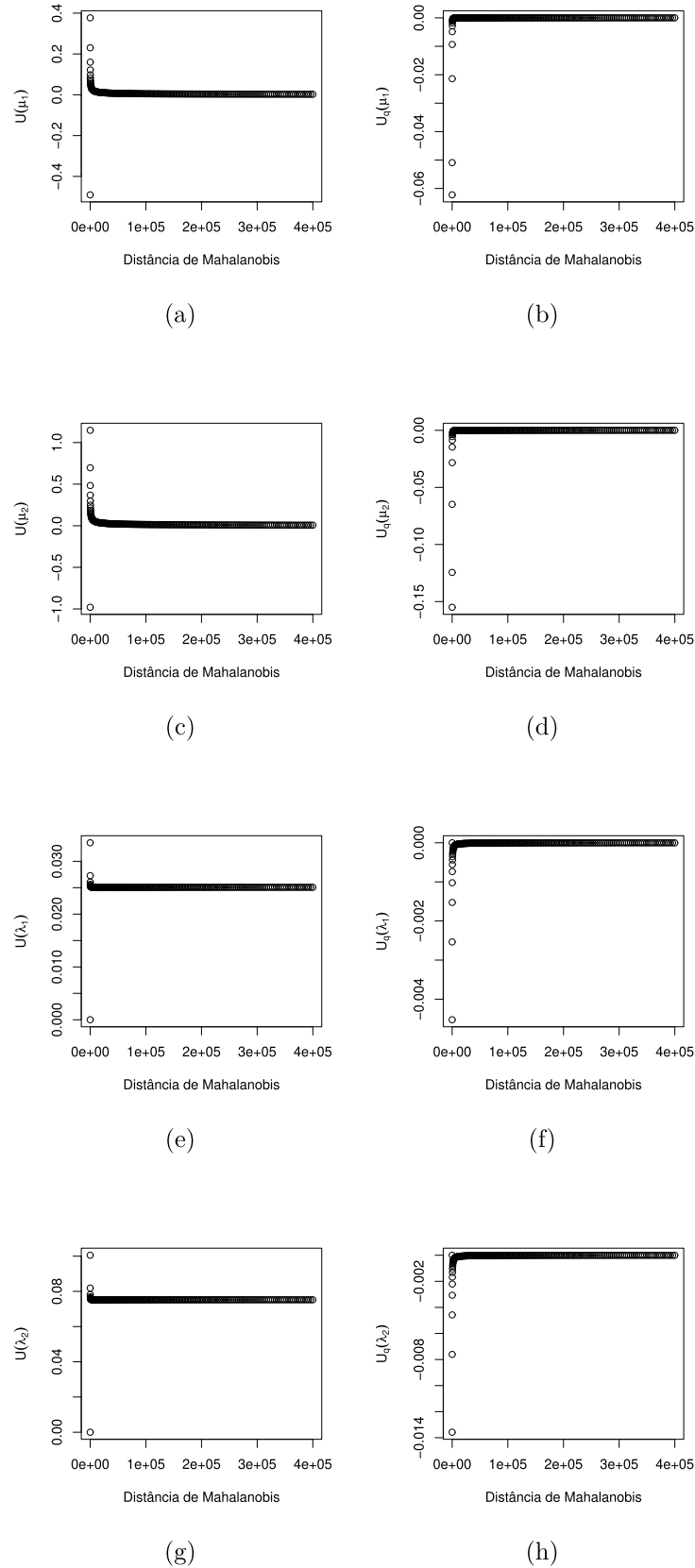
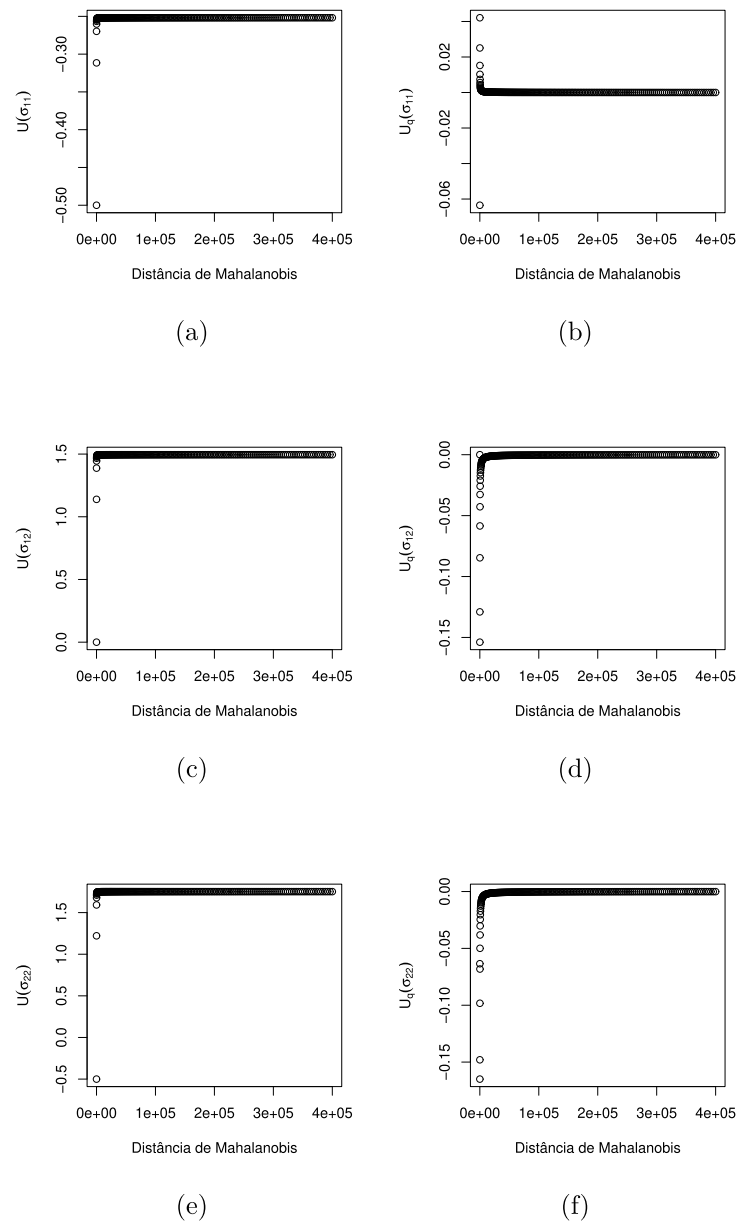


Figura 3.5 – Funções de escore para o parâmetro Σ do modelo ST, quando ν é desconhecido. Primeira coluna: caso ML e Segunda Coluna: caso ML_q , com $q = 0, 6$.



Fonte: Elaborado pela autora em 2020.

Figura 3.6 – Funções de escore para os parâmetros μ e λ do modelo ST, quando ν é desconhecido. Primeira coluna: caso ML e Segunda Coluna: caso ML_q , com $q = 0,85$.

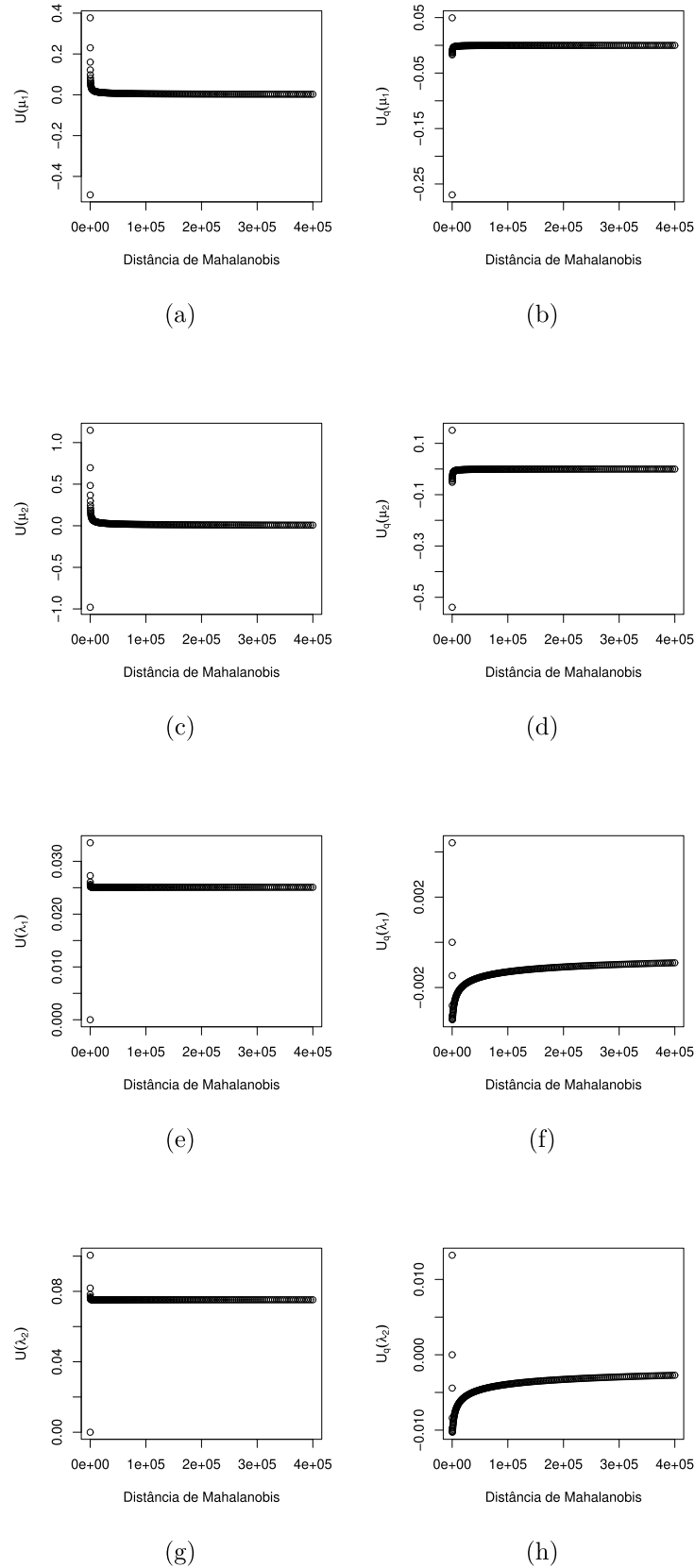
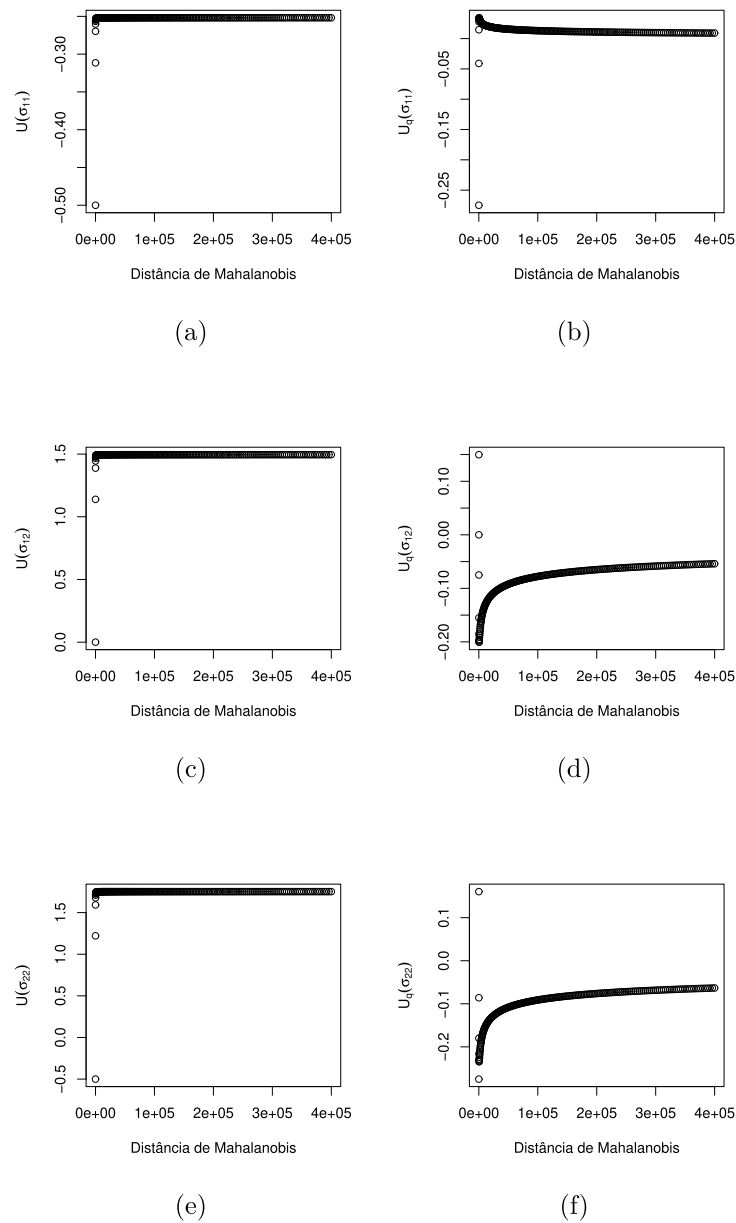
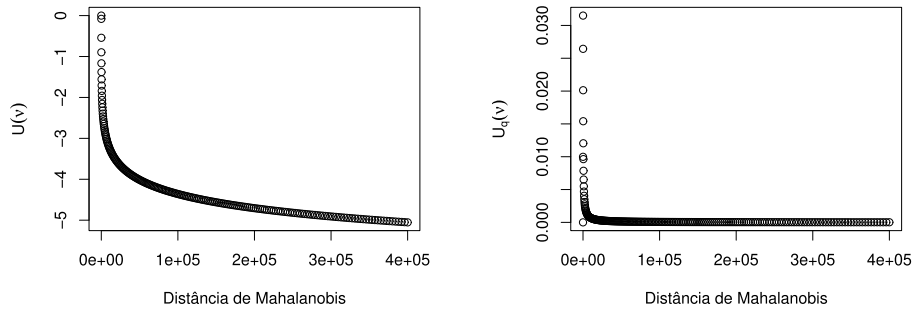


Figura 3.7 – Funções de escore para o parâmetro Σ do modelo ST, quando ν é desconhecido. Primeira coluna: caso ML e Segunda Coluna: caso ML_q , com $q = 0,85$.



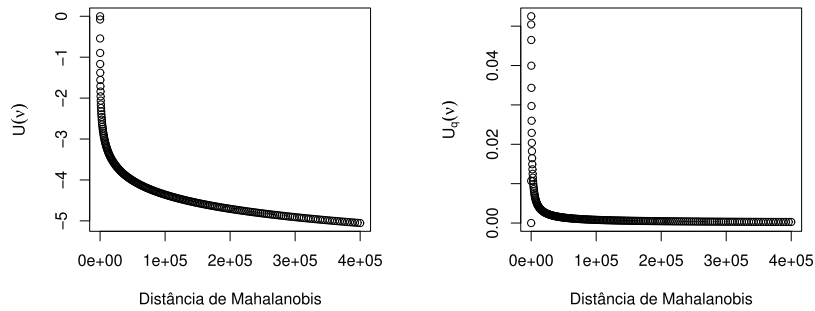
Fonte: Elaborado pela autora em 2020.

Figura 3.8 – Função de escore para o parâmetro ν do modelo ST. Primeira coluna: caso ML e Segunda Coluna: caso ML_q , com $q = 0,5$.



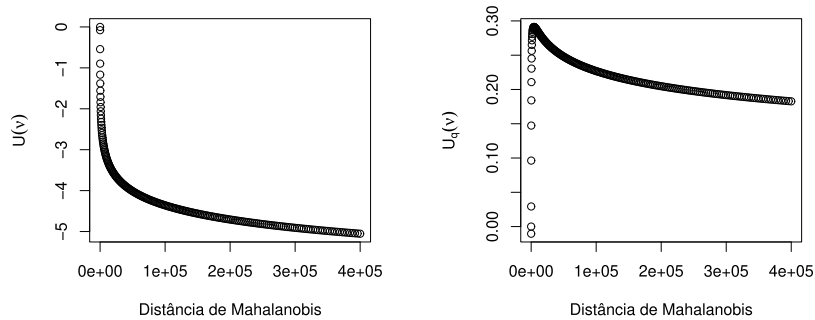
Fonte: Elaborado pela autora em 2020.

Figura 3.9 – Função de escore para o parâmetro ν do modelo ST. Primeira coluna: caso ML e Segunda Coluna: caso ML_q , com $q = 0,6$.



Fonte: Elaborado pela autora em 2020.

Figura 3.10 – Função de escore para o parâmetro ν do modelo ST. Primeira coluna: caso ML e Segunda Coluna: caso ML_q , com $q = 0,85$.



Fonte: Elaborado pela autora em 2020.

4 APLICAÇÕES NUMÉRICAS

Neste capítulo, estudos de simulações e três aplicações em dados reais serão apresentados com o propósito de ilustrar o modelo e os resultados inferenciais aqui desenvolvidos; veja as Seções 4.1 e 4.2. Neste trabalho, consideramos os dados reais Stack-Loss [15], Land Rent [62] e AIS [18] que se destacam por terem sido utilizados em ilustrações numéricas por um grande número de programas estatísticos e autores no contexto de procedimentos robustos, de presença de outliers e de questões relacionadas.

As aplicações numéricas foram realizadas no programa R; veja [58] para mais detalhes. O procedimento de estimação ML, via algoritmo EM, no modelo ST foi feito usando o pacote *mixsmsn* do R, veja [55]. Os demais procedimentos propostos neste trabalho foram programados e os códigos estão disponíveis a pedido.

4.1 ESTUDOS DE SIMULAÇÕES

Nesta seção, apresentamos estudos de simulações para avaliar o desempenho dos estimadores de ML_q propostos sobre os estimadores de ML, no contexto do modelo ST, em termos das medidas viés médio (Bias), viés mediano (MB) e erro quadrático médio (MSE), definidas na Seção 2.4. Estes estudos de simulações foram planejados para observar as mudanças nas estimativas a medida que variamos os tamanhos amostrais ($n = 25, 50, 100, 200$ e 1000) e adicionamos diferentes porcentagens de outliers aos dados (10%, 20% e 30%). Note que ao variar os tamanhos amostrais, podemos mostrar o comportamento assintótico dos estimadores de ML_q obtidos via algoritmo EM_q proposto. Adicionalmente, quando consideramos diferentes porcentagens de outliers, podemos investigar a robustez dos métodos de estimação na presença de outliers. Para cada tamanho amostral, adicionamos outliers aos dados gerados a partir de uma distribuição uniforme de -10 a 10 , do mesmo modo que [21] considerou no contexto do modelo t-Student, veja Seção 5 em [21] para mais informações. Os dados foram gerados artificialmente do modelo ST multivariado sob dois cenários com dimensões distintas, um com duas dimensões ($p = 2$) e outro com quatro dimensões ($p = 4$). Neste contexto, geramos 1000 amostras provenientes do modelo ST com as seguintes configurações:

- **Cenário 1:** $ST_2(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \nu)$, onde $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$, $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)^\top$ e $\nu = 3$.
 - **Caso 1:** $\boldsymbol{\mu} = (2, 1)^\top$, $\boldsymbol{\Sigma} = I_{2 \times 2}$ e $\boldsymbol{\lambda} = (6, -4)^\top$, onde $I_{2 \times 2}$ é a matriz identidade de dimensão 2×2 .
 - **Caso 2:** $\boldsymbol{\mu} = (-2, -1)^\top$, $\boldsymbol{\Sigma} = \begin{bmatrix} 2 & -0,5 \\ -0,5 & 1 \end{bmatrix}$ e $\boldsymbol{\lambda} = (6, 4)^\top$.

- **Cenário 2:** $ST_4(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, \nu)$, onde $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)^\top$, $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_{44} \end{bmatrix}$,

$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)^\top \text{ e } \nu = 5.$$

- **Caso 1:** $\boldsymbol{\mu} = (2, 1, 1, 1)^\top$, $\boldsymbol{\Sigma} = I_{4 \times 4}$ e $\boldsymbol{\lambda} = (6, 3, 2, 1)^\top$.

- **Caso 2:** $\boldsymbol{\mu} = (-2, -1, -1, -1)^\top$, $\boldsymbol{\Sigma} = \begin{bmatrix} 2 & -0,5 & -0,5 & -0,5 \\ -0,5 & 2 & -0,5 & -0,5 \\ -0,5 & -0,5 & 2 & -0,5 \\ -0,5 & -0,5 & -0,5 & 2 \end{bmatrix}$ e

$$\boldsymbol{\lambda} = (6, 3, -2, -1)^\top.$$

Analisando as tabelas e figuras apresentadas abaixo, para cada um dos cenários e casos, podemos concluir que, em geral, para uma determinada porcentagem de outlier, o viés e a variabilidade das estimativas dos parâmetros diminuem, quando o tamanho da amostra aumenta, sob ambos os métodos de estimação. Isso concorda essencialmente com as propriedades assintóticas dos estimadores de máxima verossimilhança. Assim, podemos concluir que os estimadores de ML_q também são consistentes. Além disso, em geral, quando o tamanho da amostra é fixo, observamos que uma porcentagem de outliers crescente, adicionada aos dados, corresponde a um aumento do viés e da variabilidade das estimativas dos parâmetros sob ambos métodos de estimação. Adicionalmente, as Figuras (A.1)-(A.4), consulte o Apêndice A, apresentam a média das medidas MSE e MB de cada parâmetro nos cenários considerados. Note que o distanciamento dos estimadores de ML_q para os de ML, em termos de viés e MSE, ocorre à medida que aumenta-se a contaminação dos dados. Assim, nota-se claramente que as estimativas de ML_q são robustas na presença de outliers, pois foram menos afetadas (menos sensíveis) pelas diferentes contaminações de outliers do que as estimativas de ML. Em outras palavras, sob o modelo ST, a observação atípica parece ter mais impacto no processo de estimação de ML.

4.1.1 Cenário 1: Caso 1

As Tabelas (4.1)-(4.3) mostram as medidas Bias, MB e MSE dos estimadores de ML e de ML_q dos parâmetros do modelo ST, com o adicional de 10%, 20% e 30% de outliers para $n = 25, 50, 100, 200$ e 1000, no contexto do cenário 1 - caso 1, descrito acima. Além disso, para facilitar a visualização dessas informações, apresentamos as Figuras (4.1)-(4.2).

Analisando as Figuras (4.2) (a), (b), (e) e (f), nota-se que o MSE e o MB dos estimadores dos parâmetros σ_{11} e σ_{22} , no caso ML, considerando as porcentagens de outliers de 10% e 20%, são menores do que as respectivas medidas desses estimadores para o

caso ML_q . Entretanto, com 30% de observações anômalas, destaca-se um comportamento diferente, ou seja, observa-se que o MSE e o MB dos estimadores dos parâmetros σ_{11} e σ_{22} , no caso ML_q , para os diversos valores de n adotados, são menores do que as respectivas medidas desses estimadores para o caso ML. É importante ressaltar que conforme visto na seção 3.5.2, os estimadores de ML_q no modelo ST são robustos contra outliers quando ν é considerado parâmetro desconhecido, e por tanto é estimado.

Tabela 4.1 – Cenário 1: Caso 1 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 10% de outliers.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
25	μ_1	2	0,7070	0,4989	0,3057	0,6410	0,4337	0,2583
	μ_2	1	0,6497	0,4976	0,3815	0,5834	0,4515	0,3495
	σ_{11}	1	0,2778	0,2218	0,1879	0,4061	0,3250	0,2720
	σ_{12}	0	0,4116	0,3326	0,2803	0,2928	0,2182	0,1655
	σ_{22}	1	0,3219	0,2431	0,1833	0,3945	0,3050	0,2385
	λ_1	6	4,7480	4,3518	4,3363	4,4285	4,0395	4,1957
	λ_2	-4	3,2421	2,8716	2,7785	3,1523	2,7877	2,7210
	ν	3	9,8702	2,1581	1,0752	1,3979	1,0501	0,9105
50	μ_1	2	0,5365	0,3720	0,2348	0,4584	0,3157	0,2048
	μ_2	1	0,5103	0,3858	0,2806	0,4747	0,3546	0,2611
	σ_{11}	1	0,2084	0,1659	0,1402	0,3873	0,3216	0,3036
	σ_{12}	0	0,3364	0,2861	0,2600	0,2520	0,1899	0,1431
	σ_{22}	1	0,2168	0,1648	0,1308	0,3587	0,2857	0,2346
	λ_1	6	4,6952	4,5333	4,5387	4,3044	4,0380	4,1778
	λ_2	-4	3,1057	2,9304	2,8337	2,9860	2,7363	2,7038
	ν	3	1,1531	1,0494	1,1033	1,1148	0,9055	0,8144
100	μ_1	2	0,4145	0,2999	0,2190	0,3446	0,2516	0,1910
	μ_2	1	0,3769	0,2818	0,2076	0,3621	0,2728	0,2028
	σ_{11}	1	0,1497	0,1220	0,1070	0,4101	0,3445	0,3154
	σ_{12}	0	0,2698	0,2445	0,2286	0,1862	0,1437	0,1118
	σ_{22}	1	0,1490	0,1155	0,0920	0,3785	0,3036	0,2515
	λ_1	6	4,6851	4,6158	4,5764	4,2278	4,0253	4,2309
	λ_2	-4	3,0463	2,9660	2,8742	2,8816	2,6973	2,7124
	ν	3	1,1291	1,0859	1,1325	1,2403	0,9986	0,8805
200	μ_1	2	0,3144	0,2421	0,1866	0,2754	0,2051	0,1531
	μ_2	1	0,2863	0,2077	0,1366	0,2919	0,2153	0,1481
	σ_{11}	1	0,1191	0,0964	0,0854	0,4132	0,3563	0,3388
	σ_{12}	0	0,2292	0,2168	0,2100	0,1443	0,1122	0,0868
	σ_{22}	1	0,1123	0,0898	0,0736	0,3918	0,3195	0,2872
	λ_1	6	4,5807	4,5452	4,4624	4,0341	3,8415	4,0528
	λ_2	-4	2,9268	2,8701	2,7925	2,7277	2,5610	2,5920
	ν	3	1,1213	1,1034	1,1174	1,3933	1,1409	0,9410
1000	μ_1	2	0,1324	0,1256	0,1250	0,1011	0,0848	0,0805
	μ_2	1	0,0729	0,0591	0,0489	0,0789	0,0611	0,0518
	σ_{11}	1	0,0662	0,0569	0,0552	0,3046	0,2580	0,2409
	σ_{12}	0	0,1379	0,1359	0,1357	0,0703	0,0573	0,0483
	σ_{22}	1	0,0468	0,0366	0,0296	0,2926	0,2317	0,2339
	λ_1	6	3,1382	3,1169	3,1432	2,2746	2,0090	2,2126
	λ_2	-4	1,5874	1,5456	1,5735	1,3353	1,1934	1,2305
	ν	3	1,1716	1,1687	1,1738	1,2726	1,0815	0,9699

Fonte: Elaborado pela autora em 2020.

