



BRUNA DA COSTA SILVA

**TESTES COMPUTACIONALMENTE INTENSIVOS E
ASSINTÓTICOS ROBUSTOS PARA ESTRUTURA DE
SIMETRIA COMPOSTA**

LAVRAS - MG

2020

BRUNA DA COSTA SILVA

**TESTES COMPUTACIONALMENTE INTENSIVOS E ASSINTÓTICOS ROBUSTOS
PARA ESTRUTURA DE SIMETRIA COMPOSTA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para obtenção do título de Mestre.

Prof. Dr. Daniel Furtado Ferreira
Orientador

LAVRAS - MG

2020

**Ficha catalográfica elaborada pela Coordenadoria de Processos Técnicos
da Biblioteca Universitária da UFLA**

Silva, Bruna da Costa

Testes computacionalmente intensivos e assintóticos robustos para estrutura de simetria composta / Bruna da Costa Silva.

– Lavras : UFLA, 2020.

77 p. : il.

Dissertação (mestrado acadêmico)–Universidade Federal de Lavras, 2020.

Orientador: Prof. Dr. Daniel Furtado Ferreira.

Bibliografia.

1. Teste de hipóteses. 2. Robustez. 3. Covariâncias. I. Ferreira, Daniel Furtado. II. Título.

BRUNA DA COSTA SILVA

**TESTES COMPUTACIONALMENTE INTENSIVOS E ASSINTÓTICOS ROBUSTOS
PARA ESTRUTURA DE SIMETRIA COMPOSTA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para obtenção do título de Mestre.

APROVADA em 7 de Fevereiro de 2020.

Prof. Dr. Daniel Furtado Ferreira UFLA
Prof. Dr. Ben Dêivide de Oliveira Batista UFSJ
Prof. Dr. Júlio Sílvio de Sousa Bueno Filho UFLA

Prof. Dr. Daniel Furtado Ferreira
Orientador

**LAVRAS - MG
2020**

AGRADECIMENTOS

A Deus, por iluminar sempre meus caminhos.

Ao meus pais, Ivânia e Lázaro, e aos meus irmãos, Barbara e Bernardo, agradeço por todo amor e por sonharem sempre comigo, vocês são minha fortaleza.

Ao meu companheiro de vida Eduardo, pelo seu imenso apoio e companheirismo durante todos estes anos de estudo.

Ao professor Dr. Daniel Ferreira Furtado, por ter me orientado durante o mestrado, seu apoio, paciência e humildade foram essenciais para construção do presente trabalho, como também para a minha formação pessoal, foi uma incrível oportunidade trabalhar ao seu lado.

Aos meus amigos Luiz Otávio Pala, Marcio Aloisio, Renata Cintra, Marcel Irving, Laíla Luana Campos, por compartilharem seus conhecimentos, além dos ótimos momentos durante esta trajetória.

Aos membros da banca, Prof. Dr. Ben Dêivide de Oliveira Batista e Prof. Dr. Júlio Sílvio de Sousa Bueno Filho, agradeço a disponibilidade e por todas as dicas, correções e críticas, que também foram fundamentais.

Agradeço também à Universidade Federal de Lavras (UFLA), aos professores do departamento de Estatística e Experimentação Agropecuária que contribuíram na minha formação em mestre.

À CAPES, pela bolsa de estudo concedida. O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

RESUMO

A matriz de covariâncias com estrutura de simetria composta é utilizada em muitos procedimentos estatísticos, como na experimentação quando o experimento é realizado com parcelas repetidas no tempo, em modelos mistos lineares ou não lineares, como também em estudos genéticos. Para realizar inferências sobre a estrutura de simetria composta o teste de hipóteses baseado na razão do máximo das funções de verossimilhanças é comumente utilizado. Este teste possui como principal limitação a complexidade de estabelecer uma distribuição para estatística do teste sob a hipótese nula. Para isso, utiliza-se, uma distribuição assintótica para a estatística do teste, além de assumir normalidade multivariada dos dados amostrais, tornando-se um teste pouco robusto quando violada tal condição. Portanto, o objetivo deste trabalho é propor três testes robustos na presença de desvios de normalidade causados por *outliers*, sendo o primeiro um teste assintótico baseado em estimadores robustos (LRTR), o segundo, um teste computacionalmente intensivo baseado em estimadores robustos (BLRTR) e o terceiro, um teste computacionalmente intensivo baseado em estimadores clássicos (BLTRO). Comparar os desempenhos dos três testes propostos com o teste original da razão de verossimilhança (LRTO) e com o teste robusto (MCPT) criado por Morris, Payton e Santorico (2011). Além de verificar se os testes que utilizaram os estimadores robustos Comedian, LRTR e BLRTR, obtiveram melhores desempenhos do que os testes LRTO, BLRTO e MCPT ao considerar desvios de normalidade dos dados causados por *outliers*, como também no cenário ideal de dados normalmente distribuídos. Os testes foram avaliados por meio da taxa de erro tipo I e poder, em simulações Monte Carlo. Conclui-se que os estimadores robustos não agregaram melhorias aos testes assintóticos nem aos testes computacionalmente intensivos. Os testes BLRTO e BLRTR obtiveram melhores desempenhos em relação aos testes LRTO, LRTR e MCPT. Portanto, os testes BLRTR e BLRTO superaram o teste original da razão de verossimilhanças em dois aspectos: controlam a taxa de erro tipo I quando há presença de *outliers* nos dados e não utilizam distribuição assintótica para estatística do teste.

Palavras-chave: Teste de hipóteses. Robustez. Covariância

ABSTRACT

The covariance matrix with a compound symmetry structure is used in many statistical procedures, as in experimentation when the experiment is carried out with plots repeated in time, in mixed linear or non-linear models, as well as in genetic studies. To make inferences about the compound symmetry structure the hypothesis test based on the ratio of the maximum likelihood functions is commonly used. This test has as its main limitation the complexity of establishing a distribution for test statistics under the null hypothesis. For this, an asymptotic distribution is used for statistics of the test, in addition to assuming multivariate normality of the sample data, making it a weak test when this condition is violated. Therefore, the objective of this work is to propose three robust tests in the presence of deviations from normality caused by outliers, the first being an asymptotic test based on robust estimators (LRTR), the second, a computationally intensive test based on robust estimators (BLRTR) and the third, a computationally intensive test based on classical estimators (BLTRO). Comparing the performances of the three proposed tests with the original likelihood ratio test (LRTO) and with the robust test (MCPT) created by Morris, Payton e Santorico (2011). In addition to verifying that the tests that used the robust Co-median estimators, LRTR and BLRTR, obtained better performances than the LRTO, BLRTO and MCPT tests when considering deviations from normality of data caused by outliers, as well as in the ideal scenario of normally distributed data. The tests were evaluated using the type I error rate and power, in Monte Carlo simulations. It is concluded that the robust estimators did not add improvements to the asymptotic tests or to the computationally intensive tests. The BLRTO and BLRTR tests performed better than the LRTO, LRTR and MCPT tests. Therefore, the BLRTR and BLRTO tests outperformed the original likelihood ratio test in two ways: they control the type I error rate when there are outliers in the data and do not use asymptotic distribution for test statistics.