Tabela 4.2 – Cenário 1: Caso 1 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 20% de outliers.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
25	μ_1	2	0,9038	0,6547	0,4320	0,7205	0,5175	0,3417
	μ_2	1	0,8407	0,6420	0,4902	0,6788	0,5214	0,4045
	σ_{11}	1	0,5503	0,3993	0,2779	0,4368	0,3569	0,3129
	σ_{12}	0	0,7676	0,6382	0,5297	0,4528	0,3461	0,2642
	σ_{22}	1	0,6281	0,4593	0,3142	0,4366	0,3492	0,2937
	λ_1	6	5,0654	4,7837	4,8629	4,6780	4,3367	4,4870
	λ_2	-4	3,4587	3,1289	3,0075	3,3473	3,0068	2,9405
	ν	3	7,7983	2,0124	1,3683	1,3455	1,1268	1,0845
50	μ_1	2	0,6057	0,4359	0,3031	0,5202	0,3743	0,2662
	μ_2	1	0,5999	0,4459	0,3226	0,5276	0,4019	0,3149
	σ_{11}	1	0,3127	0,2253	0,1602	0,4046	0,3322	0,2948
	σ_{12}	0	0,5484	0,4939	0,4564	0,3613	0,2744	0,2143
	σ_{22}	1	0,3639	0,2729	0,1999	0,3945	0,3209	0,2789
	λ_1	6	4,9527	4,8582	4,8557	4,5102	4,2435	4,4922
	λ_2	-4	3,2328	3,0990	3,0404	3,1302	2,8914	2,8843
	ν	3	1,3063	1,2373	1,3169	1,1902	1,0012	1,0050
100	μ_1	2	0,4723	0,3520	0,2572	0,4007	0,2923	0,2116
	μ_2	1	0,4239	0,3171	0,2283	0,4025	0,3046	0,2373
	σ_{11}	1	0,2021	0,1532	0,1159	0,4233	0,3600	0,3529
	σ_{12}	0	0,4612	0,4380	0,4194	0,2698	0,2009	0,1375
	σ_{22}	1	0,2629	0,2100	0,1814	0,4036	0,3399	0,3150
	λ_1	6	4,9063	4,8555	4,8100	4,3122	4,0583	4,3498
	λ_2	-4	3,1158	3,0407	3,0117	2,9694	2,7578	2,8362
	ν	3	1,3254	1,2986	1,3337	1,2716	1,0565	1,0147
200	μ_1	2	0,3770	0,2932	0,2243	0,2976	0,2251	0,1689
	μ_2	1	0,3211	0,2405	0,1677	0,3149	0,2333	0,1693
	σ_{11}	1	0,1597	0,1267	0,1051	0,4523	0,3987	0,4237
	σ_{12}	0	0,4130	0,4035	0,3980	0,1987	0,1425	0,0909
	σ_{22}	1	0,2217	0,1873	0,1748	0,4234	0,3620	0,3591
	λ_1	6	4,8599	4,8304	4,7778	4,1125	3,9136	4,1531
	λ_2	-4	3,0498	3,0011	2,9698	2,7611	2,5569	2,6268
	ν	3	1,3296	1,3187	1,3295	1,4546	1,2167	1,1366
1000	μ_1	2	0,1729	0,1662	0,1649	0,1016	0,0820	0,0710
	μ_2	1	0,0925	0,0768	0,0700	0,0973	0,0759	0,0627
	σ_{11}	1	0,0908	0,0796	0,0767	0,3554	0,3268	0,3138
	σ_{12}	0	0,3046	0,3032	0,3030	0,0809	0,0612	0,0499
	σ_{22}	1	0,1712	0,1637	0,1622	0,3477	0,3163	0,3002
	λ_1	6	3,5628	3,5442	3,5868	2,1924	1,9124	1,9737
	λ_2	-4	1,7455	1,7012	1,7358	1,3923	1,2227	1,2174
	ν	3	1,3876	1,3859	1,3907	1,1540	0,9206	0,7568

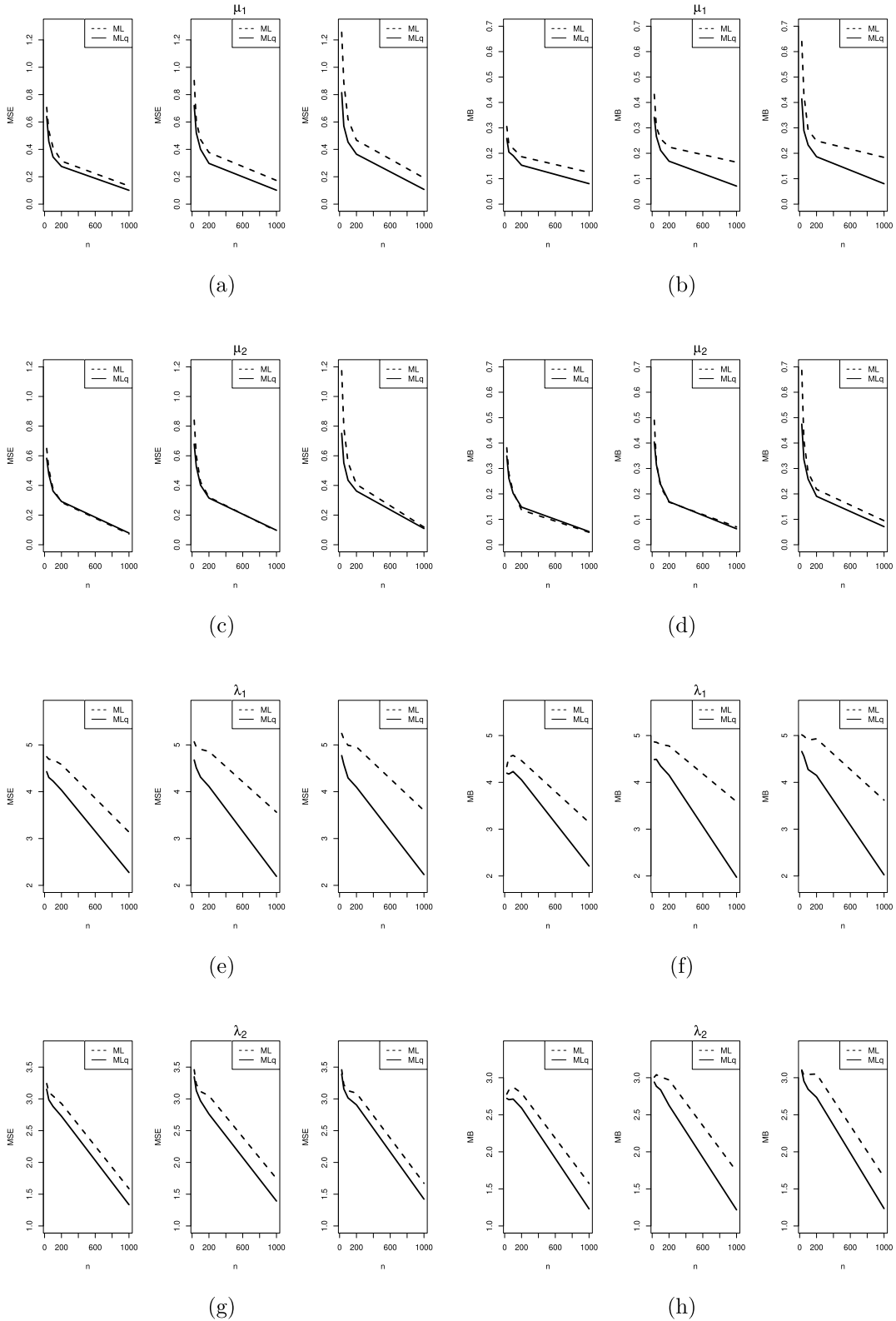
Fonte: Elaborado pela autora em 2020.

Tabela 4.3 – Cenário 1: Caso 1 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 30% de outliers.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
25	μ_1	2	1,2559	0,9308	0,6401	0,8150	0,5983	0,4140
	μ_2	1	1,1743	0,9043	0,6863	0,7514	0,5908	0,4748
	σ_{11}	1	1,1148	0,8940	0,7524	0,5712	0,4651	0,4173
	σ_{12}	0	1,3471	1,1769	1,0155	0,6610	0,4998	0,4088
	σ_{22}	1	1,1725	0,9533	0,8020	0,5978	0,4810	0,4141
	λ_1	6	5,2451	5,0164	5,0154	4,7717	4,5005	4,6590
	λ_2	-4	3,4580	3,1967	3,1037	3,3917	3,1011	3,0954
	ν	3	12,0122	3,3102	1,3685	1,4912	1,2029	1,1701
50	μ_1	2	0,8882	0,6336	0,4319	0,5707	0,4114	0,2892
	μ_2	1	0,7930	0,5841	0,4119	0,5519	0,4198	0,3354
	σ_{11}	1	0,7055	0,5721	0,5003	0,4727	0,4080	0,3946
	σ_{12}	0	0,9724	0,8811	0,7858	0,4743	0,3515	0,2514
	σ_{22}	1	0,7825	0,6502	0,5670	0,4814	0,4114	0,4006
	λ_1	6	5,1266	5,0323	4,9874	4,5857	4,3091	4,5598
	λ_2	-4	3,2258	3,1008	3,0335	3,1573	2,9113	2,9540
	ν	3	1,3240	1,1711	1,2157	1,1687	0,9729	0,9454
100	μ_1	2	0,6200	0,4344	0,2915	0,4533	0,3301	0,2326
	μ_2	1	0,5549	0,4045	0,2844	0,4351	0,3349	0,2584
	σ_{11}	1	0,5472	0,4805	0,4529	0,4445	0,3860	0,3784
	σ_{12}	0	0,8088	0,7711	0,7304	0,3800	0,2743	0,1527
	σ_{22}	1	0,6130	0,5473	0,5125	0,4486	0,3954	0,3945
	λ_1	6	4,9910	4,9418	4,9004	4,2949	4,0064	4,2746
	λ_2	-4	3,1285	3,0487	3,0440	3,0157	2,7611	2,8440
	ν	3	1,2093	1,1601	1,2178	1,2199	0,9733	0,8760
200	μ_1	2	0,4671	0,3452	0,2489	0,3648	0,2613	0,1865
	μ_2	1	0,4071	0,2975	0,2180	0,3621	0,2670	0,1905
	σ_{11}	1	0,4911	0,4544	0,4379	0,4590	0,4102	0,4209
	σ_{12}	0	0,7498	0,7340	0,7132	0,3027	0,2025	0,1039
	σ_{22}	1	0,5568	0,5217	0,5061	0,4583	0,4102	0,4071
	λ_1	6	4,9570	4,9278	4,9315	4,0974	3,8419	4,1483
	λ_2	-4	3,0905	3,0401	3,0501	2,9062	2,6874	2,7349
	ν	3	1,2109	1,1849	1,2176	1,3874	1,0746	0,9320
1000	μ_1	2	0,1912	0,1832	0,1838	0,1075	0,0890	0,0803
	μ_2	1	0,1189	0,1017	0,0945	0,1095	0,0861	0,0712
	σ_{11}	1	0,3913	0,3853	0,3845	0,3949	0,3743	0,3521
	σ_{12}	0	0,5810	0,5791	0,5779	0,1037	0,0722	0,0594
	σ_{22}	1	0,4723	0,4665	0,4640	0,3884	0,3636	0,3459
	λ_1	6	3,5980	3,5787	3,6206	2,2301	1,9676	2,0238
	λ_2	-4	1,6719	1,6237	1,6578	1,4226	1,2323	1,2347
	ν	3	1,2871	1,2835	1,2900	1,0812	0,8743	0,7293

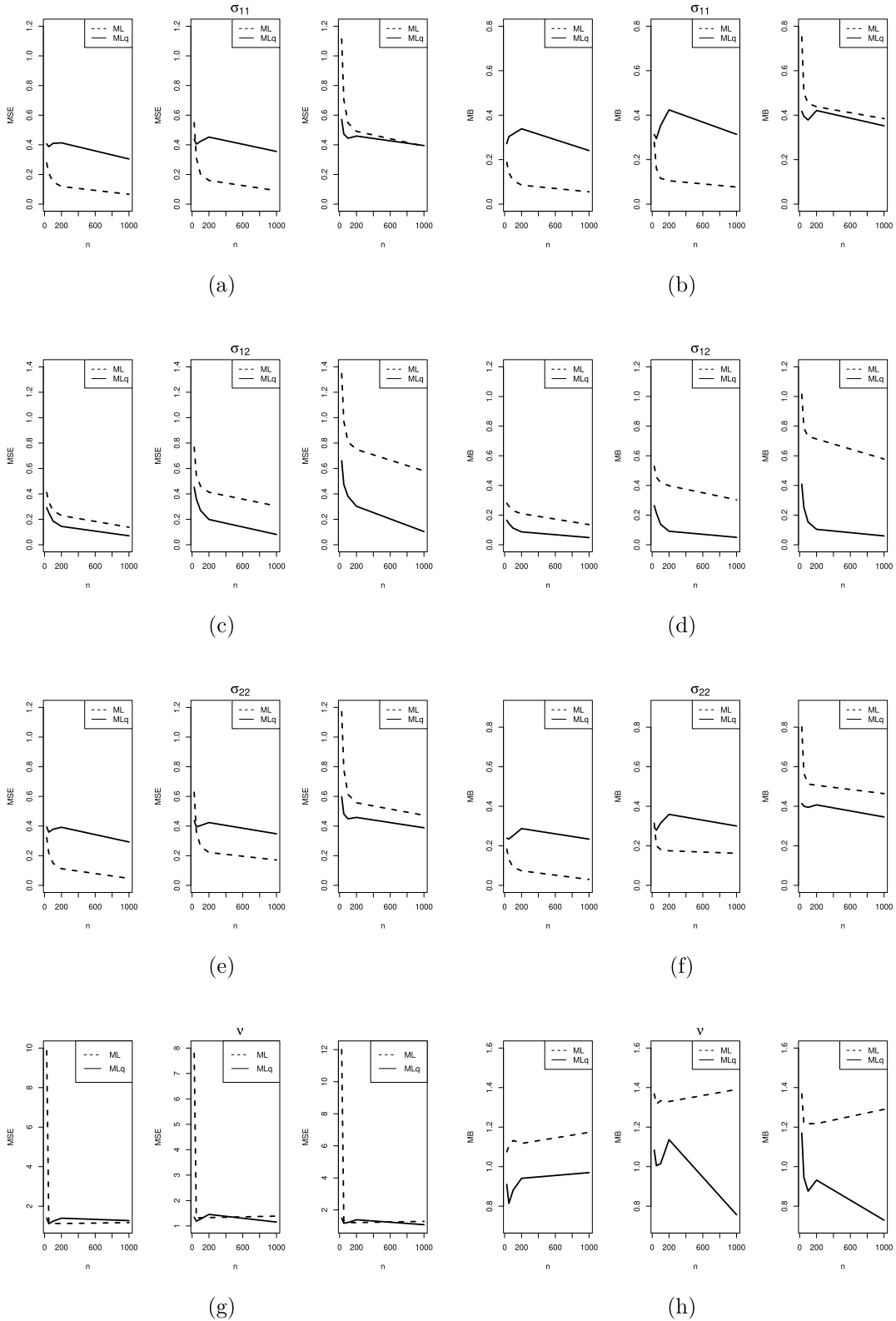
Fonte: Elaborado pela autora em 2020.

Figura 4.1 – Cenário 1: Caso 1 - MSE e MB dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros $\boldsymbol{\mu}$ e $\boldsymbol{\lambda}$ do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



Fonte: Elaborado pela autora em 2020.

Figura 4.2 – Cenário 1: Caso 1 - MSE e MB dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros Σ e ν do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



Fonte: Elaborado pela autora em 2020.

4.1.2 Cenário 1: Caso 2

As Tabelas (4.4)-(4.6) mostram as medidas Bias, MB e MSE dos estimadores de ML e de ML_q dos parâmetros do modelo ST, com o adicional de 10%, 20% e 30% de outliers para $n = 25, 50, 100, 200$ e 1000, no contexto do cenário 1 - caso 2, descrito anteriormente. Além disso, para facilitar a visualização dessas informações, apresentamos as Figuras (4.3)-(4.4).

Analisando as Figuras (4.4) (a), (b), (e) e (f), nota-se que o MSE e o MB dos estimadores dos parâmetros σ_{11} e σ_{22} , no caso ML, considerando as porcentagens de outliers de 10% e 20%, são menores do que as respectivas medidas desses estimadores para o caso ML_q . Entretanto, com 30% de observações anômalas, destaca-se um comportamento diferente, ou seja, observa-se que o MSE e o MB dos estimadores dos parâmetros σ_{11} e σ_{22} , no caso ML_q , para os diversos valores de n adotados, são menores do que as respectivas medidas desses estimadores para o caso ML.

Concluindo, pode-se observar o mesmo comportamento do cenário 1: caso 1. Os resultados obtidos no cenário 1: casos 1 e 2 concordam com os obtidos no trabalho de [56] no contexto de misturas finitas de distribuições. Segundo [56], com um valor apropriado para o parâmetro q , onde $0 < q < 1$, o estimador de ML_q irá ignorar parcialmente as observações, do conjunto de dados, nas caudas focando mais no ajuste das observações ao redor do centro. O estimador de ML_q obtido desta forma é possivelmente enviesado, em particular para os parâmetros de escala, mas será menos volátil para uma mudança significativa de dados nas caudas, portanto, um bom exemplo de troca (intercâmbio) de viés-variância. Assim, q pode ser considerado como um parâmetro de ajuste que ajusta a magnitude da compensação de viés-variância.

Tabela 4.4 – Cenário 1: Caso 2 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 10% de outliers.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
25	μ_1	-2	0,8506	0,6284	0,4238	0,7914	0,5668	0,3805
	μ_2	-1	0,6151	0,4880	0,4004	0,5586	0,4359	0,3409
	σ_{11}	2	0,7731	0,7085	0,7307	1,1132	1,0171	0,9806
	σ_{12}	-0,5	0,3833	0,3285	0,3134	0,3792	0,3357	0,3320
	σ_{22}	1	0,2854	0,2235	0,1827	0,4443	0,3518	0,2914
	λ_1	6	5,0033	4,5503	4,9227	4,3717	3,9271	4,2930
	λ_2	4	3,5876	3,1996	3,3137	3,0822	2,7061	2,8203
	ν	3	3,5351	1,2853	1,1446	1,3776	1,0458	0,8226
50	μ_1	-2	0,6712	0,4976	0,3559	0,5730	0,4134	0,2925
	μ_2	-1	0,5173	0,4033	0,3166	0,4503	0,3470	0,2847
	σ_{11}	2	0,7889	0,7514	0,7714	1,1182	1,0605	1,0655
	σ_{12}	-0,5	0,3506	0,3156	0,3068	0,3646	0,3310	0,3333
	σ_{22}	1	0,1988	0,1558	0,1320	0,3990	0,3283	0,2966
	λ_1	6	5,1206	4,9888	5,0529	4,2641	3,9535	4,1573
	λ_2	4	3,5579	3,3923	3,4023	3,0063	2,7051	2,7518
	ν	3	1,3296	1,2662	1,3425	1,3044	1,0013	0,7934
100	μ_1	-2	0,5230	0,3801	0,2643	0,3942	0,2912	0,2315
	μ_2	-1	0,4311	0,3299	0,2404	0,3670	0,2816	0,2194
	σ_{11}	2	0,7920	0,7720	0,7841	1,1352	1,0941	1,0919
	σ_{12}	-0,5	0,3164	0,2954	0,2934	0,3606	0,3360	0,3439
	σ_{22}	1	0,1443	0,1171	0,1028	0,4163	0,3494	0,3364
	λ_1	6	5,0719	5,0196	4,9946	3,9533	3,6570	3,9213
	λ_2	4	3,4955	3,4130	3,3670	2,8611	2,5796	2,5030
	ν	3	1,3728	1,3467	1,3895	1,3953	1,0671	0,8456
200	μ_1	-2	0,4079	0,2986	0,2251	0,3025	0,2247	0,1783
	μ_2	-1	0,3724	0,2741	0,1740	0,3057	0,2358	0,1858
	σ_{11}	2	0,7889	0,7778	0,7754	1,1264	1,0878	1,0755
	σ_{12}	-0,5	0,3097	0,2950	0,3030	0,3663	0,3472	0,3583
	σ_{22}	1	0,1062	0,0861	0,0747	0,4125	0,3407	0,3238
	λ_1	6	5,0105	4,9865	4,9505	3,8134	3,5389	3,9134
	λ_2	4	3,4519	3,4014	3,2437	2,7359	2,4806	2,4042
	ν	3	1,3987	1,3898	1,4075	1,5699	1,2190	0,9924
1000	μ_1	-2	0,2685	0,2581	0,2552	0,0948	0,0765	0,0654
	μ_2	-1	0,1124	0,0914	0,0793	0,0682	0,0543	0,0455
	σ_{11}	2	0,8972	0,8960	0,8981	0,9431	0,9405	0,9406
	σ_{12}	-0,5	0,3072	0,3061	0,3062	0,3527	0,3519	0,3511
	σ_{22}	2	0,1420	0,1372	0,1374	0,2406	0,2351	0,2320
	λ_1	6	4,9043	4,9004	4,9153	1,1396	0,9459	0,8209
	λ_2	4	3,2561	3,2500	3,2540	0,6958	0,5529	0,4588
	ν	3	1,5515	1,5505	1,5522	0,6951	0,5841	0,5275

Fonte: Elaborado pela autora em 2020.

Tabela 4.5 – Cenário 1: Caso 2 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 20% de outliers.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
25	μ_1	-2	1,0093	0,7559	0,5318	0,8040	0,5952	0,4287
	μ_2	-1	0,7590	0,5874	0,4483	0,5929	0,4730	0,3930
	σ_{11}	2	0,7519	0,6705	0,6958	1,1677	1,0739	1,0958
	σ_{12}	-0,5	0,6756	0,5746	0,4903	0,5018	0,4271	0,4071
	σ_{22}	1	0,5668	0,3875	0,2364	0,5079	0,4189	0,3695
	λ_1	6	5,5389	5,3003	5,4243	4,5236	4,1330	4,5277
	λ_2	4	3,9724	3,7738	3,7656	3,2854	2,9585	3,0712
	ν	3	1,7307	1,4970	1,5604	1,3649	1,0497	0,8581
50	μ_1	-2	0,8024	0,5959	0,4147	0,6115	0,4388	0,3066
	μ_2	-1	0,6319	0,5001	0,4146	0,5122	0,4094	0,3429
	σ_{11}	2	0,7036	0,6506	0,6826	1,1695	1,1111	1,1111
	σ_{12}	-0,5	0,5199	0,4746	0,4419	0,4035	0,3595	0,3506
	σ_{22}	1	0,3331	0,2383	0,1723	0,4400	0,3726	0,3414
	λ_1	6	5,5216	5,4495	5,4323	4,3071	3,9270	4,3979
	λ_2	4	3,9142	3,8171	3,8421	3,1868	2,8412	2,9514
	ν	3	1,5337	1,4803	1,5756	1,3001	1,0067	0,8207
100	μ_1	-2	0,6585	0,4971	0,3665	0,4674	0,3314	0,2346
	μ_2	-1	0,4867	0,3826	0,3029	0,3929	0,3059	0,2429
	σ_{11}	2	0,7201	0,6921	0,7106	1,1712	1,1347	1,1145
	σ_{12}	-0,5	0,4426	0,4237	0,4181	0,3685	0,3400	0,3409
	σ_{22}	1	0,1997	0,1462	0,1110	0,4196	0,3644	0,3488
	λ_1	6	5,4806	5,4547	5,3800	3,9860	3,5946	3,9074
	λ_2	4	3,8357	3,7948	3,7565	2,9401	2,5907	2,7031
	ν	3	1,6021	1,5845	1,6222	1,1509	0,8953	0,7100
200	μ_1	-2	0,5318	0,4145	0,3193	0,3181	0,2354	0,1812
	μ_2	-1	0,4314	0,3322	0,2375	0,3203	0,2400	0,1759
	σ_{11}	2	0,7228	0,7101	0,7242	1,1466	1,1215	1,0804
	σ_{12}	-0,5	0,4240	0,4149	0,4167	0,3547	0,3385	0,3503
	σ_{22}	1	0,1331	0,1027	0,0811	0,4057	0,3562	0,3234
	λ_1	6	5,4487	5,4358	5,3622	3,6544	3,2673	3,3801
	λ_2	4	3,8166	3,7897	3,6932	2,7172	2,3297	2,1853
	ν	3	1,6195	1,6142	1,6287	1,2201	0,9513	0,7579
1000	μ_1	-2	0,4184	0,4089	0,4047	0,1110	0,0914	0,0810
	μ_2	-1	0,1919	0,1707	0,1643	0,0770	0,0618	0,0538
	σ_{11}	2	0,8861	0,8849	0,8865	1,0069	1,0037	0,9992
	σ_{12}	-0,5	0,3900	0,3893	0,3897	0,3764	0,3755	0,3759
	σ_{22}	2	0,0833	0,0759	0,0760	0,2768	0,2695	0,2674
	λ_1	6	5,4577	5,4567	5,4576	1,2522	1,0478	0,9572
	λ_2	4	3,7196	3,7176	3,7107	0,7696	0,6156	0,5239
	ν	3	1,7691	1,7685	1,7704	0,6634	0,5490	0,4876

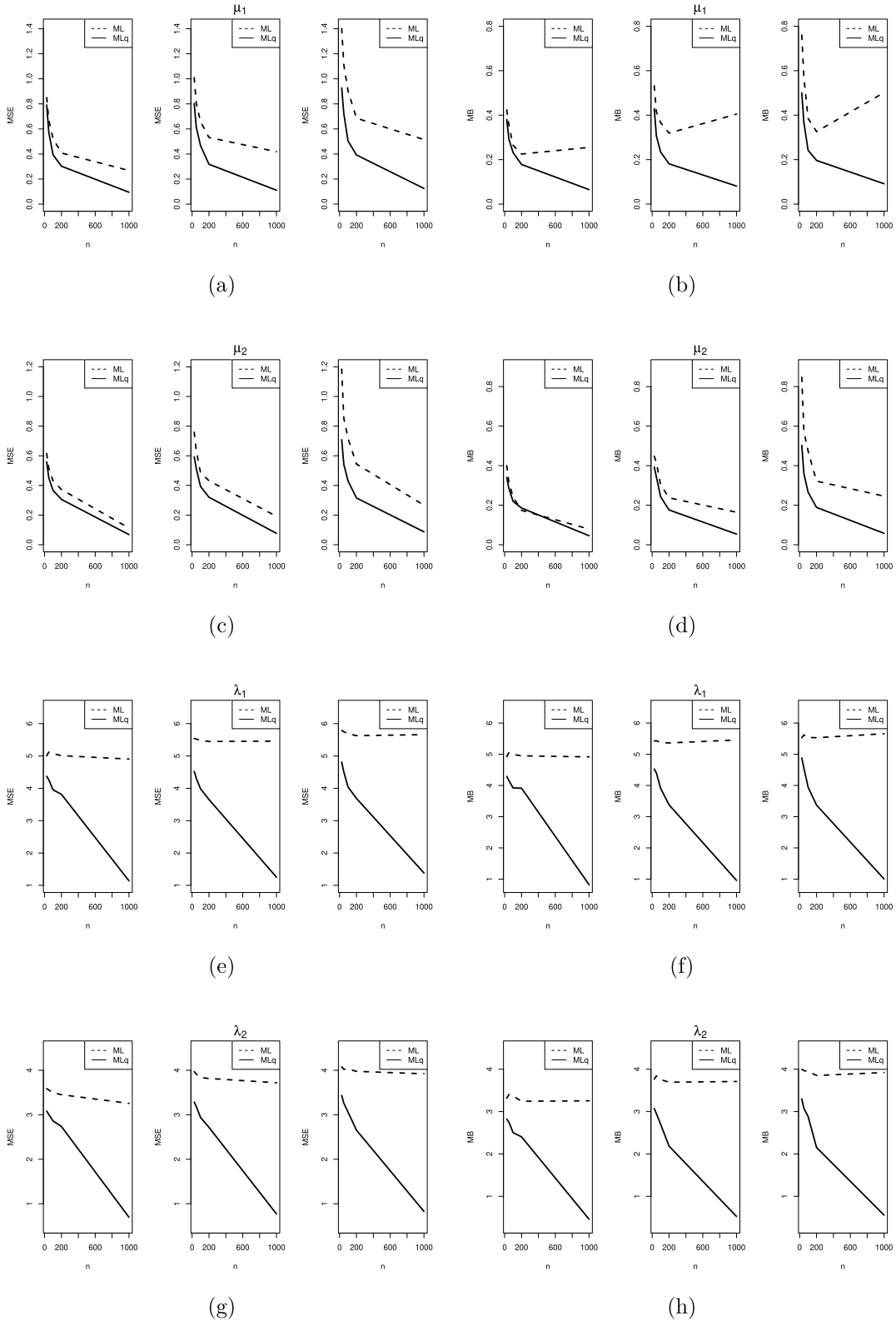
Fonte: Elaborado pela autora em 2020.

Tabela 4.6 – Cenário 1: Caso 2 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 30% de outliers.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
25	μ_1	-2	1,4019	1,0577	0,7601	0,9283	0,6931	0,4998
	μ_2	-1	1,1848	0,9589	0,8475	0,7084	0,5741	0,5027
	σ_{11}	2	0,8683	0,7078	0,6319	1,2194	1,1029	1,1252
	σ_{12}	-0,5	1,3754	1,2104	1,0585	0,6858	0,5395	0,4641
	σ_{22}	1	1,3413	1,0779	0,9566	0,6884	0,5556	0,4926
	λ_1	6	5,7858	5,6836	5,5290	4,8112	4,5003	4,8834
	λ_2	4	4,0787	3,9468	3,9973	3,4356	3,1625	3,3014
	ν	3	11,2238	3,0822	1,4696	1,6206	1,2292	1,0154
50	μ_1	-2	1,1135	0,8281	0,5692	0,7235	0,5262	0,3661
	μ_2	-1	0,8538	0,6758	0,5715	0,5425	0,4349	0,3601
	σ_{11}	2	0,6068	0,5142	0,4965	1,1990	1,1259	1,1689
	σ_{12}	-0,5	0,9987	0,9000	0,7696	0,5090	0,4239	0,3880
	σ_{22}	1	0,8953	0,7045	0,5561	0,5310	0,4483	0,4091
	λ_1	6	5,7472	5,7034	5,6191	4,5060	4,1370	4,5579
	λ_2	4	4,0238	3,9598	3,9688	3,2610	2,9637	3,0765
	ν	3	3,3669	1,3616	1,3953	1,3200	1,0168	0,8134
100	μ_1	-2	0,9032	0,6405	0,3853	0,5042	0,3543	0,2418
	μ_2	-1	0,7165	0,5691	0,4738	0,4308	0,3365	0,2655
	σ_{11}	2	0,4670	0,4029	0,3951	1,2045	1,1643	1,1601
	σ_{12}	-0,5	0,8526	0,8051	0,7357	0,4004	0,3653	0,3579
	σ_{22}	1	0,7073	0,5878	0,4940	0,4603	0,4075	0,3902
	λ_1	6	5,6871	5,6623	5,5361	4,0411	3,6032	3,9450
	λ_2	4	4,0295	3,9922	3,9469	3,0613	2,7058	2,8847
	ν	3	1,3687	1,2982	1,4143	1,1522	0,8781	0,7218
200	μ_1	-2	0,6858	0,4965	0,3260	0,3925	0,2796	0,1966
	μ_2	-1	0,5444	0,4237	0,3217	0,3153	0,2436	0,1888
	σ_{11}	2	0,4138	0,3655	0,3647	1,2022	1,1771	1,1453
	σ_{12}	-0,5	0,7588	0,7411	0,7203	0,3781	0,3623	0,3671
	σ_{22}	1	0,5536	0,4977	0,4644	0,4322	0,3846	0,3486
	λ_1	6	5,6284	5,6147	5,5307	3,6880	3,2324	3,3708
	λ_2	4	3,9760	3,9528	3,8528	2,6540	2,2555	2,1528
	ν	3	1,4236	1,3954	1,4517	1,0951	0,8518	0,6917
1000	μ_1	-2	0,5152	0,5048	0,5016	0,1250	0,1022	0,0915
	μ_2	-1	0,2667	0,2466	0,2459	0,0871	0,0692	0,0578
	σ_{11}	2	0,6882	0,6854	0,6859	1,0574	1,0546	1,0510
	σ_{12}	-0,5	0,5934	0,5924	0,5913	0,3975	0,3964	0,3967
	σ_{22}	2	0,1647	0,1551	0,1532	0,3064	0,2998	0,2942
	λ_1	6	5,6584	5,6577	5,6575	1,3824	1,1333	1,0047
	λ_2	4	3,9214	3,9201	3,9221	0,8267	0,6633	0,5558
	ν	3	1,7329	1,7320	1,7342	0,6407	0,5320	0,4689

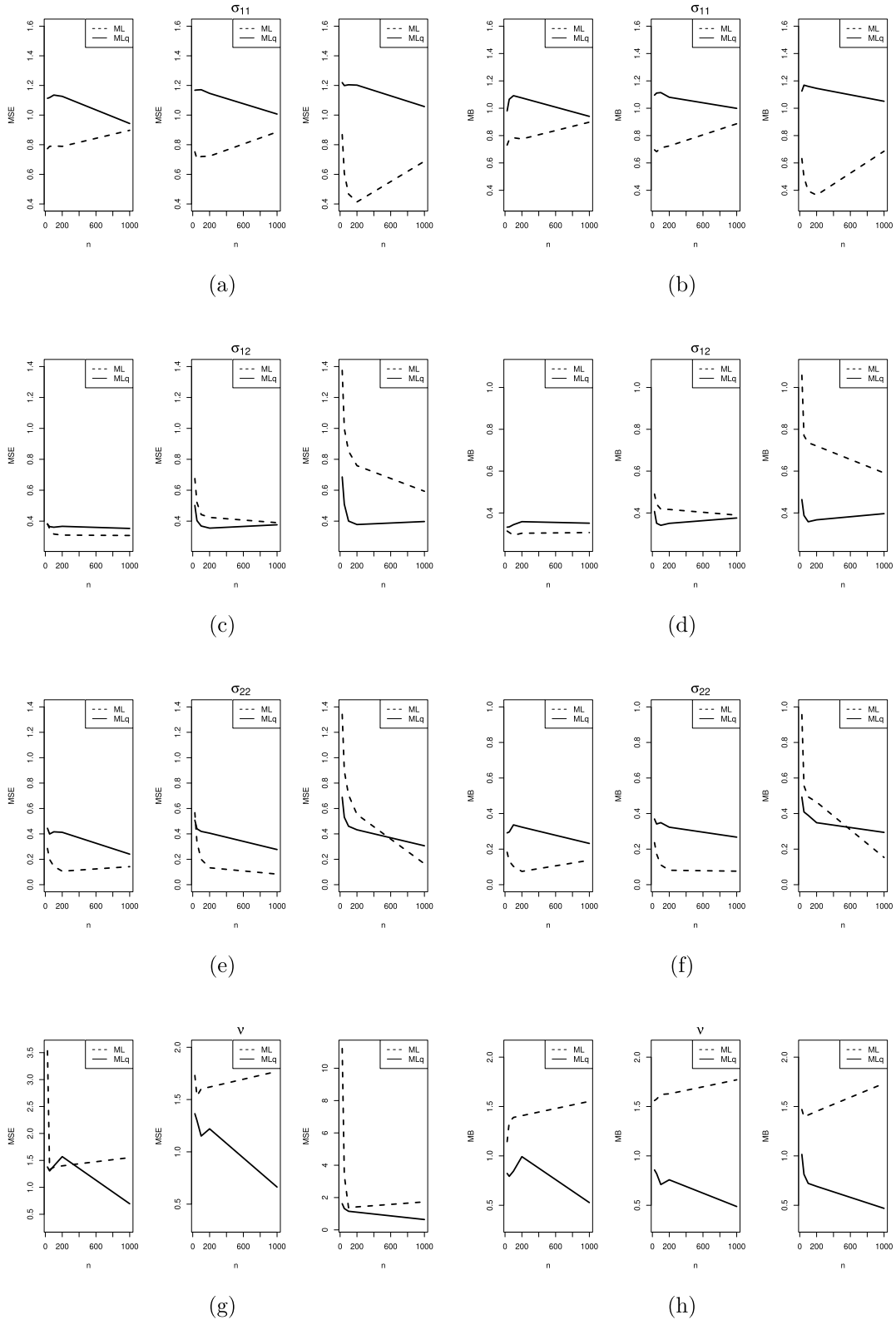
Fonte: Elaborado pela autora em 2020.

Figura 4.3 – Cenário 1: Caso 2 - MSE e MB dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros $\boldsymbol{\mu}$ e $\boldsymbol{\lambda}$ do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



Fonte: Elaborado pela autora em 2020.

Figura 4.4 – Cenário 1: Caso 2 - MSE e MB dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros Σ e ν do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



Fonte: Elaborado pela autora em 2020.

4.1.3 Cenário 2: Caso 1

As Tabelas (4.7)-(4.15) mostram as medidas Bias, MB e MSE dos estimadores de ML e de ML_q dos parâmetros do modelo ST, com o adicional de 10%, 20% e 30% de outliers para $n = 25, 50, 100, 200$ e 1000, no contexto do cenário 2 - caso 1, descrito anteriormente. Além disso, para facilitar a visualização dessas informações, apresentamos as Figuras (4.8)-(4.7).

Analisando as Figuras (4.6) (g) e (h), e (4.7), nota-se que o MSE e o MB dos estimadores dos parâmetros $\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{44}$ e λ_4 , no caso ML, considerando a porcentagem de outlier de 10%, são menores do que as respectivas medidas desses estimadores para o caso ML_q . Entretanto, com 20% e 30% de observações anômalas, destaca-se um comportamento diferente, ou seja, observa-se que o MSE e o MB dos estimadores desses parâmetros, no caso ML_q , para os diversos valores de n adotados, são menores do que as respectivas medidas desses estimadores para o caso ML. Em particular, observe que as estimativas de ML_q do λ_4 são menos voláteis para uma mudança significativa de dados nas caudas, portanto, um bom exemplo de troca (intercâmbio) de viés-variância; veja [56].

Tabela 4.7 – Cenário 2: Caso 1 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 10% de outliers, para $n = 25$.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
25	μ_1	2	0,8091	0,6155	0,4087	0,9548	0,6622	0,3573
	μ_2	1	0,6895	0,5503	0,4384	0,6278	0,5070	0,4396
	μ_3	1	0,6705	0,5559	0,4922	0,6332	0,5264	0,4515
	μ_4	1	0,6418	0,5471	0,5113	0,6354	0,5343	0,4764
	σ_{11}	1	0,3295	0,2576	0,2104	0,5239	0,4161	0,3304
	σ_{12}	0	0,3246	0,2478	0,1806	0,1955	0,1396	0,1104
	σ_{22}	1	0,3884	0,2979	0,2237	0,4957	0,3730	0,2658
	σ_{13}	0	0,3516	0,2649	0,1917	0,1992	0,1394	0,1010
	σ_{23}	0	0,3545	0,2732	0,2124	0,2022	0,1429	0,1029
	σ_{33}	1	0,4076	0,3128	0,2400	0,4974	0,3733	0,2588
	σ_{14}	0	0,3677	0,2823	0,2228	0,2045	0,1423	0,0955
	σ_{24}	0	0,3608	0,2805	0,2186	0,2053	0,1453	0,1001
	σ_{34}	0	0,3662	0,2838	0,2162	0,2110	0,1501	0,1147
	σ_{44}	1	0,4090	0,3150	0,2419	0,4944	0,3674	0,2541
	λ_1	6	5,7527	5,1166	5,3008	4,8340	4,4646	4,9151
	λ_2	3	3,4230	2,8321	2,5658	2,5129	2,1329	2,2103
	λ_3	2	2,8230	2,3939	2,0866	1,9429	1,6598	1,7128
	λ_4	1	2,3032	1,9457	1,8630	1,5119	1,2744	1,0442
	ν	5	8,5938	3,0204	1,9492	2,6002	1,9823	1,4569

Fonte: Elaborado pela autora em 2020.

Tabela 4.8 – Cenário 2: Caso 1 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 10% de outliers, para $n = 50$ e 100.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
50	μ_1	2	0,7445	0,5594	0,3510	0,7239	0,4862	0,2687
	μ_2	1	0,6169	0,4836	0,3810	0,5216	0,4111	0,3252
	μ_3	1	0,5869	0,4816	0,4237	0,5291	0,4372	0,3849
	μ_4	1	0,5708	0,4880	0,4574	0,5364	0,4483	0,4154
	σ_{11}	1	0,2408	0,1875	0,1554	0,4355	0,3590	0,3239
	σ_{12}	0	0,2393	0,1716	0,1159	0,1693	0,1243	0,0937
	σ_{22}	1	0,2833	0,2054	0,1402	0,3952	0,3016	0,2382
	σ_{13}	0	0,2569	0,1911	0,1408	0,1827	0,1319	0,0984
	σ_{23}	0	0,2690	0,2041	0,1536	0,1768	0,1280	0,0942
	σ_{33}	1	0,2948	0,2208	0,1609	0,3921	0,2973	0,2385
	σ_{14}	0	0,2761	0,2075	0,1615	0,1845	0,1294	0,0910
	σ_{24}	0	0,2791	0,2164	0,1699	0,1870	0,1366	0,1002
	σ_{34}	0	0,2752	0,2124	0,1710	0,1840	0,1322	0,0928
	σ_{44}	1	0,2903	0,2161	0,1605	0,3854	0,2877	0,2204
	λ_1	6	5,6313	5,4473	5,3745	4,8219	4,4964	4,6284
	λ_2	3	2,9953	2,7674	2,6563	2,5307	2,2002	2,1583
	λ_3	2	2,1729	1,9201	1,7206	1,8562	1,5387	1,3894
λ_4	1	1,4744	1,1979	1,1406	1,3803	1,1392	1,0000	
ν	5	2,3416	2,1884	2,3944	1,9588	1,4310	1,0200	
100	μ_1	2	0,6682	0,5130	0,3453	0,5491	0,3754	0,2031
	μ_2	1	0,5356	0,4166	0,3136	0,4547	0,3517	0,2665
	μ_3	1	0,5051	0,4146	0,3653	0,4488	0,3630	0,3123
	μ_4	1	0,4908	0,4196	0,3871	0,4431	0,3656	0,3290
	σ_{11}	1	0,1564	0,1253	0,1076	0,3935	0,3375	0,3168
	σ_{12}	0	0,1507	0,1134	0,0872	0,1307	0,1013	0,0819
	σ_{22}	1	0,1791	0,1327	0,0991	0,3255	0,2620	0,2257
	σ_{13}	0	0,1649	0,1255	0,0968	0,1259	0,0956	0,0757
	σ_{23}	0	0,1776	0,1399	0,1177	0,1226	0,0909	0,0694
	σ_{33}	1	0,1844	0,1404	0,1090	0,3142	0,2488	0,2126
	σ_{14}	0	0,1913	0,1521	0,1296	0,1289	0,0953	0,0758
	σ_{24}	0	0,1921	0,1554	0,1368	0,1303	0,0964	0,0730
	σ_{34}	0	0,1903	0,1530	0,1308	0,1307	0,0966	0,0724
	σ_{44}	1	0,1937	0,1471	0,1139	0,3103	0,2451	0,2115
	λ_1	6	5,5702	5,4916	5,3569	4,6264	4,2317	4,4688
	λ_2	3	2,8601	2,7424	2,6397	2,4919	2,1538	2,0979
	λ_3	2	2,0051	1,8577	1,7588	1,8218	1,5136	1,3059
λ_4	1	1,2015	1,0089	0,9351	1,3165	1,0700	1,0000	
ν	5	2,5001	2,4462	2,5413	1,5193	1,1625	0,9022	

Fonte: Elaborado pela autora em 2020.

Tabela 4.9 – Cenário 2: Caso 1 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 10% de outliers, para $n = 200$ e 1000.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
200	μ_1	2	0,5911	0,4655	0,2971	0,4102	0,2878	0,1905
	μ_2	1	0,4822	0,3752	0,2741	0,3805	0,2841	0,1925
	μ_3	1	0,4663	0,3887	0,3344	0,3770	0,2970	0,2326
	μ_4	1	0,4296	0,3806	0,3700	0,3583	0,2920	0,2558
	σ_{11}	1	0,1271	0,1063	0,0998	0,4045	0,3584	0,3394
	σ_{12}	0	0,1025	0,0800	0,0659	0,1060	0,0835	0,0692
	σ_{22}	1	0,1202	0,0929	0,0719	0,3334	0,2766	0,2472
	σ_{13}	0	0,1209	0,0953	0,0793	0,1033	0,0797	0,0653
	σ_{23}	0	0,1345	0,1117	0,0997	0,0933	0,0700	0,0570
	σ_{33}	1	0,1355	0,1057	0,0885	0,3267	0,2667	0,2334
	σ_{14}	0	0,1396	0,1150	0,1020	0,0986	0,0749	0,0601
	σ_{24}	0	0,1463	0,1230	0,1177	0,0969	0,0720	0,0553
	σ_{34}	0	0,1472	0,1249	0,1196	0,0945	0,0698	0,0503
	σ_{44}	1	0,1342	0,1065	0,0885	0,3241	0,2653	0,2271
	λ_1	6	5,5409	5,4945	5,2937	4,3859	3,9947	4,3937
	λ_2	3	2,8534	2,7792	2,5856	2,3829	2,0474	2,0367
	λ_3	2	1,9885	1,8882	1,7633	1,7561	1,4419	1,2706
λ_4	1	1,1294	0,9817	0,8913	1,1955	0,9809	0,8916	
ν	5	2,5793	2,5612	2,5900	1,6383	1,2944	1,0721	
1000	μ_1	2	0,5078	0,4188	0,2478	0,2920	0,2105	0,1425
	μ_2	1	0,4119	0,3075	0,1439	0,2848	0,2019	0,1217
	μ_3	1	0,3863	0,3263	0,2351	0,2783	0,2166	0,1640
	μ_4	1	0,3694	0,3422	0,3170	0,2747	0,2290	0,1998
	σ_{11}	1	0,1159	0,1000	0,0888	0,4226	0,3696	0,3521
	σ_{12}	0	0,0663	0,0567	0,0570	0,0813	0,0658	0,0559
	σ_{22}	1	0,0583	0,0474	0,0413	0,3652	0,2956	0,2136
	σ_{13}	0	0,0809	0,0699	0,0664	0,0736	0,0627	0,0571
	σ_{23}	0	0,0957	0,0837	0,0906	0,0585	0,0493	0,0445
	σ_{33}	1	0,0760	0,0664	0,0646	0,3524	0,2813	0,1935
	σ_{14}	0	0,1028	0,0906	0,0949	0,0675	0,0588	0,0568
	σ_{24}	0	0,1075	0,0962	0,1079	0,0563	0,0460	0,0426
	σ_{34}	0	0,1105	0,0995	0,1115	0,0558	0,0438	0,0361
	σ_{44}	1	0,0847	0,0750	0,0741	0,3476	0,2774	0,1882
	λ_1	6	5,5005	5,4767	5,2361	4,0764	3,7432	4,1764
	λ_2	3	2,8225	2,7804	2,4874	2,1838	1,8783	1,9621
	λ_3	2	1,9335	1,8747	1,5855	1,5924	1,2738	1,1588
λ_4	1	1,0623	0,9573	0,6806	1,0689	0,8401	0,6145	
ν	5	2,6625	2,6600	2,6681	1,8947	1,5962	1,5104	

Fonte: Elaborado pela autora em 2020.

Tabela 4.10 – Cenário 2: Caso 1 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 20% de outliers, para $n = 25$ e 50.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
25	μ_1	2	1,2637	0,9950	0,7817	1,1268	0,8360	0,5080
	μ_2	1	1,1501	0,9774	0,9122	0,7324	0,6346	0,6390
	μ_3	1	1,1454	0,9808	0,9411	0,7507	0,6532	0,6677
	μ_4	1	1,1158	0,9605	0,9441	0,7638	0,6700	0,6998
	σ_{11}	1	0,8149	0,6749	0,6405	0,6385	0,5379	0,4688
	σ_{12}	0	0,8496	0,7263	0,7067	0,3851	0,2378	0,1238
	σ_{22}	1	0,9060	0,7663	0,7338	0,6428	0,5277	0,4417
	σ_{13}	0	0,8622	0,7485	0,7379	0,3864	0,2427	0,1205
	σ_{23}	0	0,8770	0,7640	0,7585	0,4007	0,2514	0,1145
	σ_{33}	1	0,9083	0,7707	0,7491	0,6337	0,5228	0,4435
	σ_{14}	0	0,8863	0,7742	0,7544	0,3968	0,2473	0,1230
	σ_{24}	0	0,8808	0,7667	0,7417	0,4084	0,2539	0,1269
	σ_{34}	0	0,8747	0,7592	0,7298	0,3942	0,2484	0,1208
	σ_{44}	1	0,9000	0,7646	0,7442	0,6403	0,5289	0,4391
	λ_1	6	6,0356	5,7896	5,6640	5,1423	4,8936	5,4478
	λ_2	3	3,3680	3,0215	3,0817	2,6165	2,3409	2,8022
	λ_3	2	2,5977	2,2284	2,1324	1,9192	1,6743	2,0000
λ_4	1	1,8929	1,5201	1,4950	1,4073	1,1713	1,0000	
ν	5	20,2171	7,0366	2,1913	3,0064	2,4190	1,8684	
50	μ_1	2	1,1072	0,8703	0,6671	0,8614	0,6076	0,3893
	μ_2	1	1,0062	0,8461	0,7729	0,6241	0,5132	0,4463
	μ_3	1	0,9790	0,8479	0,8265	0,6258	0,5206	0,4743
	μ_4	1	0,9534	0,8353	0,8386	0,6335	0,5352	0,4989
	σ_{11}	1	0,6341	0,5277	0,4957	0,5473	0,4719	0,4374
	σ_{12}	0	0,6768	0,5906	0,5672	0,3050	0,1843	0,1008
	σ_{22}	1	0,7312	0,6331	0,6292	0,5146	0,4153	0,3328
	σ_{13}	0	0,6987	0,6175	0,5821	0,3118	0,1859	0,0942
	σ_{23}	0	0,7100	0,6269	0,5879	0,3153	0,1895	0,0984
	σ_{33}	1	0,7373	0,6414	0,6398	0,5135	0,4092	0,3313
	σ_{14}	0	0,7212	0,6448	0,6320	0,3188	0,1863	0,0856
	σ_{24}	0	0,7178	0,6396	0,6045	0,3199	0,1895	0,0905
	σ_{34}	0	0,7035	0,6256	0,5887	0,3105	0,1843	0,0896
	σ_{44}	1	0,7219	0,6331	0,6288	0,5064	0,4022	0,3160
	λ_1	6	5,9069	5,8190	5,6074	4,9915	4,6854	5,0328
	λ_2	3	3,1383	2,9916	2,9662	2,6504	2,3653	2,4889
	λ_3	2	2,2410	2,0477	2,0073	1,9628	1,6938	1,8104
λ_4	1	1,4107	1,1590	1,1906	1,4055	1,1567	1,0000	
ν	5	2,1546	1,7855	1,7734	2,2825	1,7178	1,2747	

Fonte: Elaborado pela autora em 2020.