Keywords: Hypothesis testing. Robustness. Covariance

LISTA DE FIGURAS

Figura 4.1 – Primeiro cenário: Simulação do erro tipo I para os testes de simetria composta ($p = 3$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada, considerando respectivamente $\alpha = 0,1$, $\alpha = 0,05$ e $\alpha = 0,01$	53
Figura 4.2 – Segundo cenário: Simulação do erro tipo I para os testes de simetria composta ($n = 25$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada, considerando respectivamente $\alpha = 0,1$, $\alpha = 0,05$ e $\alpha = 0,01$	54
Figura 4.3 – Primeiro cenário: Simulação do erro tipo I para os testes de simetria composta ($p = 3$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\delta = 0,10$), considerando respectivamente $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$	55
Figura 4.4 – Segundo cenário: Simulação do erro tipo I para os testes de simetria composta ($n = 25$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\delta = 0,1$), considerando respectivamente $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$	56
Figura 4.5 – Primeiro cenário: Simulação do erro tipo I para os testes de simetria composta ($p = 3$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\delta = 0,3$), considerando respectivamente $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$	57
Figura 4.6 – Segundo cenário: Simulação do erro tipo I para os testes de simetria composta ($n = 25$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\delta = 0,3$), considerando respectivamente $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$	57
Figura 4.7 – Primeiro cenário (Σ_H): Simulação do poder para os testes de simetria composta ($p = 3$; $d = 4$; $\rho = 0$) respectivamente para as distribuições normal multivariada contaminada ($\delta = 0,30$), normal multivariada contaminada ($\delta = 0,10$), normal multivariada contaminada ($\delta = 0,05$) e normal multivariada, considerando $\alpha = 0,05$	59

Figura 4.8 – Segundo cenário (Σ_{AP}): Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) respectivamente para as distribuições normal multivariada contaminada ($\delta = 0,30$), normal multivariada contaminada ($\delta = 0,10$), normal multivariada contaminada ($\delta = 0,05$) e normal multivariada, considerando $\alpha = 0,05$ 60

LISTA DE TABELAS

Tabela 4.1 – Segundo cenário (Σ_{AP}): Simulação do poder para os testes de simetria composta ($p = 20$; $n = 25$; $\sigma^2 = 9$ e $\rho = 0,6$) para as distribuições normal multivariada, normal multivariada contaminada ($\delta = 0,10$) e normal multivariada contaminada ($\delta = 0,30$), considerando $\alpha = 0,05$	61
Tabela 4.2 – Primeiro cenário (Σ_H): Simulação do poder para os testes de simetria composta ($p = 20$; $n = 25$; $d = 4$ e $\rho = 0$) respectivamente para as distribuições normal multivariada, normal multivariada contaminada ($\delta = 0,10$) e normal multivariada contaminada ($\delta = 0,30$), considerando $\alpha = 0,05$	62
Tabela 4.3 – Primeiro cenário (Σ_H): Simulação do poder para os testes de simetria composta ($p = 20$; $n = 25$; $d = 10$ e 20 e $\rho = 0$) para a distribuição normal multivariada, considerando $\alpha = 0,05$	62
Tabela A.1 – Primeiro cenário: Simulação do erro tipo I para o teste de simetria composta ($p = 3$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\delta = 0,05$), considerando respectivamente $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$	66
Tabela A.2 – Segundo cenário: Simulação do erro tipo I para o teste de simetria composta ($n = 25$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\delta = 0,05$), considerando respectivamente $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$	66
Tabela A.3 – Primeiro cenário: Simulação do poder para os testes de simetria composta ($p = 3$; $d = 4$; $\rho = 0$) com a distribuição normal multivariada contaminada ($\mu_1 = [0, 0, \dots, 0]^\top$, $\mu_2 = [10, 10, \dots, 10]^\top$, $\Sigma_1 = \Sigma_H$ e $\Sigma_2 = 2\Sigma_H$), considerando $\alpha = 0,10$, $\alpha = 0,01$ e $\delta = 0,30$	67
Tabela A.4 – Primeiro cenário: Simulação do poder para os testes de simetria composta ($p = 3$; $d = 4$; $\rho = 0$) com a distribuição normal multivariada contaminada ($\mu_1 = [0, 0, \dots, 0]^\top$, $\mu_2 = [10, 10, \dots, 10]^\top$, $\Sigma_1 = \Sigma_H$ e $\Sigma_2 = 2\Sigma_H$), considerando $\alpha = 0,10$, $\alpha = 0,01$ e $\delta = 0,10$	67
Tabela A.5 – Primeiro cenário: Simulação do poder para os testes de simetria composta ($p = 3$; $d = 4$; $\rho = 0$) com a distribuição normal multivariada contaminada ($\mu_1 = [0, 0, \dots, 0]^\top$, $\mu_2 = [10, 10, \dots, 10]^\top$, $\Sigma_1 = \Sigma_H$ e $\Sigma_2 = 2\Sigma_H$), considerando $\alpha = 0,10$, $\alpha = 0,01$ e $\delta = 0,05$	67

Tabela A.6 – Primeiro cenário: Simulação do poder para os testes de simetria composta ($p = 3$; $d = 4$; $\rho = 0$) com a distribuição normal multivariada ($\boldsymbol{\mu} = [0, 0, \dots, 0]^\top$, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_H$), considerando $\alpha = 0,10$ e $\alpha = 0,01$	68
Tabela A.7 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = 2\boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$, $\alpha = 0,01$ e $\delta = 0,30$	68
Tabela A.8 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = 2\boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$, $\alpha = 0,01$ e $\delta = 0,10$	68
Tabela A.9 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = 2\boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$, $\alpha = 0,01$ e $\delta = 0,05$	69
Tabela A.10 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada ($\boldsymbol{\mu} = [0, 0, \dots, 0]^\top$ e $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$ e $\alpha = 0,01$	69
Tabela A.11 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$, $0,01$, $0,05$ e $\delta = 0,30$	69
Tabela A.12 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$, $0,01$, $0,05$ e $\delta = 0,10$	70
Tabela A.13 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$, $0,01$, $0,05$ e $\delta = 0,05$	70