Tabela 4.11 – Cenário 2: Caso 1 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 20% de outliers, para $n = 100$ e 200.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
100	μ_1	2	1,0117	0,7865	0,5936	0,5939	0,4140	0,2641
	μ_2	1	0,8965	0,7545	0,6864	0,4883	0,3792	0,2992
	μ_3	1	0,8789	0,7559	0,7178	0,4794	0,3813	0,3165
	μ_4	1	0,8564	0,7550	0,7380	0,4684	0,3738	0,3216
	σ_{11}	1	0,5231	0,4489	0,4223	0,4459	0,3959	0,3839
	σ_{12}	0	0,5746	0,5178	0,4901	0,1724	0,1106	0,0801
	σ_{22}	1	0,6195	0,5573	0,5231	0,3798	0,3110	0,2727
	σ_{13}	0	0,6008	0,5436	0,5030	0,1699	0,1025	0,0684
	σ_{23}	0	0,6052	0,5497	0,5129	0,1788	0,1052	0,0676
	σ_{33}	1	0,6350	0,5705	0,5394	0,3696	0,2977	0,2543
	σ_{14}	0	0,6231	0,5712	0,5276	0,1786	0,1032	0,0643
	σ_{24}	0	0,6100	0,5574	0,5187	0,1783	0,1022	0,0639
	σ_{34}	0	0,6086	0,5546	0,5111	0,1791	0,1035	0,0622
	σ_{44}	1	0,6222	0,5641	0,5302	0,3680	0,2922	0,2479
	λ_1	6	5,8450	5,7968	5,5950	4,6507	4,1603	4,7329
	λ_2	3	3,0266	2,9418	2,8081	2,5470	2,2144	2,2708
	λ_3	2	2,1078	1,9880	1,9024	1,9242	1,6362	1,5756
λ_4	1	1,2382	1,0473	0,9724	1,4109	1,1618	1,0747	
ν	5	1,9473	1,7869	1,8802	1,6958	1,2756	0,9690	
200	μ_1	2	0,9636	0,7328	0,4753	0,4466	0,3091	0,1869
	μ_2	1	0,8438	0,6949	0,6440	0,3891	0,2907	0,1975
	μ_3	1	0,8251	0,7153	0,6611	0,3893	0,3009	0,2347
	μ_4	1	0,8131	0,7262	0,6701	0,3960	0,3113	0,2500
	σ_{11}	1	0,4529	0,4046	0,3766	0,4198	0,3814	0,3783
	σ_{12}	0	0,5159	0,4769	0,4482	0,1099	0,0816	0,0697
	σ_{22}	1	0,5541	0,5134	0,4813	0,3424	0,2956	0,2732
	σ_{13}	0	0,5369	0,4998	0,4654	0,1035	0,0736	0,0576
	σ_{23}	0	0,5486	0,5115	0,4720	0,1000	0,0661	0,0499
	σ_{33}	1	0,5658	0,5268	0,5018	0,3296	0,2784	0,2474
	σ_{14}	0	0,5630	0,5270	0,4933	0,1049	0,0719	0,0565
	σ_{24}	0	0,5594	0,5241	0,4940	0,1008	0,0656	0,0469
	σ_{34}	0	0,5507	0,5155	0,4819	0,1042	0,0692	0,0518
	σ_{44}	1	0,5668	0,5286	0,5019	0,3203	0,2662	0,2300
	λ_1	6	5,8301	5,7956	5,5168	4,3631	3,8631	4,4391
	λ_2	3	3,0030	2,9408	2,6909	2,4150	2,0675	2,1047
	λ_3	2	2,1186	2,0311	1,9682	1,8206	1,5170	1,4437
λ_4	1	1,2426	1,0934	1,2256	1,3679	1,1111	1,0232	
ν	5	2,0367	1,9511	2,0619	1,4924	1,1264	0,8611	

Fonte: Elaborado pela autora em 2020.

Tabela 4.12 – Cenário 2: Caso 1 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 20% de outliers, para $n = 1000$.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
1000	μ_1	2	0,7384	0,5203	0,2473	0,2976	0,2067	0,1359
	μ_2	1	0,6565	0,5535	0,4709	0,2825	0,1996	0,1158
	μ_3	1	0,6486	0,5824	0,5403	0,2762	0,2079	0,1490
	μ_4	1	0,6441	0,6025	0,5810	0,2608	0,2098	0,1778
	σ_{11}	1	0,3498	0,3373	0,3324	0,4154	0,3730	0,3424
	σ_{12}	0	0,4177	0,4039	0,4069	0,0800	0,0647	0,0556
	σ_{22}	1	0,4484	0,4393	0,4329	0,3520	0,3009	0,2228
	σ_{13}	0	0,4468	0,4346	0,4389	0,0734	0,0596	0,0545
	σ_{23}	0	0,4510	0,4392	0,4401	0,0589	0,0468	0,0420
	σ_{33}	1	0,4634	0,4546	0,4488	0,3429	0,2891	0,2176
	σ_{14}	0	0,4730	0,4620	0,4613	0,0665	0,0555	0,0533
	σ_{24}	0	0,4610	0,4496	0,4501	0,0571	0,0433	0,0381
	σ_{34}	0	0,4617	0,4507	0,4499	0,0560	0,0410	0,0341
	σ_{44}	1	0,4638	0,4550	0,4488	0,3373	0,2824	0,2138
	λ_1	6	5,7114	5,6918	5,4241	4,0025	3,5475	4,1091
	λ_2	3	2,9470	2,9073	2,6381	2,1704	1,8159	1,9289
	λ_3	2	2,0299	1,9725	1,7067	1,6088	1,2911	1,2150
	λ_4	1	1,1308	1,0253	0,7598	1,0565	0,8537	0,7098
	ν	5	2,1841	2,1702	2,1861	1,7094	1,2734	0,9412

Fonte: Elaborado pela autora em 2020.

Tabela 4.13 – Cenário 2: Caso 1 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 30% de outliers, para $n = 25$.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
25	μ_1	2	1,7451	1,4686	1,3383	1,2117	0,9501	0,6997
	μ_2	1	1,6985	1,4849	1,4405	0,8305	0,7304	0,8065
	μ_3	1	1,6829	1,4983	1,4936	0,8505	0,7537	0,8517
	μ_4	1	1,6917	1,5141	1,5064	0,8849	0,7845	0,9358
	σ_{11}	1	1,3729	1,2735	1,2725	0,7619	0,6689	0,6657
	σ_{12}	0	1,3879	1,3030	1,3072	0,6211	0,3985	0,1751
	σ_{22}	1	1,4883	1,3858	1,3661	0,7884	0,6809	0,6721
	σ_{13}	0	1,4113	1,3254	1,3270	0,6300	0,4059	0,1866
	σ_{23}	0	1,4221	1,3354	1,3305	0,6404	0,4114	0,1919
	σ_{33}	1	1,4837	1,3830	1,3649	0,7864	0,6829	0,6887
	σ_{14}	0	1,4376	1,3577	1,3437	0,6405	0,4163	0,1887
	σ_{24}	0	1,4348	1,3498	1,3488	0,6392	0,4145	0,1885
	σ_{34}	0	1,4217	1,3382	1,3331	0,6352	0,4125	0,1932
	σ_{44}	1	1,4736	1,3726	1,3548	0,7788	0,6755	0,6783
	λ_1	6	5,9474	5,8112	5,6988	5,2483	5,0728	5,3936
	λ_2	3	3,2213	2,9771	2,8152	2,6478	2,4195	2,8935
	λ_3	2	2,3525	2,0524	2,0175	1,8951	1,6531	2,0000
	λ_4	1	1,6076	1,2692	1,2610	1,3938	1,1261	1,0000
	ν	5	34,2620	15,7818	2,3772	3,1805	2,6290	2,2225

Fonte: Elaborado pela autora em 2020.

Tabela 4.14 – Cenário 2: Caso 1 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 30% de outliers, para $n = 50$ e 100.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
50	μ_1	2	1,5539	1,3039	1,1453	0,8988	0,6452	0,4349
	μ_2	1	1,4839	1,3044	1,2993	0,6652	0,5518	0,5026
	μ_3	1	1,4556	1,2952	1,3401	0,6791	0,5594	0,4943
	μ_4	1	1,4786	1,3252	1,3762	0,7234	0,6089	0,5742
	σ_{11}	1	1,1579	1,0977	1,1067	0,6159	0,5427	0,4861
	σ_{12}	0	1,1864	1,1312	1,1297	0,4342	0,2529	0,1078
	σ_{22}	1	1,2663	1,2067	1,2035	0,6064	0,4974	0,4033
	σ_{13}	0	1,2070	1,1552	1,1541	0,4411	0,2534	0,0968
	σ_{23}	0	1,2054	1,1513	1,1469	0,4426	0,2538	0,0972
	σ_{33}	1	1,2611	1,2047	1,2043	0,6047	0,4974	0,4276
	σ_{14}	0	1,2296	1,1751	1,1766	0,4584	0,2608	0,0903
	σ_{24}	0	1,2285	1,1745	1,1694	0,4546	0,2636	0,1100
	σ_{34}	0	1,2147	1,1608	1,1527	0,4509	0,2630	0,1045
	σ_{44}	1	1,2609	1,2031	1,1978	0,6020	0,4898	0,3957
	λ_1	6	5,9180	5,8562	5,6415	5,0758	4,8104	5,1214
	λ_2	3	3,0732	2,9520	2,7292	2,6880	2,4170	2,5719
	λ_3	2	2,1528	1,9890	1,8869	1,9358	1,6608	1,7186
λ_4	1	1,3430	1,0959	1,2896	1,5131	1,2073	1,0000	
ν	5	5,2165	2,1715	1,2455	2,4899	1,9564	1,5485	
100	μ_1	2	1,4669	1,2365	1,0391	0,6204	0,4410	0,3035
	μ_2	1	1,4139	1,2581	1,2651	0,5030	0,3975	0,3278
	μ_3	1	1,3912	1,2372	1,2832	0,5138	0,4102	0,3338
	μ_4	1	1,4021	1,2682	1,3575	0,5252	0,4184	0,3501
	σ_{11}	1	1,0962	1,0632	1,0518	0,5066	0,4517	0,4196
	σ_{12}	0	1,1167	1,0828	1,0922	0,2841	0,1456	0,0777
	σ_{22}	1	1,1914	1,1602	1,1478	0,4514	0,3662	0,3032
	σ_{13}	0	1,1346	1,1020	1,1092	0,2911	0,1426	0,0696
	σ_{23}	0	1,1358	1,1021	1,1111	0,2855	0,1364	0,0648
	σ_{33}	1	1,1940	1,1621	1,1530	0,4473	0,3530	0,2881
	σ_{14}	0	1,1695	1,1366	1,1365	0,2975	0,1396	0,0633
	σ_{24}	0	1,1525	1,1180	1,1125	0,2896	0,1397	0,0648
	σ_{34}	0	1,1410	1,1072	1,1104	0,2864	0,1388	0,0657
	σ_{44}	1	1,1907	1,1591	1,1438	0,4425	0,3478	0,2817
	λ_1	6	5,8912	5,8464	5,4940	4,7763	4,3669	4,8245
	λ_2	3	3,0445	2,9581	2,6668	2,6130	2,3096	2,3592
	λ_3	2	2,0948	1,9685	1,6628	1,9639	1,6986	1,6769
λ_4	1	1,2386	1,0240	0,8242	1,4477	1,1782	1,0449	
ν	5	1,8112	1,0956	0,8773	1,8463	1,4263	1,1594	

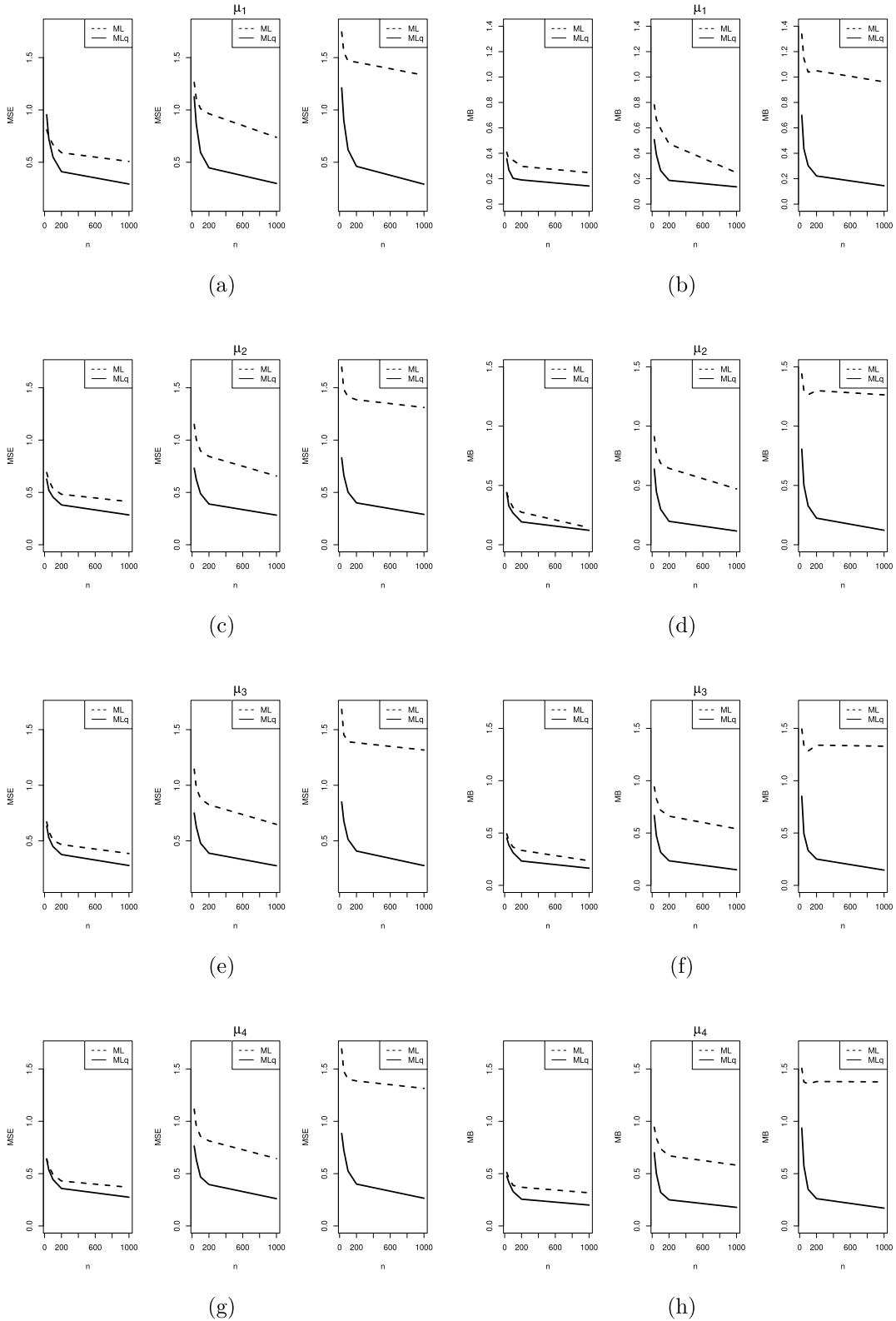
Fonte: Elaborado pela autora em 2020.

Tabela 4.15 – Cenário 2: Caso 1 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 30% de outliers, para $n = 200$ e 1000.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
200	μ_1	2	1,4566	1,2386	1,0492	0,4611	0,3281	0,2217
	μ_2	1	1,3855	1,2229	1,2993	0,4000	0,3046	0,2240
	μ_3	1	1,3827	1,2433	1,3395	0,4086	0,3152	0,2510
	μ_4	1	1,3856	1,2546	1,3800	0,4004	0,3152	0,2609
	σ_{11}	1	1,0610	1,0391	1,0456	0,4438	0,4082	0,3924
	σ_{12}	0	1,0798	1,0528	1,0790	0,1557	0,0860	0,0644
	σ_{22}	1	1,1484	1,1281	1,1189	0,3716	0,3209	0,2840
	σ_{13}	0	1,1111	1,0860	1,1061	0,1546	0,0795	0,0560
	σ_{23}	0	1,1058	1,0798	1,1005	0,1514	0,0743	0,0511
	σ_{33}	1	1,1647	1,1447	1,1479	0,3538	0,2955	0,2604
	σ_{14}	0	1,1371	1,1120	1,1268	0,1551	0,0772	0,0517
	σ_{24}	0	1,1149	1,0896	1,1074	0,1503	0,0734	0,0480
	σ_{34}	0	1,1099	1,0835	1,1091	0,1491	0,0758	0,0507
	σ_{44}	1	1,1589	1,1387	1,1432	0,3488	0,2900	0,2456
	λ_1	6	5,8642	5,8269	5,4073	4,4855	4,0182	4,6418
	λ_2	3	3,0412	2,9692	2,5724	2,4817	2,1639	2,2085
	λ_3	2	2,0904	1,9847	1,5815	1,8646	1,5768	1,4636
λ_4	1	1,2041	1,0102	0,6240	1,3161	1,0851	0,9888	
ν	5	0,9231	0,7621	0,6961	1,5352	1,1599	0,9042	
1000	μ_1	2	1,3316	1,1250	0,9620	0,2906	0,2113	0,1446
	μ_2	1	1,3118	1,1663	1,2635	0,2898	0,2075	0,1222
	μ_3	1	1,3169	1,1942	1,3296	0,2775	0,2093	0,1457
	μ_4	1	1,3145	1,1881	1,3768	0,2650	0,2105	0,1702
	σ_{11}	1	1,0172	1,0079	1,0201	0,4265	0,3940	0,3725
	σ_{12}	0	1,0357	1,0188	1,0734	0,0743	0,0615	0,0536
	σ_{22}	1	1,1078	1,0995	1,1108	0,3639	0,3222	0,2519
	σ_{13}	0	1,0658	1,0501	1,1030	0,0637	0,0529	0,0489
	σ_{23}	0	1,0657	1,0500	1,1038	0,0497	0,0416	0,0378
	σ_{33}	1	1,1200	1,1120	1,1281	0,3512	0,3070	0,2396
	σ_{14}	0	1,0862	1,0705	1,1248	0,0607	0,0517	0,0495
	σ_{24}	0	1,0687	1,0530	1,1032	0,0504	0,0416	0,0377
	σ_{34}	0	1,0631	1,0470	1,1012	0,0496	0,0405	0,0357
	σ_{44}	1	1,1093	1,1012	1,1147	0,3406	0,2965	0,2391
	λ_1	6	5,7725	5,7425	5,3486	4,0754	3,6141	4,2909
	λ_2	3	2,9870	2,9259	2,4670	2,2392	1,8825	2,0312
	λ_3	2	2,0589	1,9693	1,5041	1,6334	1,3304	1,2543
λ_4	1	1,1828	1,0201	0,5495	1,0979	0,8881	0,7549	
ν	5	0,7152	0,6429	0,6547	1,6364	1,2187	0,9231	

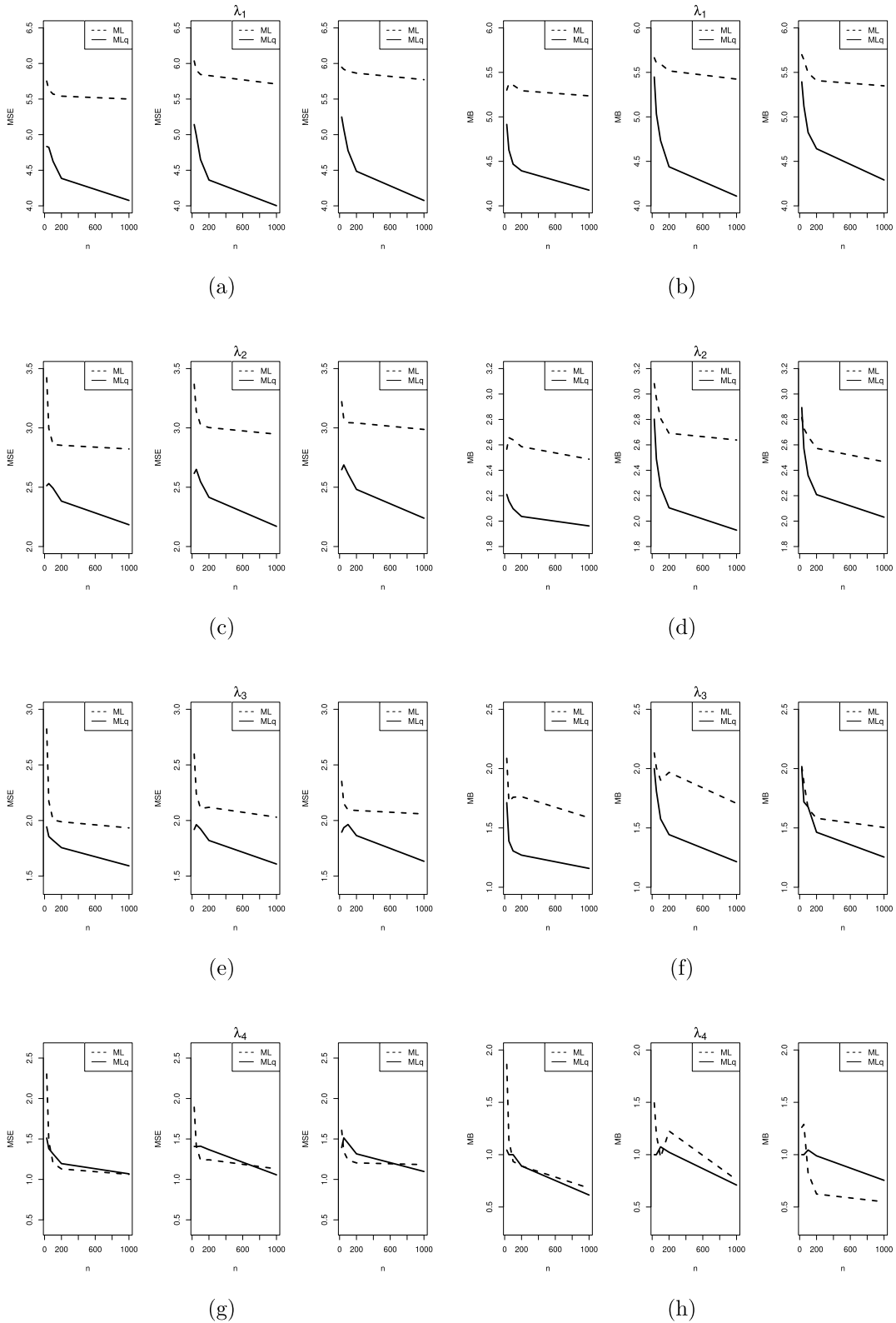
Fonte: Elaborado pela autora em 2020.

Figura 4.5 – Cenário 2: Caso 1 - MSE e MB dos estimadores de ML ($q = 1$) e de ML_q do parâmetro μ do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



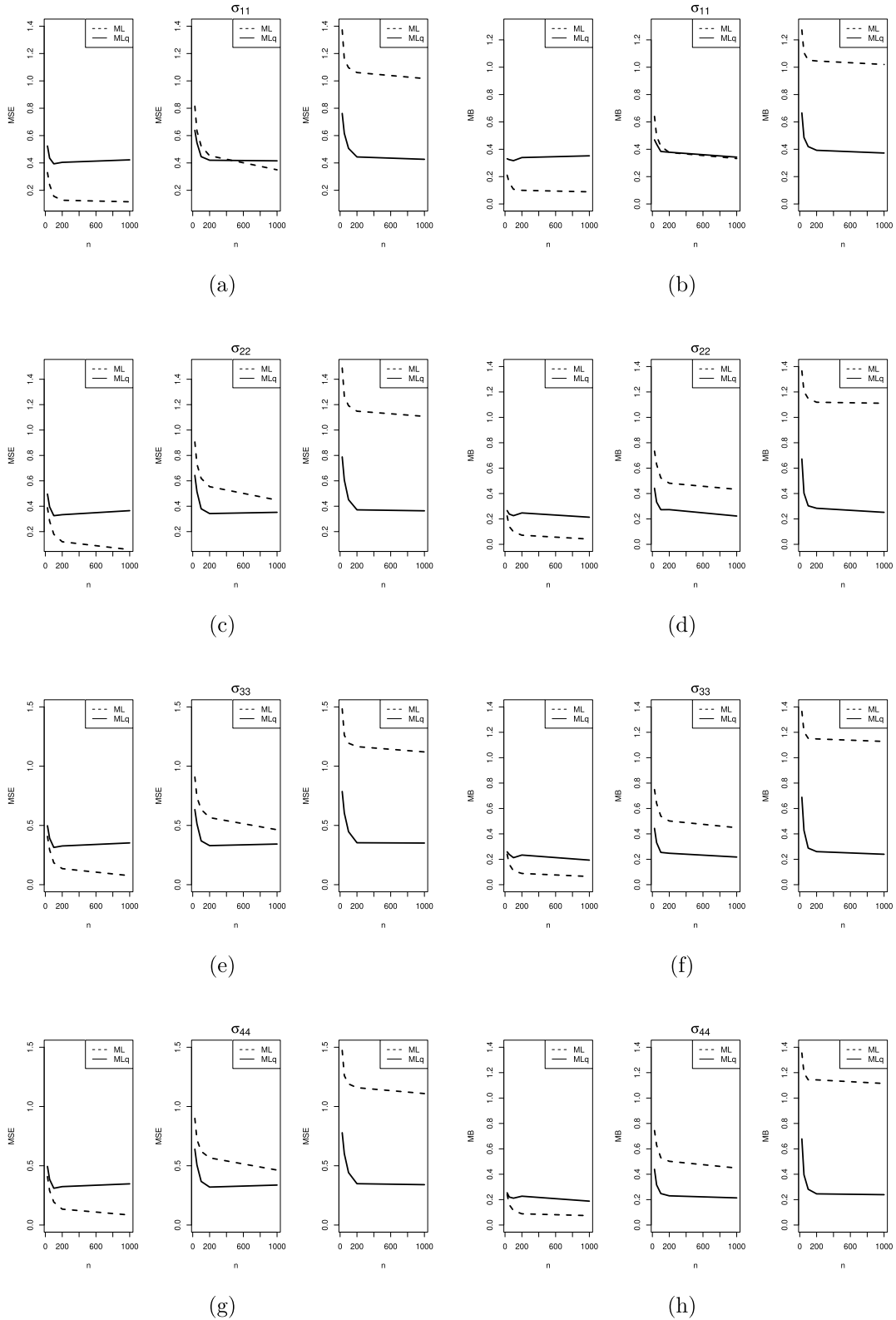
Fonte: Elaborado pela autora em 2020.

Figura 4.6 – Cenário 2: Caso 1 - MSE e MB dos estimadores de ML ($q = 1$) e de ML_q do parâmetro λ do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



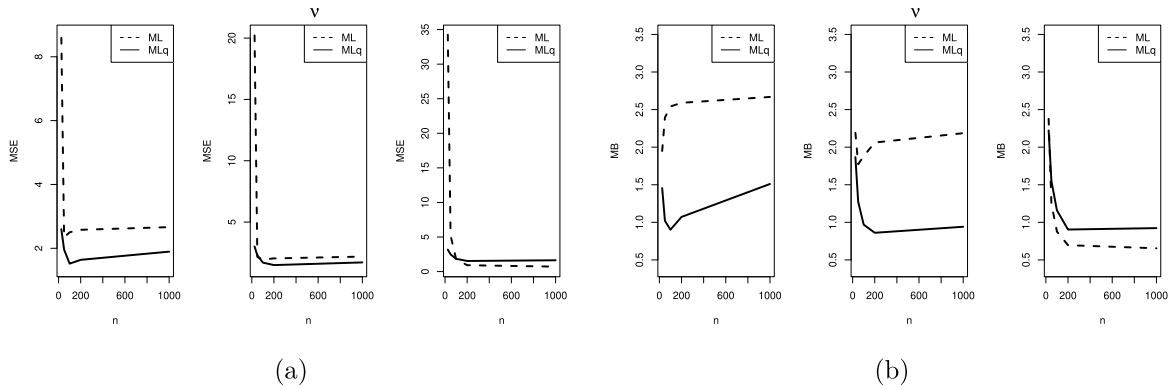
Fonte: Elaborado pela autora em 2020.

Figura 4.7 – Cenário 2: Caso 1 - MSE e MB dos estimadores de ML ($q = 1$) e de ML_q do parâmetro Σ do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



Fonte: Elaborado pela autora em 2020.

Figura 4.8 – Cenário 2: Caso 1 - MSE e MB dos estimadores de ML ($q = 1$) e de ML_q do parâmetro ν do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



Fonte: Elaborado pela autora em 2020.

4.1.4 Cenário 2: Caso 2

As Tabelas (4.16)-(4.24) mostram as medidas Bias, MB e MSE dos estimadores de ML e de ML_q dos parâmetros do modelo ST, com o adicional de 10%, 20% e 30% de outliers para $n = 25, 50, 100, 200$ e 1000, no contexto do cenário 2 - caso 2, descrito anteriormente. Além disso, para facilitar a visualização dessas informações, apresentamos as Figuras (4.10)-(4.9).

Analisando as Figuras (4.6) (g) e (h), e (4.7), nota-se que o MSE e o MB dos estimadores dos parâmetros $\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{44}, \lambda_3$ e λ_4 , no caso ML, considerando todas as porcentagens de outliers adotadas, são menores do que as respectivas medidas desses estimadores para o caso ML_q . Entretanto, em particular, observe que as estimativas de ML_q dos parâmetros $\sigma_{11}, \sigma_{22}, \sigma_{33}$ e σ_{44} são menos voláteis para uma mudança significativa de dados nas caudas, portanto, um bom exemplo de troca (intercâmbio) de viés-variância; veja [56].

Tabela 4.16 – Cenário 2: Caso 2 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 10% de outliers, para $n = 25$ e 50.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
25	μ_1	-2	0,9415	0,7355	0,5364	0,8795	0,6738	0,4991
	μ_2	-1	0,7772	0,6195	0,5229	0,7176	0,5746	0,4890
	μ_3	-1	0,8094	0,6594	0,5750	0,7821	0,6496	0,5861
	μ_4	-1	0,7618	0,6297	0,5800	0,7523	0,6281	0,5730
	σ_{11}	2	0,7441	0,6760	0,6850	1,0708	0,9994	0,9741
	σ_{12}	-0,5	0,5107	0,4258	0,3701	0,4045	0,3447	0,3297
	σ_{22}	2	0,5678	0,4925	0,4891	0,9155	0,8098	0,7667
	σ_{13}	-0,5	0,7367	0,6723	0,6414	0,5894	0,5433	0,5174
	σ_{23}	-0,5	0,6548	0,5839	0,5548	0,5255	0,4702	0,4701
	σ_{33}	2	0,5857	0,5094	0,5134	0,9210	0,8154	0,7826
	σ_{14}	-0,5	0,6861	0,6144	0,5607	0,5461	0,4931	0,4944
	σ_{24}	-0,5	0,6356	0,5614	0,5303	0,5160	0,4626	0,4633
	σ_{34}	-0,5	0,5596	0,4772	0,4373	0,4529	0,3938	0,3802
	σ_{44}	2	0,5478	0,4695	0,4653	0,8888	0,7753	0,7241
	λ_1	6	5,5732	5,0279	5,0839	4,9881	4,6699	4,8117
	λ_2	3	3,3061	2,7568	2,4743	2,7526	2,3760	2,2198
λ_3	-2	2,7079	2,1888	1,7587	2,3326	1,9449	1,7396	
λ_4	-1	2,2782	1,7995	1,5243	1,8338	1,4518	1,1997	
ν	5	8,7158	2,9245	2,1518	1,9836	1,5313	1,2375	
50	μ_1	-2	0,8794	0,6980	0,5318	0,7710	0,5982	0,4666
	μ_2	-1	0,6806	0,5535	0,4684	0,6321	0,5140	0,4484
	μ_3	-1	0,6726	0,5417	0,4570	0,6571	0,5297	0,4389
	μ_4	-1	0,6509	0,5398	0,4842	0,6309	0,5185	0,4635
	σ_{11}	2	0,7822	0,7404	0,7784	1,0644	1,0225	1,0368
	σ_{12}	-0,5	0,4146	0,3576	0,3226	0,3524	0,3014	0,2750
	σ_{22}	2	0,5893	0,5379	0,5581	0,8935	0,8258	0,8335
	σ_{13}	-0,5	0,6559	0,6166	0,5870	0,5520	0,5162	0,5050
	σ_{23}	-0,5	0,5678	0,5251	0,4913	0,4813	0,4415	0,4282
	σ_{33}	2	0,5939	0,5399	0,5703	0,8862	0,8228	0,8221
	σ_{14}	-0,5	0,6084	0,5685	0,5387	0,5107	0,4748	0,4677
	σ_{24}	-0,5	0,5560	0,5161	0,4815	0,4720	0,4345	0,4220
	σ_{34}	-0,5	0,4757	0,4245	0,3958	0,4091	0,3646	0,3579
	σ_{44}	2	0,5590	0,5076	0,5383	0,8695	0,7979	0,8048
	λ_1	6	5,5743	5,4093	5,2188	5,0693	4,8018	4,8806
	λ_2	3	3,0166	2,7848	2,7103	2,7760	2,4608	2,3791
λ_3	-2	1,9933	1,7211	1,5436	2,0367	1,7267	1,4587	
λ_4	-1	1,3692	1,0656	0,8559	1,4891	1,1607	0,9659	
ν	5	2,5336	2,3848	2,6271	1,5606	1,2227	0,9989	

Fonte: Elaborado pela autora em 2020.

Tabela 4.17 – Cenário 2: Caso 2 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 10% de outliers, para $n = 100$ e 200.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
100	μ_1	-2	0,7523	0,5937	0,4272	0,6359	0,4733	0,3235
	μ_2	-1	0,6007	0,4801	0,4061	0,5604	0,4450	0,3719
	μ_3	-1	0,5714	0,4526	0,3694	0,5573	0,4448	0,3749
	μ_4	-1	0,5474	0,4599	0,4230	0,5380	0,4388	0,3895
	σ_{11}	2	0,8013	0,7822	0,8010	1,0958	1,0661	1,0670
	σ_{12}	-0,5	0,3443	0,3133	0,3082	0,3188	0,2847	0,2882
	σ_{22}	2	0,5914	0,5646	0,5830	0,9382	0,8880	0,8917
	σ_{13}	-0,5	0,5844	0,5642	0,5661	0,5101	0,4856	0,4835
	σ_{23}	-0,5	0,5076	0,4860	0,4747	0,4508	0,4261	0,4262
	σ_{33}	2	0,6048	0,5805	0,5897	0,9407	0,8926	0,8891
	σ_{14}	-0,5	0,5386	0,5158	0,5015	0,4715	0,4454	0,4390
	σ_{24}	-0,5	0,4931	0,4707	0,4608	0,4388	0,4138	0,4113
	σ_{34}	-0,5	0,4120	0,3878	0,3799	0,3794	0,3505	0,3443
	σ_{44}	2	0,5710	0,5446	0,5569	0,9209	0,8679	0,8613
	λ_1	6	5,3961	5,3164	5,1425	4,7620	4,4944	4,5523
	λ_2	3	2,8212	2,6938	2,5690	2,6008	2,3211	2,1873
	λ_3	-2	1,8373	1,6865	1,5631	1,8805	1,6247	1,3538
λ_4	-1	1,1032	0,8715	0,6602	1,3244	1,0208	0,7802	
ν	5	2,6398	2,5929	2,7050	1,4624	1,1628	0,9597	
200	μ_1	-2	0,6885	0,5609	0,4157	0,5551	0,4187	0,3010
	μ_2	-1	0,5375	0,4236	0,3419	0,4877	0,3784	0,2863
	μ_3	-1	0,4906	0,3815	0,2867	0,4783	0,3734	0,3035
	μ_4	-1	0,4920	0,4164	0,3842	0,4846	0,4022	0,3688
	σ_{11}	2	0,8165	0,8067	0,8099	1,1337	1,1075	1,1021
	σ_{12}	-0,5	0,3134	0,2924	0,2956	0,3175	0,2906	0,3000
	σ_{22}	2	0,6051	0,5933	0,6004	0,9985	0,9520	0,9513
	σ_{13}	-0,5	0,5441	0,5309	0,5392	0,4793	0,4594	0,4544
	σ_{23}	-0,5	0,4730	0,4607	0,4639	0,4296	0,4129	0,4149
	σ_{33}	2	0,6114	0,5986	0,6065	0,9902	0,9468	0,9435
	σ_{14}	-0,5	0,4974	0,4851	0,4932	0,4471	0,4290	0,4293
	σ_{24}	-0,5	0,4542	0,4418	0,4424	0,4130	0,3967	0,4019
	σ_{34}	-0,5	0,3838	0,3684	0,3668	0,3621	0,3425	0,3499
	σ_{44}	2	0,5813	0,5681	0,5724	0,9755	0,9270	0,9150
	λ_1	6	5,3730	5,3263	5,0990	4,6074	4,3751	4,5291
	λ_2	3	2,7882	2,7103	2,4793	2,4607	2,2128	2,1182
	λ_3	-2	1,7128	1,6140	1,4002	1,7727	1,5451	1,2662
λ_4	-1	0,9753	0,7983	0,5872	1,2385	0,9664	0,7620	
ν	5	2,7186	2,7036	2,7408	1,4794	1,1599	0,9518	

Fonte: Elaborado pela autora em 2020.

Tabela 4.18 – Cenário 2: Caso 2 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 10% de outliers, para $n = 1000$.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	MB	MSE_q	$Bias_q$	MB_q
1000	μ_1	-2	0,4851	0,4041	0,3441	0,3596	0,2837	0,2695
	μ_2	-1	0,4316	0,2997	0,1164	0,3713	0,2755	0,1752
	μ_3	-1	0,3938	0,2831	0,1564	0,3796	0,2735	0,1820
	μ_4	-1	0,4078	0,3528	0,3174	0,3854	0,3153	0,2542
	σ_{11}	2	0,7972	0,7940	0,7838	1,1980	1,1681	1,0803
	σ_{12}	-0,5	0,2929	0,2815	0,3099	0,3289	0,3070	0,3182
	σ_{22}	2	0,6099	0,6072	0,6067	1,1065	1,0517	0,9785
	σ_{13}	-0,5	0,4806	0,4726	0,4349	0,4288	0,4167	0,4014
	σ_{23}	-0,5	0,4306	0,4239	0,4415	0,3984	0,3876	0,4047
	σ_{33}	2	0,6088	0,6039	0,6010	1,1001	1,0457	0,9307
	σ_{14}	-0,5	0,4448	0,4362	0,4476	0,4071	0,3936	0,3885
	σ_{24}	-0,5	0,4120	0,4059	0,4275	0,3907	0,3802	0,3993
	σ_{34}	-0,5	0,3658	0,3583	0,3683	0,3661	0,3528	0,3682
	σ_{44}	2	0,5788	0,5751	0,5785	1,0889	1,0308	0,9118
	λ_1	6	5,1315	5,1154	4,9770	4,1410	3,9052	4,4678
	λ_2	3	2,6536	2,6118	2,3534	2,2253	2,0160	2,0095
	λ_3	-2	1,5707	1,5043	1,2904	1,5875	1,3859	1,1460
	λ_4	-1	0,8308	0,6857	0,4325	1,1088	0,8303	0,5150
	ν	5	2,7761	2,7744	2,7786	1,5519	1,1830	0,9523

Fonte: Elaborado pela autora em 2020.

Tabela 4.19 – Cenário 2: Caso 2 - Bias, VMD e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 20% de outliers, para $n = 25$.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	VMD	MSE_q	$Bias_q$	VMD_q
25	μ_1	-2	1,2972	0,9936	0,7327	1,0787	0,8303	0,6412
	μ_2	-1	1,1157	0,9366	0,8663	0,8515	0,7166	0,6874
	μ_3	-1	1,1580	0,9593	0,8636	0,9383	0,7901	0,7668
	μ_4	-1	1,1067	0,9249	0,8379	0,8962	0,7595	0,7478
	σ_{11}	2	0,5602	0,4629	0,4284	1,0624	0,9127	0,9165
	σ_{12}	-0,5	1,0197	0,9090	0,8819	0,6624	0,5456	0,5000
	σ_{22}	2	0,5625	0,4558	0,4041	0,9538	0,7824	0,7118
	σ_{13}	-0,5	1,2868	1,1942	1,1739	0,8636	0,7590	0,6352
	σ_{23}	-0,5	1,1677	1,0693	1,0429	0,7816	0,6728	0,5548
	σ_{33}	2	0,5535	0,4493	0,3944	0,9586	0,7853	0,6946
	σ_{14}	-0,5	1,2224	1,1274	1,0858	0,8164	0,7096	0,5811
	σ_{24}	-0,5	1,1667	1,0691	1,0203	0,7789	0,6699	0,5551
	σ_{34}	-0,5	1,0564	0,9477	0,8877	0,7015	0,5916	0,5000
	σ_{44}	2	0,5742	0,4645	0,4136	0,9555	0,7804	0,6940
	λ_1	6	5,8320	5,5767	5,3995	5,2979	5,0690	5,2079
	λ_2	3	3,2478	2,9107	2,7178	2,8259	2,5474	2,5491
	λ_3	-2	2,3191	1,9220	1,6121	2,2888	1,9465	2,0000
	λ_4	-1	1,7613	1,3783	1,2264	1,7218	1,3525	1,0162
	ν	5	17,9053	6,0972	2,2778	2,2904	1,7885	1,4287

Fonte: Elaborado pela autora em 2020.

Tabela 4.20 – Cenário 2: Caso 2 - Bias, VMD e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 20% de outliers, para $n = 50$ e 100.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	VMD	MSE_q	$Bias_q$	VMD_q
50	μ_1	-2	1,1647	0,8850	0,5993	0,8915	0,6680	0,4770
	μ_2	-1	1,0194	0,8545	0,8083	0,7804	0,6374	0,5697
	μ_3	-1	0,9823	0,8028	0,6846	0,7329	0,5868	0,4989
	μ_4	-1	0,9375	0,8041	0,7637	0,7023	0,5805	0,5262
	σ_{11}	2	0,4682	0,3915	0,3555	1,0728	0,9881	1,0691
	σ_{12}	-0,5	0,8707	0,7989	0,7579	0,5245	0,4158	0,3311
	σ_{22}	2	0,4059	0,3346	0,3047	0,9275	0,8293	0,8754
	σ_{13}	-0,5	1,1212	1,0640	1,0252	0,7119	0,6192	0,5115
	σ_{23}	-0,5	1,0339	0,9690	0,9239	0,6487	0,5548	0,4687
	σ_{33}	2	0,3995	0,3280	0,2898	0,9246	0,8278	0,8691
	σ_{14}	-0,5	1,0636	1,0061	0,9701	0,6649	0,5752	0,4904
	σ_{24}	-0,5	1,0109	0,9499	0,9126	0,6340	0,5429	0,4538
	σ_{34}	-0,5	0,8983	0,8324	0,7963	0,5641	0,4671	0,3888
	σ_{44}	2	0,3942	0,3251	0,2994	0,9215	0,8243	0,8637
	λ_1	6	5,7308	5,6385	5,3806	5,0648	4,7518	4,9339
	λ_2	3	3,0800	2,9251	2,8969	2,8285	2,5384	2,5018
	λ_3	-2	1,9212	1,7052	1,4643	2,0199	1,7188	1,5075
λ_4	-1	1,2052	0,9493	0,7679	1,4514	1,1201	0,8939	
ν	5	2,5096	1,9427	2,0082	1,6918	1,3123	1,0276	
100	μ_1	-2	1,0187	0,7568	0,4356	0,6748	0,4997	0,3531
	μ_2	-1	0,8656	0,7239	0,6547	0,6096	0,4825	0,3815
	μ_3	-1	0,8783	0,7196	0,6458	0,5517	0,4396	0,3681
	μ_4	-1	0,8339	0,7174	0,6719	0,5425	0,4407	0,3783
	σ_{11}	2	0,4351	0,3752	0,3787	1,1313	1,0973	1,1254
	σ_{12}	-0,5	0,7600	0,7152	0,6748	0,3618	0,3091	0,2808
	σ_{22}	2	0,3227	0,2702	0,2539	0,9892	0,9434	0,9698
	σ_{13}	-0,5	1,0162	0,9768	0,9173	0,5366	0,4965	0,4664
	σ_{23}	-0,5	0,9192	0,8777	0,8245	0,4749	0,4333	0,3996
	σ_{33}	2	0,3399	0,2878	0,2717	0,9960	0,9498	0,9712
	σ_{14}	-0,5	0,9564	0,9165	0,8645	0,4996	0,4557	0,4310
	σ_{24}	-0,5	0,8946	0,8546	0,8098	0,4632	0,4197	0,3906
	σ_{34}	-0,5	0,7977	0,7534	0,7038	0,4129	0,3671	0,3377
	σ_{44}	2	0,3114	0,2613	0,2398	0,9658	0,9206	0,9378
	λ_1	6	5,6105	5,5612	5,3041	4,6503	4,2859	4,5472
	λ_2	3	2,9324	2,8463	2,6920	2,5946	2,3142	2,2614
	λ_3	-2	1,8588	1,7240	1,4985	1,8223	1,5650	1,4047
λ_4	-1	1,0635	0,8395	0,6043	1,2799	0,9778	0,7467	
ν	5	2,1647	2,0270	2,2190	1,4284	1,1390	0,9935	

Fonte: Elaborado pela autora em 2020.

Tabela 4.21 – Cenário 2: Caso 2 - Bias, VMD e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 20% de outliers, para $n = 200$ e 1000.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	VMD	MSE_q	$Bias_q$	VMD_q
200	μ_1	-2	0,9588	0,7059	0,3825	0,5447	0,4012	0,2829
	μ_2	-1	0,7802	0,6310	0,5676	0,4770	0,3645	0,2756
	μ_3	-1	0,7334	0,5891	0,4871	0,4331	0,3426	0,2780
	μ_4	-1	0,7344	0,6424	0,6131	0,4484	0,3654	0,3166
	σ_{11}	2	0,4242	0,3849	0,3988	1,1690	1,1468	1,1487
	σ_{12}	-0,5	0,7036	0,6765	0,6456	0,3285	0,3012	0,3008
	σ_{22}	2	0,2778	0,2375	0,2252	1,0371	1,0055	0,9921
	σ_{13}	-0,5	0,9363	0,9116	0,8800	0,4677	0,4469	0,4380
	σ_{23}	-0,5	0,8565	0,8311	0,7923	0,4283	0,4091	0,3983
	σ_{33}	2	0,2974	0,2553	0,2444	1,0317	1,0002	0,9939
	σ_{14}	-0,5	0,8938	0,8681	0,8407	0,4420	0,4214	0,4134
	σ_{24}	-0,5	0,8349	0,8100	0,7708	0,4147	0,3956	0,3817
	σ_{34}	-0,5	0,7425	0,7145	0,6783	0,3720	0,3481	0,3341
	σ_{44}	2	0,2682	0,2268	0,2180	1,0077	0,9750	0,9571
	λ_1	6	5,5802	5,5427	5,2589	4,4820	4,1546	4,4360
	λ_2	3	2,8888	2,8251	2,5657	2,4176	2,1414	2,0920
	λ_3	-2	1,7212	1,6205	1,3691	1,6085	1,4043	1,2737
λ_4	-1	0,9486	0,7535	0,4853	1,1516	0,9015	0,7359	
ν	5	2,2757	2,2113	2,3369	1,3925	1,0576	0,8185	
1000	μ_1	-2	0,6817	0,4970	0,2876	0,3767	0,2824	0,2204
	μ_2	-1	0,5915	0,4691	0,3089	0,3558	0,2586	0,1684
	μ_3	-1	0,5876	0,4408	0,2447	0,3178	0,2464	0,1981
	μ_4	-1	0,5822	0,5146	0,4133	0,3293	0,2723	0,2457
	σ_{11}	2	0,4282	0,4169	0,4098	1,1868	1,1659	1,1071
	σ_{12}	-0,5	0,6351	0,6260	0,6455	0,3249	0,3070	0,3174
	σ_{22}	2	0,2435	0,2274	0,2284	1,0871	1,0545	0,9519
	σ_{13}	-0,5	0,8394	0,8309	0,8334	0,4281	0,4197	0,4138
	σ_{23}	-0,5	0,7696	0,7618	0,7726	0,3987	0,3908	0,3963
	σ_{33}	2	0,2771	0,2603	0,2585	1,0822	1,0504	0,9647
	σ_{14}	-0,5	0,8002	0,7912	0,7958	0,4066	0,3976	0,4015
	σ_{24}	-0,5	0,7477	0,7399	0,7453	0,3864	0,3788	0,3841
	σ_{34}	-0,5	0,6743	0,6660	0,6805	0,3616	0,3514	0,3501
	σ_{44}	2	0,2397	0,2211	0,2237	1,0677	1,0343	0,9446
	λ_1	6	5,3838	5,3650	5,1770	4,1721	3,9283	4,3520
	λ_2	3	2,7669	2,7267	2,4734	2,1954	1,9813	1,9805
	λ_3	-2	1,6399	1,5682	1,3311	1,5023	1,3589	1,2037
λ_4	-1	0,8494	0,6985	0,4376	0,9790	0,7751	0,5663	
ν	5	2,3976	2,3898	2,3950	1,5160	1,1578	0,9442	

Fonte: Elaborado pela autora em 2020.

Tabela 4.22 – Cenário 2: Caso 2 - Bias, VMD e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 30% de outliers, para $n = 25$ e 50.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	VMD	MSE_q	$Bias_q$	VMD_q
25	μ_1	-2	1,7189	1,3681	1,0990	1,2009	0,9533	0,7583
	μ_2	-1	1,6471	1,4060	1,3442	0,9926	0,8459	0,8772
	μ_3	-1	1,6612	1,4087	1,3188	1,0664	0,9006	0,9228
	μ_4	-1	1,6289	1,4075	1,4061	1,0249	0,8772	0,9500
	σ_{11}	2	0,7732	0,6336	0,5306	1,0951	0,8788	0,7152
	σ_{12}	-0,5	1,5809	1,4966	1,4809	0,9089	0,7688	0,6253
	σ_{22}	2	0,9937	0,8234	0,7328	1,0471	0,8234	0,6700
	σ_{13}	-0,5	1,8388	1,7627	1,7481	1,0836	0,9445	0,8336
	σ_{23}	-0,5	1,7490	1,6701	1,6755	1,0223	0,8822	0,7504
	σ_{33}	2	0,8927	0,7381	0,6397	1,0395	0,8147	0,6310
	σ_{14}	-0,5	1,7895	1,7180	1,7307	1,0566	0,9176	0,7939
	σ_{24}	-0,5	1,7493	1,6657	1,6363	1,0240	0,8804	0,7556
	σ_{34}	-0,5	1,6012	1,5231	1,5254	0,9302	0,7987	0,6595
	σ_{44}	2	0,9358	0,7753	0,6920	1,0437	0,8158	0,6317
	λ_1	6	5,7803	5,6370	5,3820	5,3336	5,1370	5,2853
	λ_2	3	3,1591	2,8986	2,7035	2,8672	2,6101	2,9035
	λ_3	-2	2,1887	1,8773	1,6577	2,3283	2,0369	2,0000
λ_4	-1	1,5053	1,1978	1,2200	1,6666	1,3377	1,0000	
ν	5	34,0451	15,7331	2,3300	2,6153	2,0653	1,6506	
50	μ_1	-2	1,5278	1,2178	0,9441	0,9253	0,6922	0,4846
	μ_2	-1	1,3916	1,1893	1,1608	0,8184	0,6621	0,5604
	μ_3	-1	1,4121	1,2116	1,1804	0,8372	0,6723	0,5778
	μ_4	-1	1,3919	1,2069	1,1832	0,8567	0,6925	0,5969
	σ_{11}	2	0,5115	0,4135	0,3491	1,1166	1,0006	1,1242
	σ_{12}	-0,5	1,3718	1,3177	1,3079	0,6615	0,5086	0,3760
	σ_{22}	2	0,7085	0,5960	0,5429	1,0036	0,8881	0,9269
	σ_{13}	-0,5	1,6392	1,5960	1,5967	0,8464	0,7164	0,5396
	σ_{23}	-0,5	1,5264	1,4782	1,4751	0,7765	0,6399	0,4925
	σ_{33}	2	0,6350	0,5273	0,4941	1,0087	0,8899	0,9305
	σ_{14}	-0,5	1,5792	1,5300	1,5251	0,8033	0,6633	0,5000
	σ_{24}	-0,5	1,5273	1,4744	1,4705	0,7800	0,6401	0,5000
	σ_{34}	-0,5	1,3973	1,3453	1,3395	0,7071	0,5698	0,4408
	σ_{44}	2	0,6833	0,5719	0,5133	0,9916	0,8747	0,9138
	λ_1	6	5,7578	5,6921	5,3707	5,1031	4,8595	4,9925
	λ_2	3	3,0118	2,8854	2,6961	2,8448	2,5521	2,4714
	λ_3	-2	1,9926	1,8119	1,5506	2,0841	1,7883	1,6289
λ_4	-1	1,1997	0,9269	0,5995	1,5809	1,2219	1,0000	
ν	5	5,6083	2,1670	1,2988	2,0597	1,6196	1,3132	

Fonte: Elaborado pela autora em 2020.