SUMÁRIO

1	Introdução	11
2	Referencial Teórico	14
2.1	Teste de hipóteses	14
2.2	Covariância e correlação	15
2.3	Teste de hipóteses original para matriz de covariâncias com simetria composta (LRTO)	17
2.3.1	Máximo da função de verossimilhança no espaço paramétrico irrestrito	18
2.3.2	Máximo da função de verossimilhança no espaço paramétrico restrito	21
2.3.3	Estatística do teste	28
2.3.4	Estudos sobre o teste original da matriz de covariâncias com simetria composta (LRTO)	29
2.4	<i>Outliers</i> Multivariados	32
2.5	Estatística robusta	33
2.5.1	Estimador robusto Comedian	34
2.6	Método Monte Carlo	36
2.7	<i>Bootstrap</i>	37
2.8	Avaliação dos testes de hipóteses	38
2.9	Distribuições multivariadas	39
2.9.1	Distribuição normal multivariada	41
2.9.2	Distribuição normal multivariada contaminada	42
3	Métodos	43
3.1	Teste de hipóteses	43
3.1.1	Teste da razão de verossimilhanças original (LRTO)	44
3.1.2	Teste da razão de verossimilhanças baseado em estimadores robustos (LRTR)	44
3.1.3	Teste <i>bootstrap</i> paramétrico baseado em estimadores robustos (BLRTR)	45
3.1.4	Teste <i>bootstrap</i> paramétrico baseado em estimadores clássicos (BLRTO)	46
3.2	Avaliação do desempenho dos testes	46
3.2.1	Erro tipo I	47
3.2.2	Poder do teste	48
4	Resultados e discussão	51
4.1	Taxa de erro tipo I dos testes	51

4.2	Poder dos testes	58
5	Conclusões	63
A	APÊNDICE - Resultados da simulação	66
B	APÊNDICE - Comandos usados no R	71

1 INTRODUÇÃO

Ao considerar uma amostra aleatória de tamanho n , $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ no hiperespaço de uma população normal multivariada, é possível realizar inferências sobre os parâmetros populacionais: vetor de médias $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$. A matriz de covariâncias é uma matriz quadrada simétrica, composta por variâncias e covariâncias associadas as p variáveis, expressando como as p variáveis se relacionam.

Existem diversas estruturas para as matrizes de covariâncias como a estrutura autor-regressiva de primeira ordem, em que a correlação entre as observações tomadas em tempos próximos é maior que quando tomadas em tempos mais distantes. A estrutura de independência é representada por possuir covariâncias nulas e variâncias diferentes. A estrutura de simetria composta, possui correlações iguais (ρ) e as variâncias homogêneas (σ^2). Já a matriz com estrutura de esfericidade considera covariâncias nulas e variâncias homogêneas, o que não deixa de ser um caso de simetria composta em que as covariâncias são iguais a zero.

Este estudo aborda a matriz de covariâncias com estrutura de simetria composta, cuja estrutura de covariâncias é utilizada em muitos procedimentos estatísticos. Uma das aplicações desta matriz é na experimentação quando o experimento é realizado com parcelas repetidas no tempo, pois casualiza-se as parcelas e aloca as subparcelas de acordo com o tempo, considerando as medidas no tempo como os tratamentos das subparcelas. Neste tipo de experimento, o teste F, em relação à parcela, tem distribuição F exata, mas em relação à subparcela, terá distribuição F exata se a matriz de covariâncias dos erros tiver estrutura de simetria composta, sendo esta uma condição suficiente.

Uma condição mais geral para a estrutura de covariação em relação as subparcelas repetidas no tempo foi criada por Huynh e Feldt (1970), para que a análise de variância tenha distribuição F exata em relação as subparcelas. Esta condição, denotada de H-F, é equivalente a especificar que variâncias da diferença entre pares de erros sejam todas iguais, podendo ser verificada ao calcular todas as variâncias das diferenças dos possíveis pares de erros. Se as variâncias das diferenças forem iguais a uma constante para todo par, esta matriz é considerada do tipo H-F, e se as variâncias desta matriz forem todas homogêneas, então esta condição é equivalente à de simetria composta.

Ainda no contexto da experimentação com medidas repetidas ao longo do tempo, os modelos mistos lineares ou não lineares são muito utilizados, pois possuem uma parte aleatória para explicar a estrutura de covariância dessas medidas, além de realizar análises com dados

incompletos e desbalanceados. A estrutura de simetria composta é muito utilizada neste tipo de modelo, representando a estrutura de covariância da variável resposta. No estudo de Ueda et al. (2003), a matriz de covariâncias com estrutura de simetria composta foi a que obteve melhor resultado para verificar a proporção de severidade da epidemia míldio-da-batateira em quatro cultivares de batata, em um modelo não linear misto.

A matriz de covariâncias com estrutura de simetria composta também pode ser abordada fora do contexto da experimentação, como por exemplo em um estudo genético realizado por Morris, Payton e Santorico (2011), que identificou que três genes específicos (*NM_001081*, *NM_002125* e *V00522*) que possuem ligação no mesmo local do cromossomo 6, tinham esta estrutura de correlação.

Para realizar inferências sobre a estrutura de simetria composta o teste de hipóteses baseado na razão do máximo das funções de verossimilhanças é comumente utilizado. Neste caso, sua hipótese nula é que a matriz de covariâncias populacional é igual a uma matriz com estrutura de simetria composta, e na hipótese alternativa, a estrutura de covariação da matriz de covariâncias populacional é geral.

Este teste possui como principal limitação a complexidade de estabelecer uma distribuição para estatística do teste sob a hipótese nula. Para isso, utiliza-se, uma distribuição assintótica para a estatística do teste. Além do mais, para aplicar este teste, assume-se normalidade multivariada dos dados amostrais. Sendo assim, o teste de hipóteses para matriz de covariâncias com simetria composta é pouco robusto na presença de desvios de normalidade dos dados (FERREIRA, 2018).