Tabela 4.23 – Cenário 2: Caso 2 - Bias, VMD e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 30% de outliers, para $n = 100$ e 200.

n	Parâmetro	Valor Real	ML			ML_q		
			MSE	Bias	VMD	MSE_q	$Bias_q$	VMD_q
100	μ_1	-2	1,4060	1,1082	0,8341	0,6891	0,5144	0,3731
	μ_2	-1	1,2871	1,1148	1,0696	0,6414	0,5006	0,4060
	μ_3	-1	1,3442	1,1416	1,1059	0,6275	0,4894	0,3977
	μ_4	-1	1,3110	1,1499	1,1709	0,6231	0,4903	0,4183
	σ_{11}	2	0,4179	0,3436	0,3050	1,1552	1,1122	1,1624
	σ_{12}	-0,5	1,2989	1,2634	1,2589	0,4477	0,3503	0,2934
	σ_{22}	2	0,5976	0,5245	0,5106	1,0219	0,9786	0,9965
	σ_{13}	-0,5	1,5614	1,5306	1,5212	0,6098	0,5333	0,4755
	σ_{23}	-0,5	1,4487	1,4167	1,4044	0,5508	0,4757	0,4171
	σ_{33}	2	0,5334	0,4559	0,4297	1,0172	0,9717	0,9929
	σ_{14}	-0,5	1,5216	1,4895	1,4940	0,5792	0,5009	0,4441
	σ_{24}	-0,5	1,4419	1,4087	1,3960	0,5502	0,4704	0,4131
	σ_{34}	-0,5	1,3171	1,2823	1,2681	0,4923	0,4135	0,3571
	σ_{44}	2	0,5870	0,5115	0,5003	0,9978	0,9505	0,9721
	λ_1	6	5,7070	5,6591	5,2936	4,7848	4,4790	4,7400
	λ_2	3	2,9747	2,8857	2,5891	2,6406	2,3697	2,3086
	λ_3	-2	1,9565	1,8182	1,4955	1,8498	1,6031	1,4398
	λ_4	-1	1,1271	0,8836	0,5541	1,3552	1,0540	0,8806
	ν	5	1,6429	1,1547	1,0067	1,5815	1,2564	1,0717
200	μ_1	-2	1,3736	1,0701	0,7588	0,5349	0,3995	0,2866
	μ_2	-1	1,2333	1,0703	1,0569	0,4976	0,3897	0,3274
	μ_3	-1	1,2945	1,1053	1,0891	0,4711	0,3686	0,3031
	μ_4	-1	1,2750	1,1196	1,1547	0,4630	0,3662	0,3145
	σ_{11}	2	0,3540	0,2954	0,2708	1,1818	1,1617	1,1692
	σ_{12}	-0,5	1,2583	1,2314	1,2366	0,3385	0,3015	0,2876
	σ_{22}	2	0,5281	0,4789	0,4613	1,0558	1,0308	1,0187
	σ_{13}	-0,5	1,5299	1,5081	1,5196	0,4908	0,4632	0,4510
	σ_{23}	-0,5	1,4143	1,3905	1,3868	0,4495	0,4220	0,4071
	σ_{33}	2	0,4785	0,4262	0,4037	1,0480	1,0234	1,0242
	σ_{14}	-0,5	1,4786	1,4555	1,4656	0,4610	0,4326	0,4203
	σ_{24}	-0,5	1,3955	1,3722	1,3608	0,4281	0,4032	0,3906
	σ_{34}	-0,5	1,2853	1,2596	1,2666	0,3893	0,3606	0,3420
	σ_{44}	2	0,5283	0,4792	0,4702	1,0292	1,0035	0,9911
	λ_1	6	5,6643	5,6247	5,2450	4,4982	4,1717	4,4988
	λ_2	3	2,9518	2,8770	2,4713	2,5142	2,2349	2,2223
	λ_3	-2	1,8875	1,7665	1,3463	1,6537	1,4333	1,2794
	λ_4	-1	1,0936	0,8701	0,4829	1,1510	0,9159	0,7990
	ν	5	1,0547	0,8978	0,8531	1,4350	1,1179	0,9394

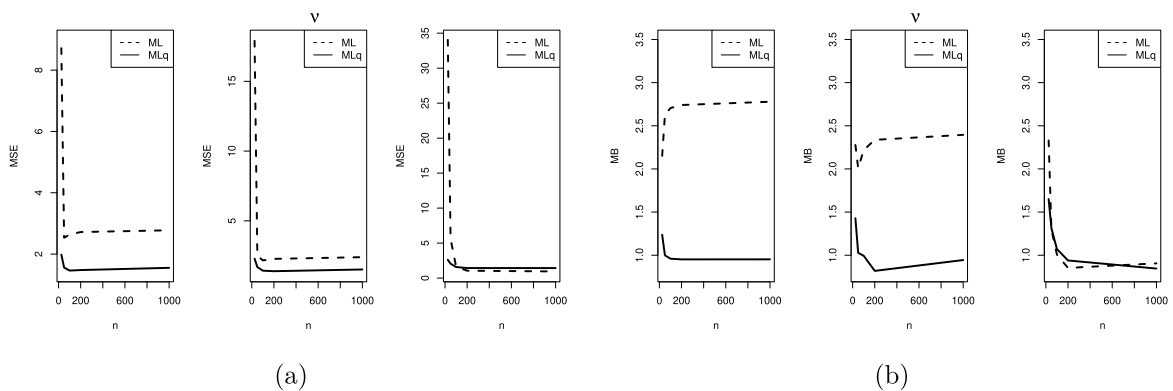
Fonte: Elaborado pela autora em 2020.

Tabela 4.24 – Cenário 2: Caso 2 - Bias, MB e MSE dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros do modelo ST, com o adicional de 30% de outliers, para $n = 1000$.

n	Parâmetro	Valor Real	ML			ML _q		
			MSE	Bias	MB	MSE _q	Bias _q	MB _q
1000	μ_1	-2	1,1067	0,8111	0,6048	0,3714	0,2741	0,2102
	μ_2	-1	1,0804	0,9187	0,9304	0,3490	0,2585	0,1678
	μ_3	-1	1,1519	0,9059	0,6635	0,2970	0,2225	0,1649
	μ_4	-1	1,1347	0,9610	0,9023	0,3138	0,2534	0,2166
	σ_{11}	2	0,2809	0,2532	0,2373	1,2034	1,1870	1,1321
	σ_{12}	-0,5	1,2149	1,1984	1,2217	0,3189	0,3041	0,3068
	σ_{22}	2	0,4674	0,4501	0,4383	1,1107	1,0854	0,9894
	σ_{13}	-0,5	1,4425	1,4259	1,4496	0,4279	0,4220	0,4130
	σ_{23}	-0,5	1,3525	1,3376	1,3582	0,3995	0,3942	0,3952
	σ_{33}	2	0,4002	0,3829	0,3752	1,1063	1,0827	1,0003
	σ_{14}	-0,5	1,3967	1,3807	1,4069	0,4069	0,4002	0,3982
	σ_{24}	-0,5	1,3329	1,3188	1,3304	0,3853	0,3808	0,3809
	σ_{34}	-0,5	1,2317	1,2164	1,2442	0,3606	0,3535	0,3473
	σ_{44}	2	0,4517	0,4353	0,4267	1,0880	1,0635	0,9848
	λ_1	6	5,4588	5,4304	5,1717	4,1211	3,8436	4,2695
	λ_2	3	2,8169	2,7530	2,3580	2,2333	2,0135	2,0312
	λ_3	-2	1,7543	1,6438	1,2594	1,4314	1,3068	1,2109
	λ_4	-1	0,9901	0,7736	0,3398	0,9857	0,7788	0,6164
ν	5	0,9482	0,8947	0,9054	1,4307	1,0751	0,8455	

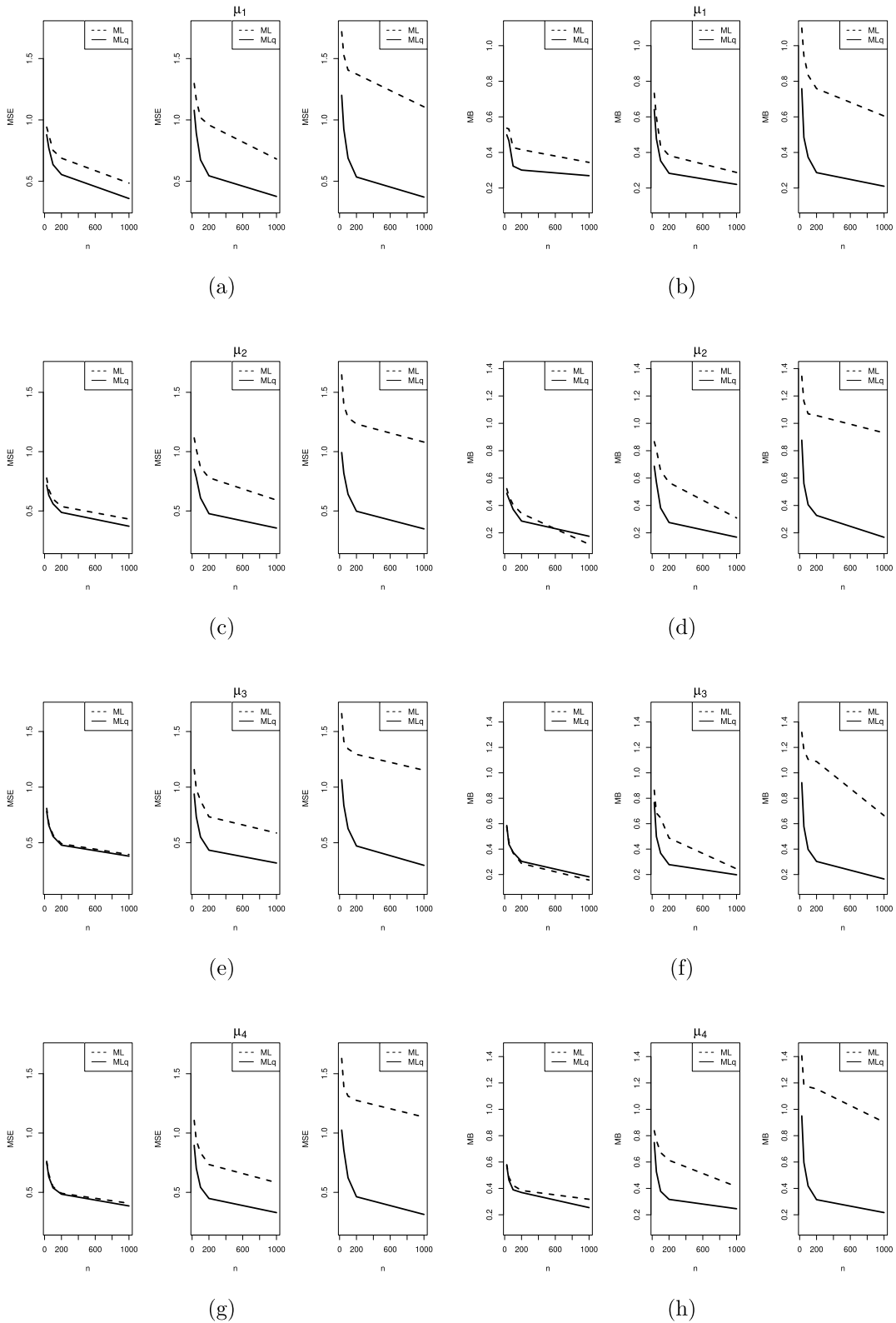
Fonte: Elaborado pela autora em 2020.

Figura 4.9 – Cenário 2: Caso 2 - MSE e VMD dos estimadores de ML ($q = 1$) e de ML_q do parâmetro ν do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



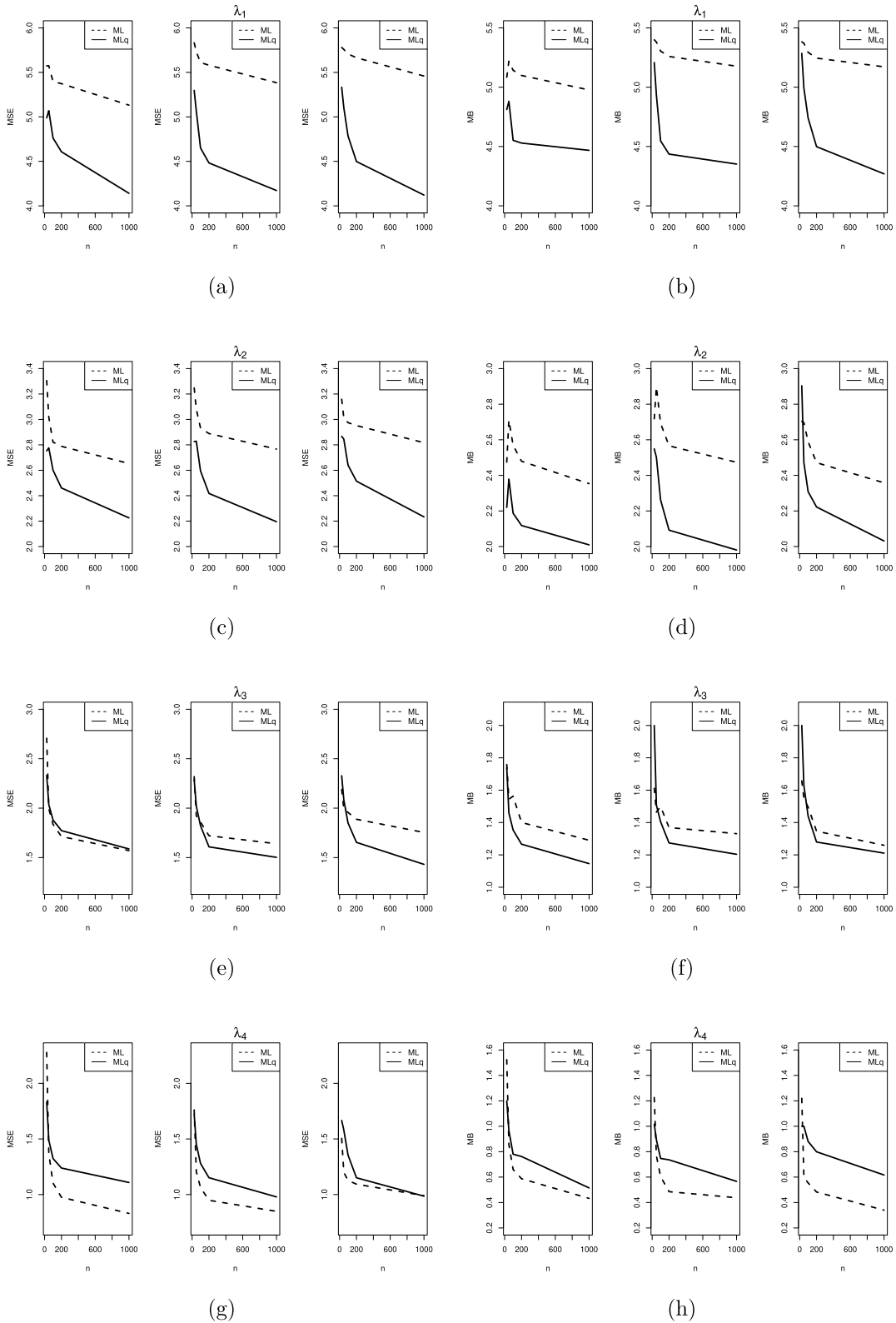
Fonte: Elaborado pela autora em 2020.

Figura 4.10 – Cenário 2: Caso 2 - MSE e VMD dos estimadores de ML ($q = 1$) e de ML_q do parâmetro μ do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



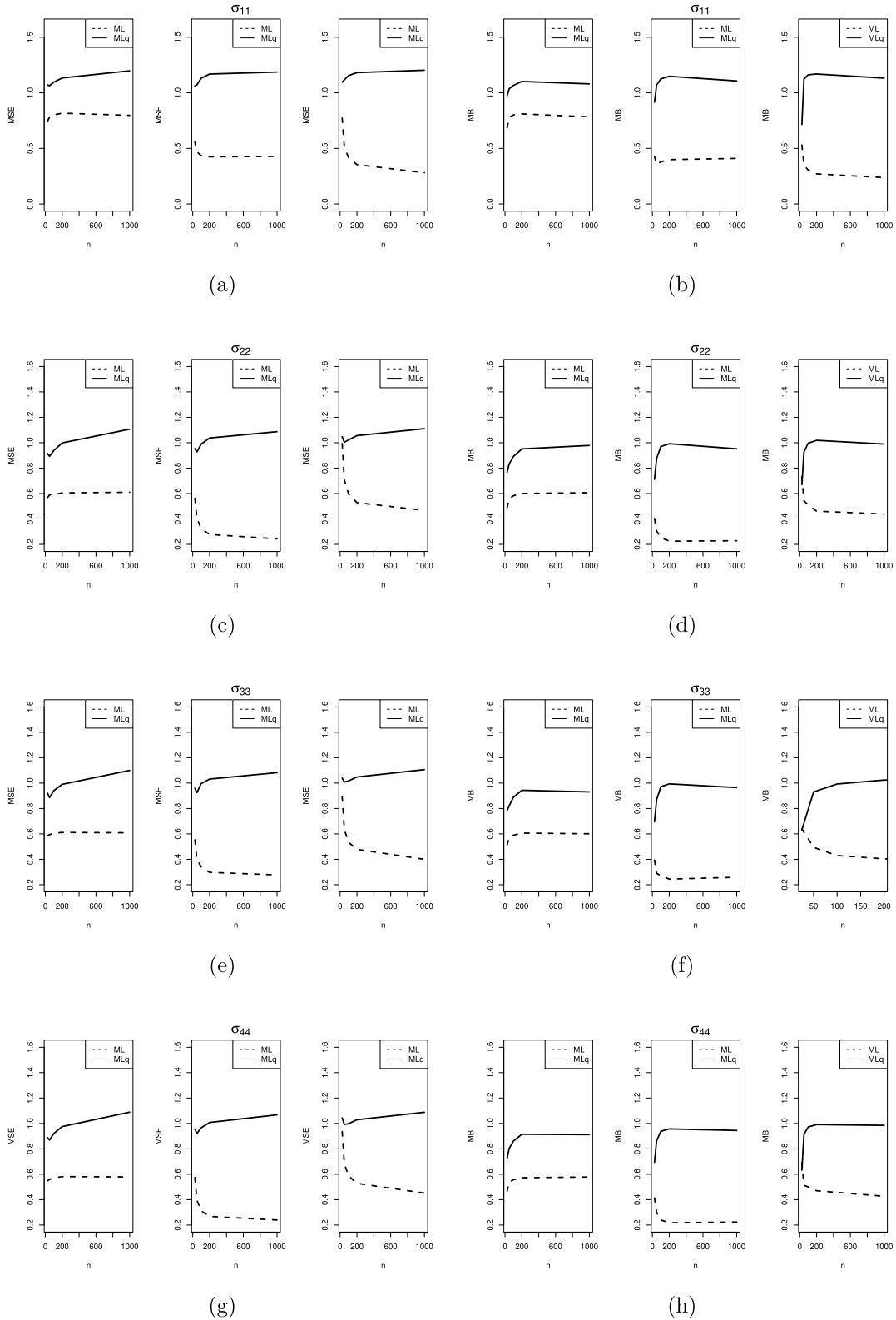
Fonte: Elaborado pela autora em 2020.

Figura 4.11 – Cenário 2: Caso 2 - MSE e VMD dos estimadores de ML ($q = 1$) e de ML_q do parâmetro λ do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



Fonte: Elaborado pela autora em 2020.

Figura 4.12 – Cenário 2: Caso 2 - MSE e VMD dos estimadores de ML ($q = 1$) e de ML_q do parâmetro Σ do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



Fonte: Elaborado pela autora em 2020.

4.2 APLICAÇÕES COM DADOS REAIS

Nesta seção, mostramos a aplicabilidade dos resultados inferenciais desenvolvidos, neste trabalho, em três conjuntos de dados reais, sendo eles Stack-Loss, Land rent e AIS, que serão devidamente apresentados com maiores detalhes abaixo. Como foi dito anteriormente, na Seção 3.4, assumimos o valor de q , onde $0 < q < 1$, que fornece o modelo de caudas mais pesadas, ou seja, com menor estimativa de ML_q para o parâmetro grau de liberdade do modelo ST ajustado. As estimativas de ML e de ML_q dos parâmetros do modelo ST são obtidas via algoritmo EM, descritos nas Seções 2.4.1.1 e 3.3.1, respectivamente, e seus desvios padrões correspondentes (SE) calculados por meio do procedimento paramétrico de bootstrap; veja [22] para mais informações. Na versão paramétrica do método bootstrap, considera-se $\hat{\boldsymbol{\theta}}$ como o valor verdadeiro do parâmetro a fim de gerar 1000 amostras do modelo ST. Esta estimativa também foi usada como chute inicial do algoritmo EM para obter cada amostra de bootstrap. Sendo assim, calcula-se o desvio padrão amostral dessas replicações. De acordo com os resultados obtidos na Seção 2.3, identificamos as observações atípicas considerando a distância de Mahalanobis. Conforme sugerido por [41], usamos as quantidades TRC (mudança relativa total) e MRC (máximo da mudança relativa) para revelar o impacto da observação atípica, possível observação influente, detectada nas estimativas de ML e ML_q de $\boldsymbol{\theta}$. Elas são definidas como

$$TRC = \sum_{j=1}^{n_p} \left| \frac{\hat{\boldsymbol{\theta}}_j - \hat{\boldsymbol{\theta}}_{[i]j}}{\hat{\boldsymbol{\theta}}_j} \right| \text{ e } MRC = \max_{j=1, \dots, n_p} \left| \frac{\hat{\boldsymbol{\theta}}_j - \hat{\boldsymbol{\theta}}_{[i]j}}{\hat{\boldsymbol{\theta}}_j} \right| \quad (4.1)$$

onde n_p é a dimensão de $\boldsymbol{\theta}$ e o $\hat{\boldsymbol{\theta}}_{[i]}$ é a estimativa de máxima verossimilhança de $\boldsymbol{\theta}$ sem a i -ésima observação.

4.2.1 Stack-Loss

O primeiro conjunto de dados que analisamos foi Stack-Loss, que é um conjunto de dados clássico, no contexto de procedimentos robustos e de presença de outliers, e foi originalmente apresentado por [15]. Este conjunto está relacionado a observação de um processo químico no decorrer de 21 dias, com quatro variáveis observadas x_1 , x_2 , x_3 relacionadas, respectivamente, ao fluxo de ar, temperatura da água e concentração de ácido e uma variável y chamada de stack-loss, conforme apresentado com mais detalhes por [20].

Apesar desse conjunto de dados ser comumente usado no contexto de regressão, não iremos abordá-lo neste cenário. Consideramos \mathbf{Y} um vetor aleatório tridimensional, seguindo os casos descritos abaixo, pois ao usá-lo como quadrimensional, observamos que, sob ambos métodos de estimação, a estimativa de ν estava muito alta e assim, estaríamos no contexto do modelo SN, retirando-nos do objetivo principal deste trabalho. Lembrando que quando $\nu \rightarrow \infty$, temos que a distribuição ST tende à distribuição SN,

como foi mostrado na Figura 2.2. Então, por este motivo, dividimos esta aplicação em dois casos:

- Caso 1: $\mathbf{Y}_i = (x_{1i}, x_{2i}, y_i)^\top$;
- Caso 2: $\mathbf{Y}_i = (x_{1i}, x_{3i}, y_i)^\top$,

com $i = 1, \dots, 21$.

4.2.1.1 Caso 1

A Tabela 4.25 apresenta as estimativas de máxima verossimilhança ML e ML_q e seus respectivos desvios padrões (SD) sob o modelo ST ajustado aos dados Stack-Loss - caso 1. Na coluna denominada amostra, temos os valores da média amostral, desvio amostral e a assimetria amostral, como se trata de dados reais não temos um valor base para os graus de liberdade. Observa-se que as estimativa de ML e ML_q para ν foram 7,7840 e 1,0264 (evidenciando caudas mais pesadas), respectivamente. Podemos observar também que, em relação ao desvio padrão, os estimadores de ML_q mostraram ser mais precisos em relação aos de ML em quase todos os parâmetros, exceto para o parâmetro de assimetria λ .

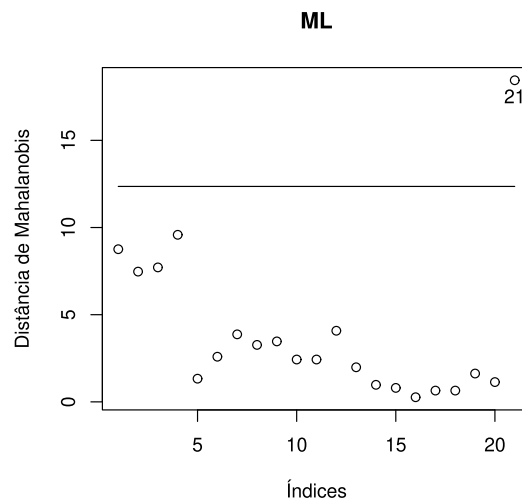
Tabela 4.25 – Stack-Loss - caso 1: Estimativas de ML ($q = 1$) e ML_q ($q = 0,51$) para os parâmetros do modelo ST ajustado com os correspondentes desvios padrões, SE e SE_q .

Parâmetro	Amostra	ML	ML_q	SE	SE_q
μ_1	60,4286	51,6037	47,7664	1,4410	1,2768
μ_2	21,0952	18,3052	18,1479	1,8951	1,5662
μ_3	17,5238	7,3151	5,8940	1,0616	0,8755
σ_{11}	84,0571	8,0871	4,6976	0,5866	0,4213
σ_{12}	22,6571	1,6983	1,5070	0,4887	0,3116
σ_{22}	9,9905	2,1315	0,9350	0,6264	0,4326
σ_{13}	85,7643	6,7789	4,0277	0,5241	0,3299
σ_{23}	28,1476	2,4719	1,6160	0,5092	0,3026
σ_{33}	103,4619	9,7723	3,9544	0,6460	0,3927
λ_1	0,8736	2,2131	2,7016	1,3521	3,2363
λ_2	0,5044	0,8502	0,6555	1,3854	3,3308
λ_3	1,2442	3,8866	3,2156	1,2646	2,8689
ν	—	7,7840	1,0264	2,4648	2,3352

Fonte: Elaborado pela autora em 2020.

A Figura 4.13 ilustra a distância de Mahalanobis para o modelo ajustado ST sob o método de estimação ML. Note que a observação 21 aparece como outlier, possível observação influente.

Figura 4.13 – Stack-Loss - caso 1: Gráfico da distância de Mahalanobis para o modelo ST ajustado, sob o método de estimação ML, considerando $H = 0,95$.



Fonte: Elaborado pela autora em 2020.