Morris, Payton e Santorico (2011) criaram um teste MCPT, não paramétrico para simetria composta, por meio do uso da permutação, com intuito de obter robustez na presença de desvios de normalidade, considerando que os desvios eram causados por dados provenientes de outras distribuições de probabilidade. O presente estudo irá focar nos desvios de normalidade, mas diferente do estudo de Morris, Payton e Santorico (2011), estes desvios terão como causa valores discrepantes, denominados de *outliers*. Assim, serão criados testes robustos sobre este aspecto. Isto é, mesmo que hajam violações das pressuposições de normalidade pela presença de *outliers*, os desempenhos dos testes potencialmente continuam bom ou muito bom, quando comparado com os desempenhos em situações ótimas para suas aplicações.

Portanto, o objetivo deste trabalho é propor três testes robustos, sendo o primeiro um teste assintótico baseado em estimadores robustos (LRTR), o segundo, um teste computacio-

nalmente intensivo baseado em estimadores robusto (BLRTR) e o terceiro, um teste computacionalmente intensivo baseado em estimadores clássicos (BLTRO). Comparar os desempenhos dos três testes propostos com o do teste original da razão de verossimilhanças (LRTO) e com o do teste robusto (MCPT) criado por Morris, Payton e Santorico (2011). Além de verificar se os testes que utilizaram os estimadores robustos Comedian, LRTR e BLRTR, obtiveram melhores desempenhos do que os testes LRTO, BLRTO e MCPT ao considerar desvios de normalidade dos dados causados por *outliers*, como também no cenário ideal de dados normalmente distribuídos.

Este estudo foi subdividido em: seção 1 introdução; seção 2 com o referencial teórico, descrevendo os principais conceitos estatísticos sobre o teste de hipóteses para matriz de covariâncias com simetria composta, juntamente com os principais conceitos necessários para realizar os testes de hipóteses propostos; seção 3 metodológica, descrevendo todos os passos necessários para a realização dos testes e avaliação de seus desempenhos; seção 4, discussão dos resultados obtidos; seção 5, conclusão; e os apêndices com os demais resultados das simulações e comandos usados do software R.

2 REFERENCIAL TEÓRICO

A seção a seguir apresentará o desenvolvimento teórico do teste de hipóteses original sobre a matriz de covariâncias com simetria composta, baseado no teste da razão de verossimilhanças, apresentando posteriormente suas principais limitações e estudos já realizados sobre este assunto. Em seguida, são exibidos os principais conceitos utilizados para realizar os testes de hipóteses assintóticos e computacionalmente intensivos baseados em estimadores robustos e clássicos propostos neste trabalho.

2.1 Teste de hipóteses

O teste de hipóteses é um método inferencial que permite fazer afirmações e tomar decisões a respeito de parâmetros populacionais, o qual é constituído essencialmente de duas hipóteses. A hipótese nula H_0 , representando o que pretende-se testar, ou seja, a afirmação realizada sobre o vetor de parâmetros $\boldsymbol{\theta}$, correspondendo ao espaço paramétrico restrito $\boldsymbol{\theta} \in \Omega_0$. Já na hipótese alternativa H_1 , o vetor de parâmetros $\boldsymbol{\theta}$ é geral, correspondendo ao espaço paramétrico irrestrito $\boldsymbol{\theta} \in \Omega$ (FERREIRA, 2018). Assim, as hipóteses nula e alternativa são:

$$\begin{cases} H_0 : \boldsymbol{\theta} \in \Omega_0 \\ H_1 : \boldsymbol{\theta} \in \Omega. \end{cases}$$

O espaço paramétrico restrito, Ω_0 , corresponde as restrições impostas ao realizar uma afirmação sobre o parâmetro populacional. Já o espaço paramétrico irrestrito, Ω , representa todo o espaço paramétrico sem restrição sobre $\boldsymbol{\theta}$.

Portanto, o objetivo do teste é afirmar se $\boldsymbol{\theta} \in \Omega_0$ ou se $\boldsymbol{\theta} \in \Omega$. Como critério de decisão, rejeita-se H_0 em favor de H_1 , se o resultado da estatística do teste pertencer a região de rejeição R , representado o conjunto de valores assumidos pela estatística do teste para os quais a hipótese nula é rejeitada, ou seja:

$$\sup_{\boldsymbol{\theta} \in \Omega_0} P(\mathbf{X} \in R; \boldsymbol{\theta}) = \alpha,$$

em que $\mathbf{X}_{n \times p}$ é a matriz de dados e α é o nível de significância do teste, que assume valores entre $[0,1]$, dado que a região de rejeição é baseada na premissa do erro tipo I fixado em α , que corresponde $P(\text{Rejeitar } H_0 | H_0 \text{ verdadeira}) = \alpha$ (FERREIRA, 2018).

Para obter a região de rejeição é necessário identificar a distribuição da estatística do teste Λ , que representa a razão do máximo das funções de verossimilhanças, sendo dada por:

$$\Lambda = \frac{L_{\Omega_0}(\mathbf{X}; \hat{\boldsymbol{\theta}})}{L_{\Omega}(\mathbf{X}; \hat{\boldsymbol{\theta}})},$$

cujos numerador é o máximo da função de verossimilhança do espaço paramétrico restrito, representando a probabilidade máxima da amostra observada, sendo o máximo calculado em relação aos parâmetros do modelo subjacente restrito às condições impostas pela hipótese nula. O denominador é o máximo da função de verossimilhança do espaço paramétrico irrestrito, representando a probabilidade máxima de ter observado a amostra em questão em relação aos parâmetros sem qualquer restrição a eles (CASELLA; BERGER, 2014).

No entanto, devido a dificuldade de estabelecer a distribuição nula de Λ , Ferreira (2018) comenta que se $\Omega_0 \subset \Omega$, com $\Omega_0 \subset \mathbb{R}^s$ e $\Omega \subset \mathbb{R}^r$, sob determinadas condições de regularidade $-2\ln(\Lambda)$ possui distribuição assintoticamente qui-quadrado com $r - s$ graus de liberdade. A região de rejeição é obtida por: $R = \{\mathbf{X} | \lambda = -2\ln[\Lambda(\mathbf{X})] > \chi_{\alpha, r-s}^2\}$ em que $\chi_{\alpha, r-s}^2$ é o quantil superior 100 $\alpha\%$ da distribuição qui-quadrado com $r - s$ graus de liberdade. Como a distribuição é assintoticamente qui-quadrado, a região de rejeição terá tamanho assintoticamente igual a α .