Quando excluimos a observação 21 do conjunto de dados, observamos que houve alterações nas estimativas de ML e de ML_q , com um destaque para a estimativa de ML do ν , uma vez que ao retirar esta observação atípica, seu valor foi bem próximo de 100, indicando que não há distribuição de cauda pesada. Além disso, notamos também que houve um aumento no valor na estimativa de ML_q do ν , porém em uma escala bem menor, obtendo o valor de 3,4608, com $q = 0,64$, reforçando a robustez desses novos estimadores em relação às observações atípicas. Para avaliar o impacto da observação atípica detectada nas estimativas de ML e ML_q , a Tabela 4.26 apresenta as medidas TRC e MRC, quando excluimos a observação 21. Dessa forma, comparando as mudanças relativas nas estimativas de ML e ML_q em termos de TRC e MRC, podemos observar que o valor de TRC, sob o método de estimação ML, é mais que o dobro do valor de TRC no caso ML_q . E o valor de MRC no caso ML é quase o sêxtuplo do valor atribuído à quantidade correspondente no contexto de ML_q .

Tabela 4.26 – Stack-Loss- caso 1: MCR e TRC para o modelo ST ajustado, sob ambos métodos de estimação, quando a observação 21 é excluída. Índice q indica que as medidas foram obtidas sob o método de estimação ML_q .

	TRC	TRC _q	MRC	MRC _q
θ	13,1252	5,3839	11,8468	2,3716

Fonte: Elaborado pela autora em 2020.

Proveniente dos algoritmos EM e EM_q , os pesos estimados \hat{u}_i e \hat{u}_{qi} da observação $i = 21$ são, respectivamente, 0,4216 e 0,0058, corroborando a ideia de que o ML_q é um método iterativo de dupla ponderação como foi comentado anteriormente. Como

esperado, os resultados indicam que as estimativas de ML_q são menos sensíveis à presença de dados atípicos.

4.2.1.2 Caso 2

A Tabela 4.27 apresenta as estimativas de máxima verossimilhança ML e ML_q e seus respectivos desvios padrões (SD) sob o modelo ST ajustado aos dados Stack-Loss - caso 2, e adicionalmente, estão as medidas amostrais dos dados reais, utilizadas como chute inicial dos algoritmos. Observa-se que, para este caso do Stack-Loss, a diferença entre as estimativas do ν , sob os ambos os métodos de estimação, não é tão grande como no caso anterior, isto é, as estimativas de ML e ML_q para ν foram 5,9456 e 4,6516, respectivamente. Percebe-se que ocorreu uma mudança de sinal na estimativa do parâmetro σ_{23} , sob os métodos de estimação considerados. Podemos observar também que, em relação ao desvio padrão, os estimadores de ML_q mostraram ser mais precisos em relação aos de ML em todos os parâmetros, ressaltando que o desvio padrão do estimador do parâmetro ν , que pelo caso ML, foi 2,4626 que é mais que o triplo do valor obtido pelo caso ML_q , dado por 0,8187.

Tabela 4.27 – Stack-Loss - caso 2: Estimativas de ML ($q = 1$) e ML_q ($q = 0,57$) para os parâmetros do modelo ST ajustado com os correspondentes desvios padrões, SE e SE_q .

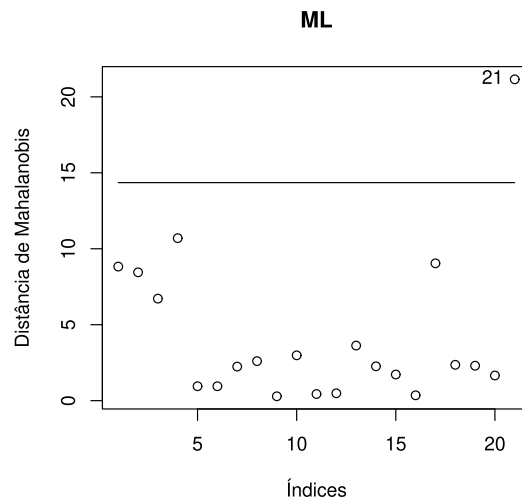
Parâmetro	Amostra	ML	ML_q	SE	SE_q
μ_1	60,4286	52,9072	55,8773	1,5772	0,8435
μ_2	86,2857	85,9485	88,1304	1,3756	0,8580
μ_3	17,5238	8,8095	11,6026	0,9726	0,7055
σ_{11}	84,0571	7,4974	3,4798	0,5954	0,8910
σ_{12}	24,5714	1,2480	0,3870	0,4767	0,4380
σ_{22}	28,7143	4,6922	2,9787	0,5697	0,6486
σ_{13}	85,7643	6,3844	2,8293	0,5260	0,4313
σ_{23}	21,7929	0,8015	-0,1619	0,4788	0,4113
σ_{33}	103,4619	9,2585	3,8042	0,5878	0,7572
λ_1	0,8736	1,3834	0,5430	1,4568	1,3171
λ_2	-0,9392	-0,3155	-0,7254	1,3488	0,6845
λ_3	1,2442	2,0081	1,6651	1,3358	1,2642
ν	—	5,9456	4,6516	2,4626	0,8187

Fonte: Elaborado pela autora em 2020.

A Figura 4.14 ilustra a distância de Mahalanobis para o modelo ajustado ST sob o método de estimação ML. Note que a observação 21 novamente aparece como outlier, possível observação influente. Ao notar que ambos casos do Stack-Loss obtiveram a mesma observação atípica decidimos averiguar o que esta observação em relação às outras possui de diferencial e percebemos que os seus valores das variáveis x_1 e x_3 se encontram acima do terceiro quartil das respectivas variáveis, com o adicional de que o seu valor observado

da variável x_3 está muito próximo ao máximo dos dados e ainda que os seus valores das variáveis x_2 e y são iguais as medianas das respectivas variáveis.

Figura 4.14 – Stack-Loss - caso 2: Gráfico da distância de Mahalanobis para o modelo ST ajustado, sob o método de estimação ML, considerando $II = 0,95$.



Fonte: Elaborado pela autora em 2020.

Quando excluimos a observação 21 do conjunto de dados, observamos que houve alterações nas estimativas de ML e de ML_q , com um destaque para a estimativa de ML do ν , uma vez que ao retirar esta observação atípica, seu valor foi bem próximo de 14. Além disso, notamos também que houve um aumento no valor na estimativa de ML_q do ν , porém em uma escala bem menor, obtendo o valor de 5,5416, com $q = 0,59$, reforçando a robustez desses novos estimadores em relação às observações atípicas. Para avaliar o impacto da observação atípica detectada nas estimativas de ML e ML_q , a Tabela 4.28 apresenta as medidas TRC e MRC, quando excluimos a observação 21. Dessa forma, comparando as mudanças relativas nas estimativas de ML e ML_q em termos de TRC e MRC, é interessante notar que as maiores modificações ocorrem sob o método de estimação ML.

Tabela 4.28 – Stack-Loss - caso 2: MCR e TRC para o modelo ST ajustado, sob ambos métodos de estimação, quando a observação 21 é excluída. Índice q indica que as medidas foram obtidas sob o método de estimação ML_q .

	TRC	TRC _q	MRC	MRC _q
θ	4,1060	2,2854	1,3377	0,9222

Fonte: Elaborado pela autora em 2020.

Proveniente dos algoritmos EM e EM_q , os pesos estimados \hat{u}_i e \hat{u}_{qi} da observação $i = 21$ são, respectivamente, 0,3412 e 0,0680, corroborando a ideia de que o ML_q é

um método iterativo de dupla reponderação como foi comentado anteriormente. Como esperado, os resultados indicam que as estimativas de ML_q são menos sensíveis à presença de dados atípicos.

4.2.2 Land Rent

[62] relatou o conjunto de dados Land rent, relacionado ao aluguel de terras, que posteriormente foi analisado por [65] no contexto de modelo de regressão sob a distribuição t-Student. Este conjunto de dados inclui as seguintes variáveis: y é a variável relacionada a renda média paga pelo aluguel de acre de plantio de alfafa, x_1 está relacionada ao aluguel médio pago por terra cultivável, x_2 é a densidade de vacas leiteiras (número de vacas por milha quadrada), x_3 é a proporção de terras agrícolas usadas para pasto e x_4 como variável binária sobre a necessidade do preparo do solo ou não para o plantio.

Nesta aplicação, usamos quatro dessas cinco variáveis, desconsiderando a variável x_4 , pois é uma variável booleana. Apesar desse conjunto de dados ser comumente usado no contexto de regressão, não iremos abordá-lo neste cenário. Seguindo [21], consideramos \mathbf{Y} um vetor aleatório quadridimensional, ou seja, $\mathbf{Y}_i = (y_i, x_{1i}, x_{2i}, x_{3i})^\top$, com $i = 1, \dots, 67$.

A Tabela 4.29 apresenta as estimativas de máxima verossimilhança ML e ML_q e seus respectivos desvios padrões (SD) sob o modelo ST ajustado aos dados Land rent. Na coluna denominada amostra, temos os valores da média amostral, desvio amostral e a assimetria amostral, como se trata de dados reais não temos um valor base para os graus de liberdade. Observa-se que as estimativas de ML e ML_q para ν foram 4,3956 e 2,6482, respectivamente. Percebe-se que ocorreu uma mudança de sinal na estimativa do parâmetro λ_4 , sob os métodos de estimação considerados. Podemos observar também que, em relação ao desvio padrão, os estimadores de ML_q mostraram ser mais precisos em relação aos de ML em quase todos os parâmetros, exceto para o parâmetro de assimetria λ .

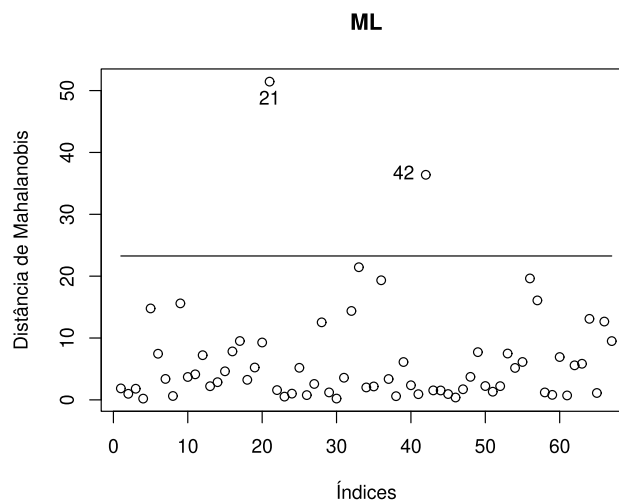
A Figura 4.15 ilustra a distância de Mahalanobis para o modelo ajustado ST sob o método de estimação ML. Note que as observações 21 e 42 aparecem como outliers. Decidimos então averiguar o que essas duas observações possuem para torná-las atípicas, para a componente y observamos que os valores de ambas observações se encontram abaixo do primeiro quartil, já para x_1 a observação 42 está abaixo da mediana enquanto que a observação 21 está abaixo do primeiro quartil. Para a componente x_2 , o valor obtido pela observação 21 se encontra entre a mediana e o terceiro quartil e é maior que a média, enquanto que o valor obtido para a observação 42 está acima do terceiro quartil, para a componente x_3 o valor relacionado a observação 21 é o máximo enquanto que o valor relacionado a observação 42 está acima do terceiro quartil.

Tabela 4.29 – Land rent: Estimativas de ML ($q = 1$) e ML_q ($q = 0, 50$) para os parâmetros do modelo ST ajustado com os correspondentes desvios padrões, SE e SE_q .

Parâmetro	Amostra	ML	ML_q	SE	SE_q
μ_1	42,1661	27,7449	24,2149	1,8057	0,6855
μ_2	43,6398	33,0639	33,9225	1,9062	0,7959
μ_3	20,5633	8,4320	1,9929	1,2718	0,3772
μ_4	0,1697	0,1075	0,0909	0,1240	0,0291
σ_{11}	510,1541	19,1219	8,0786	0,5745	0,4510
σ_{12}	418,3721	12,2971	6,5890	0,5353	0,3533
σ_{22}	447,3442	17,4646	8,4595	0,5425	0,4725
σ_{13}	106,8864	5,7408	3,1573	0,5335	0,2014
σ_{23}	15,8028	1,4430	1,0717	0,5333	0,1625
σ_{33}	235,1974	15,1266	5,0644	0,5054	0,2971
σ_{14}	-1,0554	-0,0001	0,0111	0,0623	0,0171
σ_{24}	-1,5275	-0,0332	-0,0140	0,0602	0,0155
σ_{34}	1,1581	0,0819	0,0221	0,0645	0,0184
σ_{44}	0,0209	0,0653	0,0149	0,0324	0,0183
λ_1	0,5607	0,6755	1,9885	0,7623	0,7856
λ_2	0,1022	0,5721	0,7123	0,8071	0,6009
λ_3	0,7378	1,0673	4,0984	0,6765	0,9410
λ_4	1,8078	0,3211	-0,5465	0,6474	0,5276
ν	—	4,3956	2,6482	1,5717	1,3069

Fonte: Elaborado pela autora em 2020.

Figura 4.15 – Land rent: Gráfico da distância de Mahalanobis para o modelo ST ajustado, sob o método de estimação ML, considerando $\Pi = 0.95$.



Fonte: Elaborado pela autora em 2020.

Ao observar os pesos \hat{u}_i e \hat{u}_{qi} , que podem ser vistos na tabela 4.30, provenientes dos algoritmos EM e EM_q , respectivamente, atribuídos a tais observações atípicas,

notamos que, como esperado, ambos métodos de estimação, no contexto do modelo ST, podem atribuir pesos diferentes para cada observação, sendo ainda menores para o caso ML_q . Consequentemente, o modelo ST controla melhor a influência dessas observações no processo de estimação ML_q .

Tabela 4.30 – Land rent: Pesos atribuídos as observações atípicas.

Índice i	\hat{u}_i	\hat{u}_{qi}
21	0,1578	0,0061
42	0,2154	0,0080

Fonte: Elaborado pela autora em 2020.

Como este conjunto de dados possui apenas 2 observações atípicas, realizamos a exclusão conjunta e individual delas. A comparação das medidas TRC e MRC, calculadas no modelo ST, sob ambos métodos de estimação estão apresentadas na Tabela 4.31. Os resultados mostram que as estimativas dos parâmetros são menos sensíveis à presença de dados atípicos e/ou influentes quando distribuição com caudas mais pesadas são usadas, ou seja, no contexto do processo de estimação de ML_q .

Tabela 4.31 – Land rent: MCR e TRC para o modelo ST ajustado, sob ambos métodos de estimação, quando observações atípicas são excluídas. Índice q indica que as medidas foram obtidas sob o método de estimação ML_q .

Observações excluídas	TRC	TRC $_q$	MRC	MRC $_q$
21	2,2606	0,6880	1,5298	0,1342
42	18,4253	0,7738	17,8412	0,1394
21 e 42	12,5354	2,9292	11,2690	0,5197

Fonte: Elaborado pela autora em 2020.

4.2.3 AIS

Em [18] é apresentado o conjunto de dados AIS composto por variáveis biomédicas de 202 atletas da elite australiana que treinam no AIS (Australian Institute of Sport), sendo igualmente divididos entre homens e mulheres. Como em [7], consideramos quatro variáveis deste conjunto de dados, sendo elas $\{BMI, Bfat, ssf, LBM\}$ que representam o índice de massa corporal (BMI), a porcentagem de gordura corporal ($Bfat$), a soma das dobras cutâneas (ssf) e a massa magra (LBM).

A Tabela 4.32 apresenta as estimativas de máxima verossimilhança ML e ML_q e seus respectivos desvios padrões (SD) sob o modelo ST ajustado aos dados AIS, e adicionalmente, estão as medidas amostrais dos dados reais, utilizadas como chute inicial dos algoritmos. Observa-se que as estimativa de ML e ML_q para ν foram 12,1074 e 2,4626 (evidenciando caudas mais pesadas), respectivamente. Percebe-se que ocorreu um

mudança de sinal nas estimativas dos parâmetros σ_{24} e σ_{34} , sob os métodos de estimação considerados. Podemos observar também que, em relação ao desvio padrão, os estimadores de ML_q mostraram ser mais precisos em relação aos de ML em quase todos os parâmetros, exceto para o parâmetro de assimetria λ .

Tabela 4.32 – AIS: Estimativas de ML ($q = 1$) e ML_q ($q = 0, 47$) para os parâmetros do modelo ST ajustado com os correspondentes desvios padrões, SE e SE_q .

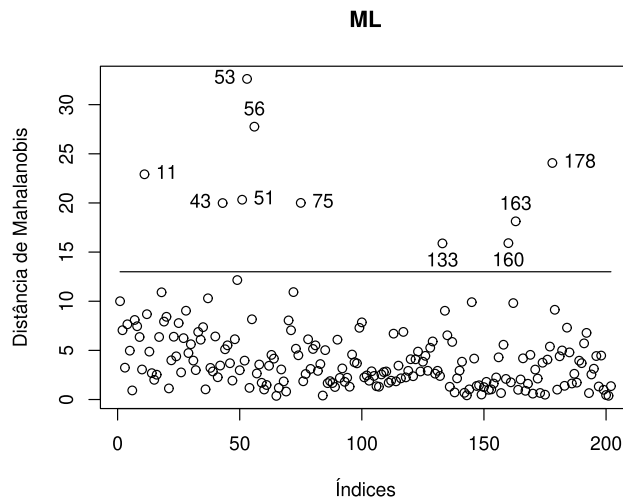
Parâmetro	Amostra	ML	ML_q	SE	SE_q
μ_1	22,9559	20,7937	20,7858	0,4793	0,3807
μ_2	13,5074	9,9519	5,6580	0,5812	0,1253
μ_3	69,0218	49,9101	29,5222	1,8971	0,5934
μ_4	64,8737	61,2288	66,5692	1,6992	1,0633
σ_{11}	8,2021	1,9398	0,7684	0,0822	0,0997
σ_{12}	3,3249	0,2787	0,3153	0,1104	0,0449
σ_{22}	38,3139	2,4004	0,4138	0,0856	0,0394
σ_{13}	29,9489	1,5749	1,2444	0,2026	0,1425
σ_{23}	194,1189	5,9912	2,0015	0,2000	0,0752
σ_{33}	1060,5009	32,9368	12,4912	0,3795	0,3484
σ_{14}	26,7213	2,1763	0,2782	0,1780	0,0974
σ_{24}	-29,2745	-0,9216	0,2874	0,1947	0,0618
σ_{34}	-88,4253	-0,4219	2,5499	0,4705	0,2480
σ_{44}	170,8301	12,7962	3,0054	0,2405	0,2022
λ_1	0,9465	1,0220	1,3026	0,4678	1,4445
λ_2	0,7595	0,5852	1,0511	0,5832	1,1229
λ_3	1,1747	1,1502	6,0762	0,6631	2,2710
λ_4	0,3585	0,6275	1,0025	0,7674	1,5905
ν	—	12,1074	2,6165	3,9704	2,1865

Fonte: Elaborado pela autora em 2020.

A Figura 4.16 ilustra a distância de Mahalanobis para o modelo ajustado ST sob o método de estimação ML. Note que as observações 11, 43, 51, 53, 56, 75, 133, 160, 163 e 178 se destacam como outliers. Do mesmo modo como foi feito para os outros casos estaremos averiguando o que levou cada observação a ser atípica. Para a componente *BMI* notamos que as observações que obtiveram o valor acima da mediana foram as observações 11 e 43, já as observações 51, 53, 56, 75, 133, 160 e 178 obtiveram valores acima do terceiro quartil enquanto que a observação 163 obteve o valor máximo. Para a componente *Bfat* notamos que as observações 11, 43, 51, 53 e 75 obtiveram valores acima do terceiro quartil, a observação 53 obteve o valor máximo, as observações 133, 163 e 178 possuem valores acima da média e da mediana enquanto que a observação 160 está acima do primeiro quartil. Para a componente *ssf* temos que os valores relacionados as observações 11, 43, 51, 53, 75 e 178 se encontram acima do terceiro quartil, a observação 56 obteve valor máximo, as observações 133 e 163 receberam valores acima da média e da mediana enquanto que a observação 160 obteve um valor abaixo da media porém acima

do primeiro quartil. Para a última componente, *LBM*, observamos que as observações 11 e 75 obtiveram valores maiores que a média e a mediana, já as observações 43, 51, 53 e 75 se encontram abaixo do primeiro quartil, as observações 133, 160 e 178 receberam valores que se encontram acima do terceiro quartil e a observação 163 recebeu o valor máximo.

Figura 4.16 – AIS: Gráfico da distância de Mahalanobis para o modelo ST ajustado, sob o método de estimação ML, considerando $\Pi = 0,95$.



Fonte: Elaborado pela autora em 2020.

Sem perda de generalidade, para ilustrar o impacto das observações atípicas detectadas nas estimativas de ML e ML_q, a Tabela 4.33 apresenta as medidas TRC e MRC, quando excluímos todas as observações que se destacaram na Figura 4.16 conjuntamente e a observação 163 individualmente. Note que as maiores mudanças ocorreram no modelo ST ajustado sob o método de estimação ML.

Tabela 4.33 – AIS: MCR e TRC para o modelo ST ajustado, sob ambos métodos de estimação, quando observações atípicas são excluídas. Índice q indica que as medidas foram obtidas sob o método de estimação ML_q.

Observações excluídas	TRC	TRC _q	MRC	MRC _q
163	0,8005	0,6676	0,3954	0,1242
11, 43, 51, 53, 56, 75, 133, 160, 163 e 178	9,1757	4,1131	7,2594	1,0954

Fonte: Elaborado pela autora em 2020.

Na tabela 4.34, temos os pesos atribuídos \hat{u}_i e \hat{u}_{qi} , provenientes dos algoritmos EM e EM_q, respectivamente, às observações atípicas identificadas na Figura 4.16. Note que os pesos \hat{u}_{qi} são os menores, confirmando a flexibilidade das estimativas de ML_q do modelo ST com caudas pesadas contra observações atípicas. Note que para o modelo ST, os pesos \hat{u}_i e \hat{u}_{qi} , obtidos na etapa E, são inversamente proporcionais a distância de Mahalanobis,

isto é, estes processos de estimação tendem a atribuir menor peso às observações atípicas de acordo com a distância de Mahalanobis. Entretanto, o algoritmo EM_q proposto, neste trabalho, fornece estimadores duplamente reponderados para os todos os parâmetros da distribuição ST e assim, este processo de estimação de θ tende a atribuir pesos menores ainda para as observações anômalas no sentido da distância de Mahalanobis.

Tabela 4.34 – AIS: Pesos atribuídos às observações atípicas.

Índice i	\hat{u}_i	\hat{u}_{qi}
11	0,4697	0,0122
43	0,5134	0,0131
51	0,5082	0,0132
53	0,3683	0,0099
56	0,4114	0,0038
75	0,5067	0,0613
133	0,5896	0,0706
160	0,5802	0,0093
163	0,5378	0,0254
178	0,4505	0,0059

Fonte: Elaborado pela autora em 2020.

5 CONCLUSÕES E TRABALHOS FUTUROS

5.1 CONCLUSÕES

Neste trabalho, desenvolvemos o método de estimação via máxima verossimilhança L_q , que pode ser considerado uma generalização da estimação de máxima verossimilhança usual, para todos os parâmetros da distribuição Skew-t. Fornecemos um algoritmo EM modificado para o cálculo dos estimadores ML_q , mostrando a robustez do método e um melhor ajuste do modelo ST às observações atípicas, ao se estimar de forma simultânea os graus de liberdade com os demais parâmetros do modelo de interesse. Para uma ilustração do desempenho do método ML_q sobre o método ML, foram feitos estudos de simulação considerando modelos ST de dimensões variadas com diferentes configurações para os valores dos parâmetros, onde pudemos observar que os estimadores ML_q mostra um desempenho superior superior, em termos de viés e EQM, de um modo geral. Entretanto, é importante ressaltar que todas as estimativas obtidas pelo ML_q são robustas contra outliers diferentemente das estimativas obtidas pelo ML.

Observamos também a aplicabilidade do método ao estudá-lo em três conjuntos de dados reais, onde observamos que o ML_q fornece um menor peso às observações atípicas durante o processo de estimação via algoritmo EM_q , pois os estimadores resultantes são duplamente reponderados. Avaliamos também a influência dos outliers sobre o conjunto de dados comparando as mudanças relativas nas estimativas de máxima verossimilhança em termos das medidas TRC e MRC. Este trabalho generaliza os resultados obtidos em [21] que estudou o método de máxima verossimilhança L_q para a distribuição t-Student.

5.2 TRABALHOS FUTUROS

Para trabalhos futuros, pretendemos estender o método de estimação ML_q via algoritmo EM_q para a classe de distribuições misturas de escala skew-normal, inclusive no contexto de regressão ao considerar que os erros seguem uma distribuição nesta classe assimétrica. Estudaremos também outras formas de escolha para o parâmetro de distorção q e buscaremos um maior conhecimento sobre as propriedades assintóticas do estimador ML_q , tais como, a eficiência assintótica.

REFERÊNCIAS

- [1] ADCOCK, Chris J. Asset pricing and portfolio selection based on the multivariate extended skew-Student-t distribution. **Annals of operations research**, v. 176, n. 1, p. 221-234, 2010.
- [2] AKAIKE, Hirotogu. Information theory and an extension of the maximum likelihood principle. In: **Proceeding of the Second International Symposium on Information Theory**, p. 267–281. 1973.
- [3] ALYAKIN, Anton; QIN, Yichen; PRIEBE, Carey E. LqRT: Robust Hypothesis Testing of Location Parameters using Lq-Likelihood-Ratio-Type Test in Python. **arXiv preprint arXiv:1911.11922**, 2019.
- [4] ANDREWS, David F.; MALLOWS, Colin L. Scale mixtures of normal distributions. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 36, n. 1, p. 99-102, 1974.
- [5] ARELLANO-VALLE, Reinaldo B.; GENTON, Marc G. On fundamental skew distributions. **Journal of Multivariate Analysis**, v. 96, n. 1, p. 93-116, 2005.
- [6] AZZALINI, Adelchi. A class of distributions which includes the normal ones. **Scandinavian journal of statistics**, p. 171-178, 1985.
- [7] AZZALINI, Adelchi; CAPITANIO, Antonella. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 65, n. 2, p. 367-389, 2003.
- [8] AZZALINI, Adelchi; CAPITANIO, Antonella. Statistical applications of the multivariate skew normal distribution. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 61, n. 3, p. 579-602, 1999.
- [9] AZZALINI, Adelchi; CAPITANIO, Antonella. The Skew Normal and Related Distributions. 2014.
- [10] AZZALINI, Adelchi. Flexible distributions as an approach to robustness: the skew-t case. In: **Recent Advances in Robust Statistics: Theory and Applications**. Springer, New Delhi, 2016. p. 1-16.
- [11] AZZALINI, Adelchi; GENTON, Marc G. Robust likelihood methods based on the skew-t and related distributions. **International Statistical Review**, v. 76, n. 1, p. 106-129, 2008.
- [12] AZZALINI, Adelchi; SALEHI, Mahdi. Some computational aspects of maximum likelihood estimation of the skew-t distribution. In: **Computational and Methodological Statistics and Biostatistics**. Springer, Cham, 2020. p. 3-28.
- [13] AZZALINI, Adelchi; VALLE, A. Dalla. The multivariate skew-normal distribution. **Biometrika**, v. 83, n. 4, p. 715-726, 1996.
- [14] BRANCO, Márcia D.; DEY, Dipak K. A general class of multivariate skew-elliptical distributions. **Journal of Multivariate Analysis**, v. 79, n. 1, p. 99-113, 2001.