2.2 Covariância e correlação

De acordo com Casella e Berger (2014), a força da relação entre duas variáveis aleatórias X e Y pode ser medida por meio da covariância e da correlação. Assumindo que $E(X) = \mu_X$, $E(Y) = \mu_Y$, $VAR(X) = \sigma_X^2$ e $VAR(Y) = \sigma_Y^2$, a covariância de X e Y e a correlação de X e Y são definidas, respectivamente, por:

$$COV(X, Y) = E((X - \mu_X)(Y - \mu_Y)),$$

$$\rho_{X, Y} = \frac{COV(X, Y)}{\sigma_Y \sigma_X}.$$

Se grandes valores de X tendem a ser observados com grandes valores de Y e pequenos valores de X tendem a ser observados com pequenos valores de Y , então a $COV(X, Y)$ terá valor positivo. Por outro lado, se grandes valores de X tendem a ser observados com pequenos valores de Y e pequenos valores de X tendem a ser observados com grandes valores de Y ,

então a $COV(X,Y)$ terá valor negativo, fornecendo assim informações sobre a relação de X e Y (CASELLA; BERGER, 2014).

O valor resultante da $COV(X,Y)$ pode ser qualquer número da escala real, não fornecendo assim a força da relação entre X e Y . Já a correlação, sempre resulta em valores no intervalo $[-1,1]$. Se o resultado da correlação for -1 ou 1 indicará uma perfeita correlação linear entre X e Y (CASELLA; BERGER, 2014).

De acordo com Casella e Berger (2014), se X e Y são variáveis aleatórias independentes, tem-se que $COV(X,Y) = \rho_{X,Y} = 0$. Entretanto, a covariância ou a correlação medem somente um tipo particular de relação, que é linear. A recíproca não é verdadeira, sendo que duas variáveis aleatórias podem ter uma forte relação, mas a covariância ou correlação será zero, porque a relação não é linear.

De acordo com Ferreira (2018), na estatística multivariada analisa-se p variáveis em n unidades amostrais, em que as n unidades amostrais são independentes, mas entre as p variáveis existem correlação. Desta forma trabalha-se com uma matriz X de dados multivariados dada por:

$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix}.$$

As linhas da matriz de dados representam vetores aleatórios p -dimensionais independentes. Se um vetor qualquer (linha da matriz de dados) for representado por \mathbf{X} , então sua matriz de covariâncias é quadrada de dimensões $(p \times p)$, positiva definida e expressa por:

$$\begin{aligned} COV(\mathbf{X}) &= \begin{bmatrix} E(X_1 - \mu_1)^2 & \dots & E(X_1 - \mu_1)(X_p - \mu_p) \\ E(X_2 - \mu_2)(X_1 - \mu_1) & \dots & E(X_2 - \mu_2)(X_p - \mu_p) \\ \vdots & \ddots & \vdots \\ E(X_p - \mu_p)(X_1 - \mu_1) & \dots & E(X_p - \mu_p)^2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix}. \end{aligned}$$

Na diagonal da matriz $COV(\mathbf{X})$, obtém-se a covariância de uma variável com ela mesma, que é de fato a variância da variável correspondente.

Ao utilizar-se de uma amostra aleatória multivariada, $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ de vetores p -dimensionais, a matriz de covariâncias amostral é expressa por:

$$\mathbf{S} = \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1p} \\ S_{21} & S_{22} & \dots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \dots & S_{pp} \end{bmatrix}, \quad (2.1)$$

dado que $l, k = 1, 2, \dots, p$ e $\bar{X}_{.k}$ é a média amostral da k -ésima variável, calcula-se o k -ésimo elemento da diagonal de \mathbf{S} por:

$$S_{kk} = \frac{1}{n-1} \sum_{j=1}^n (X_{jk} - \bar{X}_{.k})^2,$$

a covariância amostral entre a k -ésima e l -ésima variável por:

$$S_{kl} = \frac{1}{n-1} \sum_{j=1}^n (X_{jk} - \bar{X}_{.k})(X_{jl} - \bar{X}_{.l}).$$

2.3 Teste de hipóteses original para matriz de covariâncias com simetria composta (LRTO)

Wilks et al. (1946) foram os primeiros a desenvolver critérios para realizar o teste de simetria composta, por meio do teste da razão de verossimilhanças, que sob a hipótese nula, obteve uma distribuição aproximada da estatística do teste para grandes amostras. Para realização deste teste considera-se inicialmente uma amostra aleatória de tamanho n dada por $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_i, \dots, \mathbf{X}_n$, em que \mathbf{X}_i é uma variável aleatória p -dimensional, retirada de uma população com distribuição normal multivariada, com parâmetros $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$, sendo ambos desconhecidos. Portanto, as hipóteses a serem testadas são $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ versus $H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0$, dado que:

$$\boldsymbol{\Sigma}_0 = \sigma^2 \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{bmatrix} = \sigma^2 [(1-\rho)\mathbf{I} + \rho\mathbf{J}], \quad (2.2)$$

em que \mathbf{J} é uma matriz com dimensões $(p \times p)$ e composta por valores iguais a 1 e \mathbf{I} é uma matriz identidade com dimensões $(p \times p)$ (FERREIRA, 2018).

Dado que o teste é realizado utilizando-se de uma amostra aleatória de tamanho n , então, sob a hipótese nula, o estimador não viesado de Σ é:

$$\mathbf{S}_0 = S_c^2 \begin{bmatrix} 1 & r & \dots & r \\ r & 1 & \dots & r \\ \vdots & \vdots & \ddots & \vdots \\ r & r & \dots & 1 \end{bmatrix} = S_c^2[(1-r)\mathbf{I} + r\mathbf{J}], \quad (2.3)$$

em que, S_c^2 é o estimador não viesado da variância populacional σ^2 e r é o estimador comum não viesado da correlação populacional ρ , apresentados respectivamente pelas expressões (2.4) e (2.5):

$$S_c^2 = \frac{1}{p} \sum_{k=1}^p S_{kk}, \quad (2.4)$$

$$r = \frac{2}{p(p-1)S_c^2} \sum_{i=1}^{p-1} \sum_{k=i+1}^p S_{ik}. \quad (2.5)$$

Para um melhor entendimento sobre a construção do teste de hipóteses, será abordado inicialmente a obtenção do máximo da função de verossimilhança com espaço paramétrico irrestrito $L_{\Omega}(\mathbf{X}; \hat{\boldsymbol{\mu}}; \hat{\Sigma})$ e restrito $L_{\Omega_0}(\mathbf{X}; \hat{\boldsymbol{\mu}}; \Sigma_0)$. Em seguida, será obtida a estatística do teste (Λ) da razão de verossimilhanças e apresentada a correção de Box (1949) para $-2\ln(\Lambda)$, com a finalidade de se obter uma melhor aproximação assintótica da estatística à distribuição qui-quadrado, além de outros estudos realizados com intuito de aprimorar o teste já existente.

2.3.1 Máximo da função de verossimilhança no espaço paramétrico irrestrito

De acordo com Ferreira (2018), para calcular o máximo da função de verossimilhança no espaço paramétrico irrestrito, considera-se inicialmente a distribuição conjunta das n observações p -dimensionais, de uma amostra aleatória normalmente distribuída. A distribuição conjunta é obtida por meio da função de verossimilhança calculada pelo produto das densidades, $f(\mathbf{x})$, de cada observação p -dimensional independente e identicamente distribuída. Entretanto,

é considerada a função de verossimilhança aleatória (MITTELHAMMER, 2013) dada por:

$$L_{\Omega}(\mathbf{X}; \boldsymbol{\mu}; \boldsymbol{\Sigma}) = \prod_{i=1}^n f_{\mathbf{X}}(\mathbf{X}_i) = (2\pi)^{-\frac{np}{2}} |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})^{\top} \boldsymbol{\Sigma}^{-1} (\mathbf{X}_i - \boldsymbol{\mu}) \right\}. \quad (2.6)$$

Diante da função de verossimilhança é necessário obter os estimadores $\hat{\boldsymbol{\mu}}$ e $\hat{\boldsymbol{\Sigma}}$ de forma que a maximize. Para isso, realiza-se derivadas parciais de primeira ordem na função de verossimilhança em relação aos dois parâmetros. Em seguida, as funções obtidas serão igualadas a zero para resolver o sistema de equações e assim obter os estimadores de máxima verossimilhança (FERREIRA, 2018).

Aplica-se o logaritmo na função de verossimilhança para facilitar a obtenção das derivadas parciais, chamando assim de função suporte $g(\mathbf{X}; \boldsymbol{\mu}; \boldsymbol{\Sigma})$ ou função log-verossimilhança. As soluções em relação aos estimadores que levam aos máximos das funções de verossimilhança e da função suporte não diferem, dado que a função logaritmo é monótona e crescente (FERREIRA, 2018). Assim,

$$\begin{aligned} g_{\Omega}(\mathbf{X}; \boldsymbol{\mu}; \boldsymbol{\Sigma}) &= \ln[L(\mathbf{X}; \boldsymbol{\mu}; \boldsymbol{\Sigma})] = \ln \left[\prod_{i=1}^n f_{\mathbf{X}}(\mathbf{X}_i) \right] \\ &= -\frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln|\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})^{\top} \boldsymbol{\Sigma}^{-1} (\mathbf{X}_i - \boldsymbol{\mu}). \end{aligned} \quad (2.7)$$

Toma-se a derivada parcial de primeira ordem em relação a $\boldsymbol{\mu}$:

$$\frac{\partial g_{\Omega}(\mathbf{X}; \boldsymbol{\mu}; \boldsymbol{\Sigma})}{\partial \boldsymbol{\mu}} = \sum_{i=1}^n \boldsymbol{\Sigma}^{-1} (\mathbf{X}_i - \boldsymbol{\mu}),$$

igualando-se a zero e obtém-se o estimador de máxima verossimilhança para o vetor de médias, resultando em:

$$\hat{\boldsymbol{\mu}} = \frac{\sum_{i=1}^n \mathbf{X}_i}{n} = \bar{\mathbf{X}}. \quad (2.8)$$

De acordo com Ferreira (2018), algumas modificações são feitas em relação ao expoente exp na função de verossimilhança para facilitar a derivada em relação ao parâmetro Σ :

$$\begin{aligned} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{X}_i - \boldsymbol{\mu}) &= \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}}. + \bar{\mathbf{X}}. - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}. + \bar{\mathbf{X}}. - \boldsymbol{\mu}) \\ &= \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}}.)^\top \Sigma^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}.) + \\ &\quad + 2 \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}}.)^\top \Sigma^{-1} (\bar{\mathbf{X}}. - \boldsymbol{\mu}) + \\ &\quad + \sum_{i=1}^n (\bar{\mathbf{X}}. - \boldsymbol{\mu})^\top \Sigma^{-1} (\bar{\mathbf{X}}. - \boldsymbol{\mu}), \end{aligned}$$

em seguida, aplica-se a propriedade do traço, $tr(\mathbf{AB}) = tr(\mathbf{BA})$, dado que os elementos da soma são escalares:

$$\begin{aligned} \sum_{i=1}^n tr \left[\Sigma^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}.) (\mathbf{X}_i - \bar{\mathbf{X}}.)^\top \right] + 2 \sum_{i=1}^n tr \left[\Sigma^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}.) (\bar{\mathbf{X}}. - \boldsymbol{\mu})^\top \right] + \\ + \sum_{i=1}^n tr \left[\Sigma^{-1} (\bar{\mathbf{X}}. - \boldsymbol{\mu}) (\bar{\mathbf{X}}. - \boldsymbol{\mu})^\top \right], \end{aligned}$$

ao abordar a propriedade, traço da soma é a soma do traço, considerar que $2 \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}}.) (\bar{\mathbf{X}}. - \boldsymbol{\mu})^\top$, é uma matriz de zeros e que $(n-1)\mathbf{S} = \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}}.) (\mathbf{X}_i - \bar{\mathbf{X}}.)^\top$, a função suporte ficará:

$$\begin{aligned} g_{\Omega}(\mathbf{X}; \boldsymbol{\mu}; \Sigma) &= -\frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln|\Sigma| - \\ &\quad - \frac{1}{2} tr \left[(n-1) \Sigma^{-1} \mathbf{S} \right] - \frac{1}{2} tr \left[n \Sigma^{-1} (\bar{\mathbf{X}}. - \boldsymbol{\mu}) (\bar{\mathbf{X}}. - \boldsymbol{\mu})^\top \right]. \end{aligned} \quad (2.9)$$