- [15] BROWNLEE, Kenneth Alexander. **Statistical theory and methodology in science and engineering**. New York: Wiley, 1965.
- [16] CASELLA, George; BERGER, Roger L. **Statistical inference**. Pacific Grove, CA: Duxbury, 2002.
- [17] COOK, R. D. Assessment of local influence. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 48, n. 2, p. 133–169, 1986.
- [18] COOK, R. Dennis; WEISBERG, Sanford. **An introduction to regression graphics**. John Wiley & Sons, 2009.
- [19] DEMPSTER, Arthur P.; LAIRD, Nan M.; RUBIN, Donald B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 39, n. 1, p. 1-22, 1977.
- [20] DODGE, Yadolah. The guinea pig of multiple regression. In: **Robust statistics, data analysis, and computer intensive methods**. Springer, New York, NY, 1996. p. 91-117.
- [21] DOĞRU, Fatma Zehra; BULUT, Y. Murat; ARSLAN, Olcay. Doubly reweighted estimators for the parameters of the multivariate t-distribution. **Communications in Statistics-Theory and Methods**, v. 47, n. 19, p. 4751-4771, 2018.
- [22] EFRON, Bradley. Computers and the theory of statistics: thinking the unthinkable. **SIAM review**, v. 21, n. 4, p. 460-480, 1979.
- [23] FANG, K. T.; KOTZ, S.; NG, K. W. **Simetric multivariate and related distributions**. Chapman and Hall, London, 1990.
- [24] FERRARI, Davide; PATERLINI, Sandra. The maximum Lq-likelihood method: an application to extreme quantile estimation in finance. **Methodology and Computing in Applied Probability**, v. 11, n. 1, p. 3-19, 2009.
- [25] FERRARI, Davide; YANG, Yuhong. Maximum Lq-likelihood estimation. **The Annals of Statistics**, v. 38, n. 2, p. 753-783, 2010.
- [26] GELL-MANN, Murray; TSALLIS, Constantino (Ed.). **Nonextensive entropy: interdisciplinary applications**. Oxford University Press on Demand, 2004.
- [27] GIUZIO, Margherita; FERRARI, Davide; PATERLINI, Sandra. Sparse and robust normal and t-portfolios by penalized Lq-likelihood minimization. **European Journal of Operational Research**, v. 250, n. 1, p. 251-261, 2016.
- [28] HAMPEL, Frank R. et al. **Robust statistics: the approach based on influence functions**. John Wiley & Sons, 2011.
- [29] HAMPEL, Frank R. The influence curve and its role in robust estimation. **Journal of the american statistical association**, v. 69, n. 346, p. 383-393, 1974.
- [30] HAVRDA, Jan; CHARVÁT, František. Quantification method of classification processes. Concept of structural α -entropy. **Kybernetika**, v. 3, n. 1, p. (30)-35, 1967.
- [31] HAWKINS, D. M. **Identification of Outliers**. Chapman and Hall, London, New York, 1980.

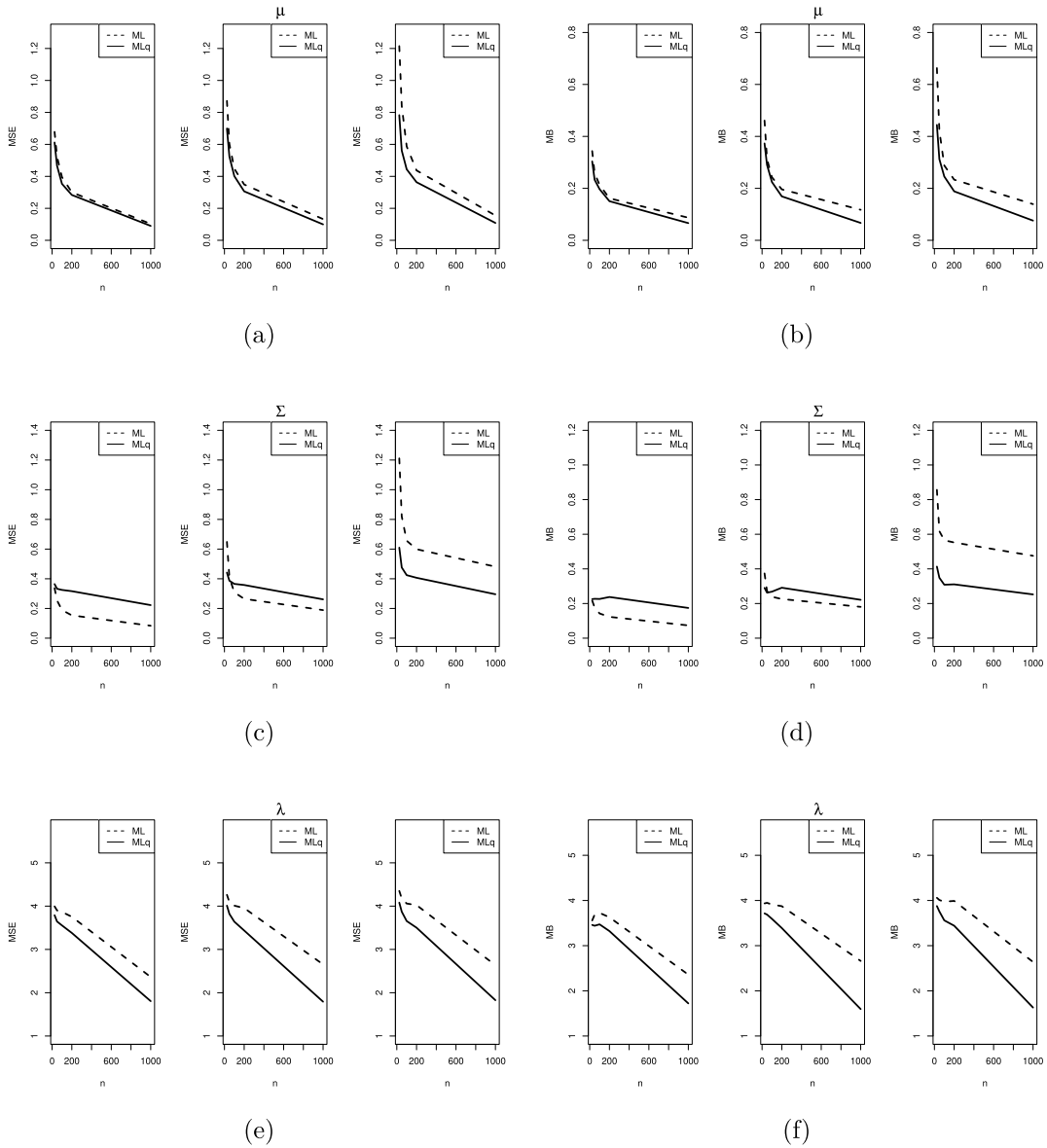
- [32] JAYNES, Edwin T. Information theory and statistical mechanics. **Physical review**, v. 106, n. 4, p. 620, 1957.
- [33] JAYNES, Edwin T. Information theory and statistical mechanics. II. **Physical review**, v. 108, n. 2, p. 171, 1957.
- [34] JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous univariate distributions**. John Wiley, New York, 1994.
- [35] KANFER, Frans; MILLARD, Sollie. A Scale Mixture Approach to t-Distributed Mixture Regression. In: *Computational and Methodological Statistics and Biostatistics*. Springer, Cham, 2020. p. 329-347.
- [36] LACHOS, Victor H.; LABRA, F. V. **Skew-normal/independent distributions, with applications**. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica, 2007.
- [37] LACHOS, Victor H.; GHOSH, Pulak; ARELLANO-VALLE, Reinaldo B. Likelihood based inference for skew-normal independent linear mixed models. **Statistica Sinica**, p. 303-322, 2010.
- [38] LAMPEL, Joseph; SHAMSIE, Jamal; SHAPIRA, Zur. Experiencing the improbable: Rare events and organizational learning. *Organization science*, v. 20, n. 5, p. 835-845, 2009. LAMPEL, Joseph; SHAMSIE, Jamal; SHAPIRA, Zur. Experiencing the improbable: Rare events and organizational learning. **Organization science**, v. 20, n. 5, p. 835-845, 2009.
- [39] LANGE, Kenneth L.; LITTLE, Roderick JA; TAYLOR, Jeremy MG. Robust statistical modeling using the t distribution. **Journal of the American Statistical Association**, v. 84, n. 408, p. 881-896, 1989.
- [40] LANGE, Kenneth; SINSHEIMER, Janet S. Normal/independent distributions and their applications in robust regression. **Journal of Computational and Graphical Statistics**, v. 2, n. 2, p. 175-198, 1993.
- [41] LEE, Sik-Yum; LU, Bin; SONG, Xin-Yuan. Assessing local influence for nonlinear structural equation models with ignorable missing data. **Computational statistics & data analysis**, v. 50, n. 5, p. 1356-1377, 2006.
- [42] LEYS, Christophe et al. Detecting multivariate outliers: Use a robust variant of the Mahalanobis distance. **Journal of Experimental Social Psychology**, v. 74, p. 150-156, 2018.
- [43] LIU, Chuanhai; RUBIN, Donald B. The ECME algorithm: a simple extension of EM and ECM with faster monotone convergence. **Biometrika**, v. 81, n. 4, p. 633-648, 1994.
- [44] LUCAS, André. Robustness of the student t based M-estimator. **Communications in Statistics-Theory and Methods**, v. 26, n. 5, p. 1165-1182, 1997.
- [45] MAHALANOBIS, Prasanta Chandra. On test and measures of group divergence: theoretical formulae. 1930.

- [46] MATLAB. 9.7.0.1190202 (R2019b). Natick, Massachusetts: The MathWorks Inc.; 2018.
- [47] MARONNA, Ricardo Antonio. Robust M-estimators of multivariate location and scatter. **The annals of statistics**, p. 51-67, 1976.
- [48] MEUCCI, Attilio. Beyond Black-Litterman: Views on non-normal markets. **Available at SSRN 848407**, 2005.
- [49] MENG, Xiao-Li; RUBIN, Donald B. Maximum likelihood estimation via the ECM algorithm: A general framework. **Biometrika**, v. 80, n. 2, p. 267-278, 1993.
- [50] MONTGOMERY, Douglas C.; PECK, Elizabeth A.; VINING, G. Geoffrey. **Introduction to linear regression analysis**. John Wiley & Sons, 2012.
- [51] PAULA, Gilberto A. et al. Robust statistical modeling using the Birnbaum-Saunders-t distribution applied to insurance. **Applied Stochastic Models in Business and Industry**, v. 28, n. 1, p. 16-34, 2012.
- [52] PAWITAN, Y. (2001). **In All Likelihood: Statistical Modelling and Inference Using Likelihood**. Oxford, Clarendon Press.
- [53] PINHEIRO, José C.; LIU, Chuanhai; WU, Ying Nian. Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution. **Journal of Computational and Graphical Statistics**, v. 10, n. 2, p. 249-276, 2001.
- [54] PITT, Ivan L. Superstar effects on royalty income in a performing rights organization. *Journal of Cultural Economics*, v. 34, n. 3, p. 219-236, 2010.
- [55] PRATES, Marcos Oliveira; LACHOS, Victor Hugo; CABRAL, C. mixsmsn: Fitting finite mixture of scale mixture of skew-normal distributions. **Journal of Statistical Software**, v. 54, n. 12, p. 1-20, 2013.
- [56] QIN, Yichen; PRIEBE, Carey E. Maximum L q-Likelihood estimation via the Expectation-Maximization algorithm: a robust estimation of mixture models. **Journal of the American Statistical Association**, v. 108, n. 503, p. 914-928, 2013.
- [57] RITTER, Gunter. **Robust cluster analysis and variable selection**. CRC Press, 2014.
- [58] TEAM, R. Core et al. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>, 2020.
- [59] TSALLIS, Constantino; MENDES, RenioS; PLASTINO, Anel R. The role of constraints within generalized nonextensive statistics. **Physica A: Statistical Mechanics and its Applications**, v. 261, n. 3-4, p. 534-554, 1998.
- [60] TSALLIS, Constantino. Possible generalization of Boltzmann-Gibbs statistics. **Journal of statistical physics**, v. 52, n. 1-2, p. 479-487, 1988.
- [61] WALLS, W. David. Modelling heavy tails and skewness in film returns. **Applied Financial Economics**, v. 15, n. 17, p. 1181-1188, 2005.

- [62] WEISBERG, Sanford. **Applied linear regression**. John Wiley & Sons, 2005.
- [63] XU, Lin; XIANG, Sijia; YAO, Weixin. Robust maximum L_q-likelihood estimation of joint mean–covariance models for longitudinal data. **Journal of Multivariate Analysis**, v. 171, p. 397-411, 2019.
- [64] ZELLER, Camila B.; LACHOS, Victor H.; VILCA-LABRA, Filidor E. Local influence analysis for regression models with scale mixtures of skew-normal distributions. **Journal of Applied Statistics**, v. 38, n. 2, p. 343-368, 2011.
- [65] ZHAO, Weihua; ZHANG, Riquan. Variable selection of varying dispersion student-t regression models. **Journal of Systems Science and Complexity**, v. 28, n. 4, p. 961-977, 2015.

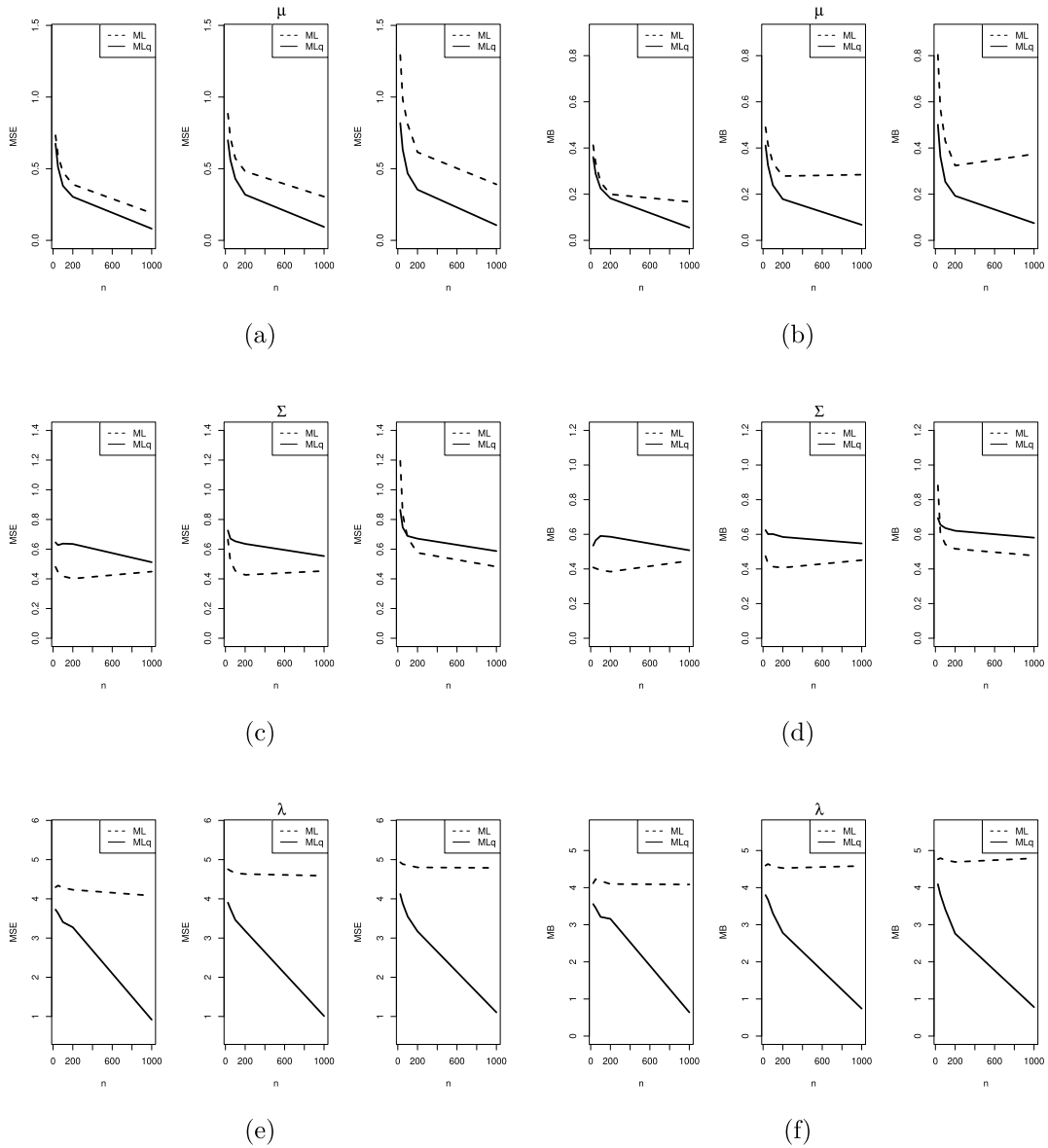
APÊNDICE A – RESULTADOS ADICIONAIS DO CAPÍTULO 4

Figura A.1 – Cenário 1: Caso 1 - Média das medidas MSE e MB dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros μ , Σ e λ do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



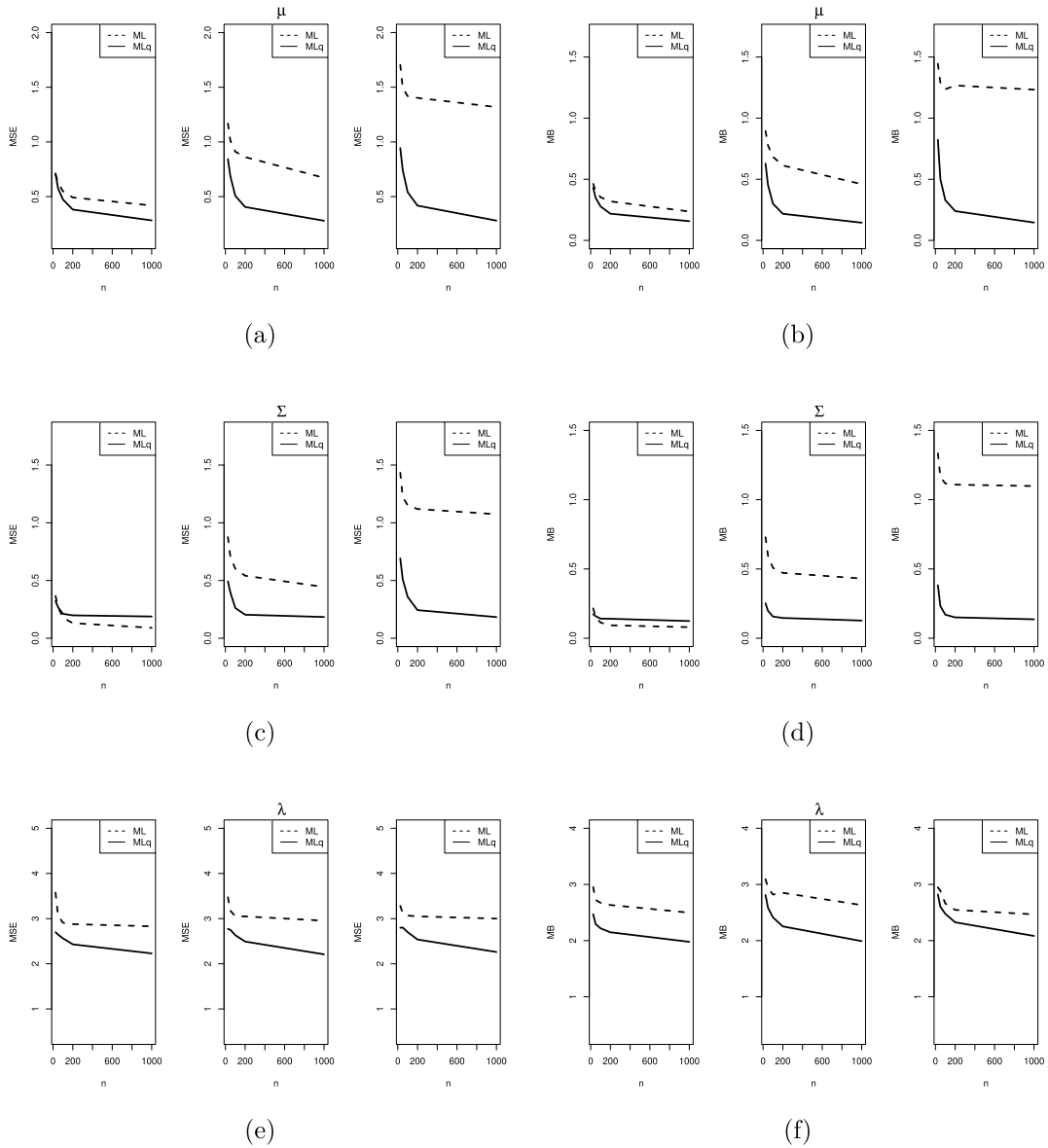
Fonte: Elaborado pela autora em 2020.

Figura A.2 – Cenário 1: Caso 2 - Média das medidas MSE e MB dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros μ , Σ e λ do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



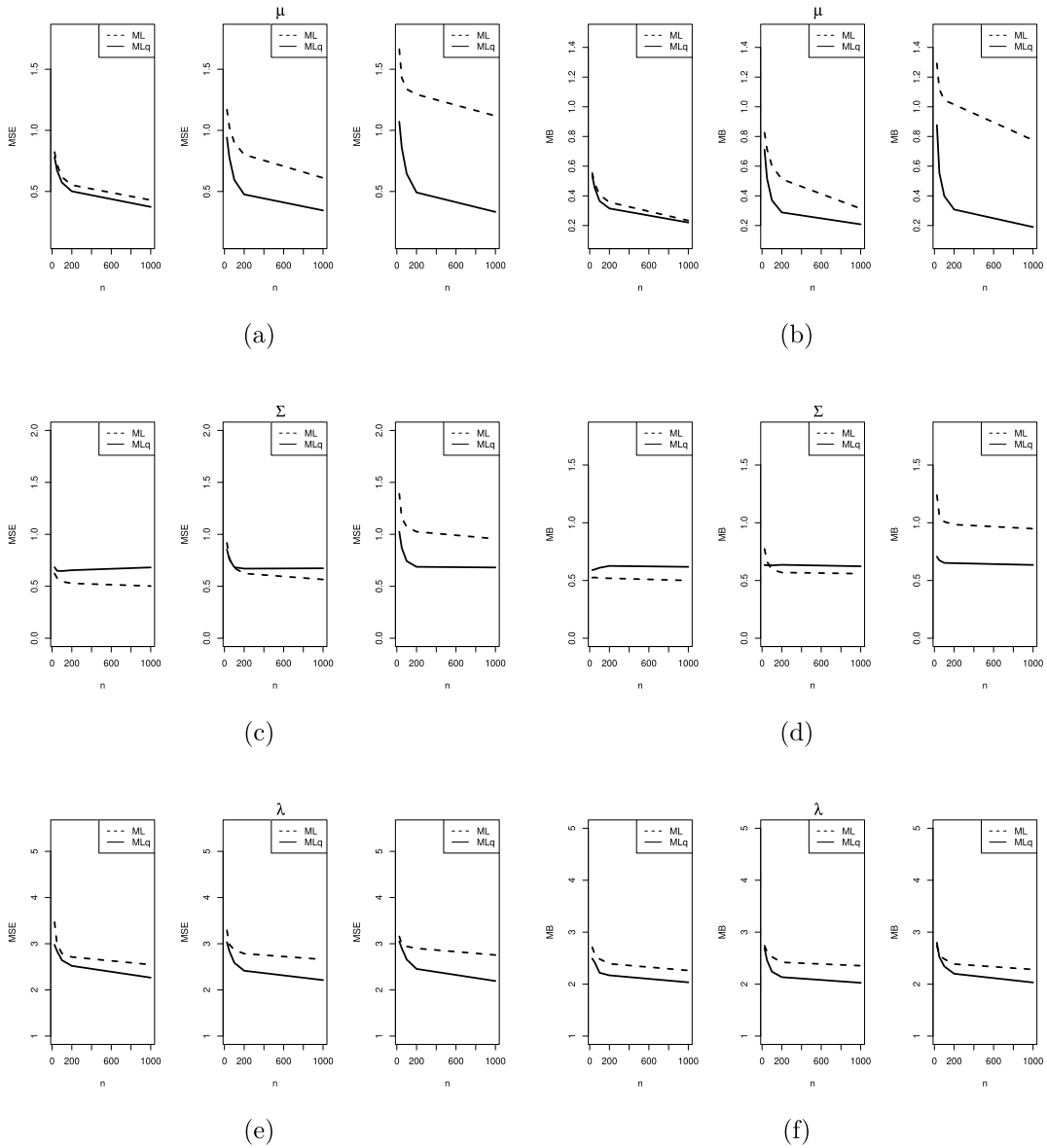
Fonte: Elaborado pela autora em 2020.

Figura A.3 – Cenário 2: Caso 1 - Média das medidas MSE e MB dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros μ , Σ e λ do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



Fonte: Elaborado pela autora em 2020.

Figura A.4 – Cenário 2: Caso 2 - Média das medidas MSE e MB dos estimadores de ML ($q = 1$) e de ML_q dos parâmetros μ , Σ e λ do modelo ST, com o adicional de 10% (primeira coluna), 20% (segunda coluna) e 30% (terceira coluna) de outliers.



Fonte: Elaborado pela autora em 2020.