Pode-se observar que $tr \left[n \Sigma^{-1} (\bar{\mathbf{X}}. - \boldsymbol{\mu}) (\bar{\mathbf{X}}. - \boldsymbol{\mu})^\top \right] = n (\bar{\mathbf{X}}. - \boldsymbol{\mu})^\top \Sigma^{-1} (\bar{\mathbf{X}}. - \boldsymbol{\mu}) \geq 0$, sendo a igualdade alcançada quando o estimador de $\boldsymbol{\mu}$ for igual a $\bar{\mathbf{X}}.$, este resultado é recorrente do fato de se tratar de uma forma quadrática tendo como núcleo uma matriz positiva definida (FERREIRA, 2018). Desta forma, o máximo da $g_{\Omega}(\mathbf{X}; \boldsymbol{\mu}; \Sigma)$ é obtido ao substituir $\boldsymbol{\mu}$ pela seu estimador de máxima verossimilhança, independentemente do valor da matriz simétrica positiva definida Σ .

Tomando a derivada de primeira ordem na função suporte modificada (2.9), em relação a Σ , de forma a maximizá-la em relação a este parâmetro obtém-se:

$$\begin{aligned} \frac{\partial g_{\Omega}(\mathbf{X}; \boldsymbol{\mu}; \Sigma)}{\partial \Sigma} &= -\frac{n}{2}(\Sigma^{-1})^{\top} + \frac{(n-1)}{2}\Sigma^{-1}\mathbf{S}\Sigma^{-1} + \frac{n}{2}\Sigma^{-1}(\bar{\mathbf{X}} \cdot - \boldsymbol{\mu})(\bar{\mathbf{X}} \cdot - \boldsymbol{\mu})^{\top}\Sigma^{-1} \\ &= -\frac{n}{2}\Sigma^{-1} + \frac{(n-1)}{2}\Sigma^{-1}\mathbf{S}\Sigma^{-1} + \frac{n}{2}\Sigma^{-1}(\bar{\mathbf{X}} \cdot - \boldsymbol{\mu})(\bar{\mathbf{X}} \cdot - \boldsymbol{\mu})^{\top}\Sigma^{-1}, \end{aligned} \quad (2.10)$$

em seguida, a expressão (2.10) é igualada a zero e substitui-se o estimador de $\boldsymbol{\mu}$, que já foi calculado na expressão (2.8), para então obter o estimador de máxima verossimilhança para Σ :

$$\begin{aligned} -\frac{n}{2}\hat{\Sigma}^{-1} + \frac{(n-1)}{2}\hat{\Sigma}^{-1}\mathbf{S}\hat{\Sigma}^{-1} + \frac{n}{2}\hat{\Sigma}^{-1}(\bar{\mathbf{X}} \cdot - \hat{\boldsymbol{\mu}})(\bar{\mathbf{X}} \cdot - \hat{\boldsymbol{\mu}})^{\top}\hat{\Sigma}^{-1} &= 0 \\ -n\hat{\Sigma}^{-1} + (n-1)\hat{\Sigma}^{-1}\mathbf{S}\hat{\Sigma}^{-1} + n\hat{\Sigma}^{-1}(\bar{\mathbf{X}} \cdot - \bar{\mathbf{X}} \cdot)(\bar{\mathbf{X}} \cdot - \bar{\mathbf{X}} \cdot)^{\top}\hat{\Sigma}^{-1} &= 0 \\ -n\hat{\Sigma}^{-1} + (n-1)\hat{\Sigma}^{-1}\mathbf{S}\hat{\Sigma}^{-1} &= 0 \\ -n\hat{\Sigma}\hat{\Sigma}^{-1}\hat{\Sigma} + (n-1)\hat{\Sigma}\hat{\Sigma}^{-1}\mathbf{S}\hat{\Sigma}^{-1}\hat{\Sigma} &= 0 \\ n\hat{\Sigma} &= (n-1)\mathbf{S} \\ \hat{\Sigma} &= \frac{(n-1)\mathbf{S}}{n} = \mathbf{S}_n. \end{aligned} \quad (2.11)$$

Após encontrar os estimadores de máxima verossimilhança do vetor de médias e da matriz de covariâncias, o máximo da função de verossimilhança com espaço paramétrico irrestrito é dado por:

$$\begin{aligned} L_{\Omega}(\mathbf{X}; \hat{\boldsymbol{\mu}}; \hat{\Sigma}) &= (2\pi)^{-\frac{np}{2}} |\mathbf{S}_n|^{-\frac{n}{2}} \exp \left\{ \frac{-tr[n\mathbf{S}_n^{-1}\mathbf{S}_n] - tr[n\mathbf{S}_n^{-1}(\bar{\mathbf{X}} \cdot - \bar{\mathbf{X}} \cdot)(\bar{\mathbf{X}} \cdot - \bar{\mathbf{X}} \cdot)^{\top}]}{2} \right\} \\ &= (2\pi)^{-\frac{np}{2}} |\mathbf{S}_n|^{-\frac{n}{2}} \exp \left\{ -\frac{np}{2} \right\}. \end{aligned} \quad (2.12)$$

2.3.2 Máximo da função de verossimilhança no espaço paramétrico restrito

De acordo com Ferreira (2018), para calcular o máximo da função de verossimilhança com espaço paramétrico restrito, inicialmente é preciso obter os autovalores e autovetores da matriz de covariâncias sob H_0 , com estrutura de simetria composta Σ_0 , realizados por meio do

método iterativo *power method*, cujos resultados são:

$$\begin{aligned} \lambda_1 &= \sigma^2[1 + (p-1)\rho] \text{ e } \lambda_k = \sigma^2(1-\rho) && \text{para } 2 \leq k \leq p, \\ \mathbf{e}_1 &= \frac{1}{\sqrt{p}}[1, 1, \dots, 1]^\top \text{ e } \mathbf{e}_k = \frac{1}{\sqrt{k(k-1)}}[1, \dots, 1, -(k-1), 0, \dots, 0]^\top && \text{para } 2 \leq k \leq p, \end{aligned}$$

em que, λ_1 é o primeiro autovalor, λ_k é o k -ésimo autovalor, \mathbf{e}_1 é o primeiro autovetor e \mathbf{e}_k é o k -ésimo autovetor.

Em seguida, a função de verossimilhança do espaço paramétrico restrito sob H_0 passará por algumas modificações por meio da decomposição espectral dos autovalores e autovetores, cujo intuito é facilitar a obtenção dos estimadores de máxima verossimilhança, para então obter seu máximo (FERREIRA, 2018). No espaço paramétrico restrito a função de verossimilhança é:

$$L_{\Omega_0}(\mathbf{X}; \boldsymbol{\mu}; \boldsymbol{\Sigma}_0) = (2\pi)^{-\frac{np}{2}} |\boldsymbol{\Sigma}_0|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}_0^{-1} (\mathbf{X}_i - \boldsymbol{\mu}) \right\}. \quad (2.13)$$

Diante da matriz de covariâncias $\boldsymbol{\Sigma}_0$ simétrica positiva definida, realiza-se a decomposição espectral $\boldsymbol{\Sigma}_0 = \mathbf{P}_0 \boldsymbol{\Lambda}_0 \mathbf{P}_0^\top$, para definir a matriz de covariâncias inversa como $\boldsymbol{\Sigma}_0^{-1} = \mathbf{P}_0 \boldsymbol{\Lambda}_0^{-1} \mathbf{P}_0^\top$. Assim, a primeira modificação é realizada no expoente exp neperiano por meio da decomposição espectral, juntamente com as combinações lineares dos autovetores e os autovalores originais de $\boldsymbol{\Sigma}_0$.

$$\begin{aligned} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}_0^{-1} (\mathbf{X}_i - \boldsymbol{\mu}) &= \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})^\top \mathbf{P}_0 \boldsymbol{\Lambda}_0^{-1} \mathbf{P}_0^\top (\mathbf{X}_i - \boldsymbol{\mu}) && (2.14) \\ &= \sum_{i=1}^n \left\{ \frac{[\mathbf{e}_1^\top (\mathbf{X}_i - \boldsymbol{\mu})]^2}{\sigma^2[1 + (p-1)\rho]} + \frac{1}{\sigma^2(1-\rho)} \sum_{k=2}^p [\mathbf{e}_k^\top (\mathbf{X}_i - \boldsymbol{\mu})]^2 \right\} \\ &= \frac{1}{\sigma^2[1 + (p-1)\rho]} \sum_{i=1}^n \text{tr}[(\mathbf{X}_i - \boldsymbol{\mu})^\top \mathbf{e}_1 \mathbf{e}_1^\top (\mathbf{X}_i - \boldsymbol{\mu})] + \\ &\quad + \frac{1}{\sigma^2(1-\rho)} \sum_{i=1}^n \sum_{k=2}^p \text{tr}[(\mathbf{X}_i - \boldsymbol{\mu})^\top \mathbf{e}_k \mathbf{e}_k^\top (\mathbf{X}_i - \boldsymbol{\mu})] \\ &= \frac{1}{\sigma^2[1 + (p-1)\rho]} \sum_{i=1}^n \text{tr}[\mathbf{e}_1 \mathbf{e}_1^\top (\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})^\top] + \\ &\quad + \frac{1}{\sigma^2(1-\rho)} \sum_{k=2}^p \sum_{i=1}^n \text{tr}[\mathbf{e}_k \mathbf{e}_k^\top (\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})^\top], \end{aligned}$$

em que \mathbf{P}_0 é a matriz dos autovetores de $\mathbf{\Sigma}_0$ e $\mathbf{\Lambda}_0$ a matriz diagonal dos autovalores de $\mathbf{\Sigma}_0$. De acordo com a ortogonalidade de \mathbf{P} obtém-se:

$$\sum_{i=1}^p \mathbf{e}_i \mathbf{e}_i^\top = \mathbf{P} \mathbf{P}^\top = \mathbf{I},$$

portanto,

$$\mathbf{e}_1 \mathbf{e}_1^\top = \frac{1}{p} \mathbf{J}, \quad (2.15)$$

logo:

$$\sum_{i=2}^p \mathbf{e}_i \mathbf{e}_i^\top = \mathbf{I} - \frac{1}{p} \mathbf{J}. \quad (2.16)$$

Utilizando a propriedade de que a soma dos traços é igual ao traço da soma, em seguida colocando em pós evidência o somatório em relação a i da segunda parcela da soma, abordando juntamente os resultados das equações 2.15 e 2.16, obtém-se:

$$\begin{aligned} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})^\top \mathbf{\Sigma}_0^{-1} (\mathbf{X}_i - \boldsymbol{\mu}) &= \frac{1}{\sigma^2 [1 + (p-1)\rho]} \text{tr} \left[\mathbf{e}_1 \mathbf{e}_1^\top \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}) (\mathbf{X}_i - \boldsymbol{\mu})^\top \right] + \\ &+ \frac{1}{\sigma^2 (1-\rho)} \text{tr} \left[\sum_{k=2}^p \mathbf{e}_k \mathbf{e}_k^\top \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}) (\mathbf{X}_i - \boldsymbol{\mu})^\top \right] \\ &= \frac{1}{\sigma^2 [1 + (p-1)\rho]} \text{tr} \left[\frac{\mathbf{J}}{p} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}) (\mathbf{X}_i - \boldsymbol{\mu})^\top \right] + \\ &+ \frac{1}{\sigma^2 (1-\rho)} \text{tr} \left\{ \left[\sum_{k=2}^p \mathbf{e}_k \mathbf{e}_k^\top \right] \left[\sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}) (\mathbf{X}_i - \boldsymbol{\mu})^\top \right] \right\} \\ &= \frac{1}{\sigma^2 [1 + (p-1)\rho]} \text{tr} \left[\frac{\mathbf{J}}{p} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}) (\mathbf{X}_i - \boldsymbol{\mu})^\top \right] + \\ &+ \frac{1}{\sigma^2 (1-\rho)} \text{tr} \left\{ \left(\mathbf{I} - \frac{\mathbf{J}}{p} \right) \left[\sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}) (\mathbf{X}_i - \boldsymbol{\mu})^\top \right] \right\}. \end{aligned}$$

Com intuito de simplificar mais o resultado acima, soma-se e subtrai-se a média amostral, $\bar{\mathbf{X}}$, na parcela $\sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})^\top$, da seguinte forma:

$$\begin{aligned} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})^\top &= \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}} + \bar{\mathbf{X}} - \boldsymbol{\mu})(\mathbf{X}_i - \bar{\mathbf{X}} + \bar{\mathbf{X}} - \boldsymbol{\mu})^\top \\ &= \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^\top + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})^\top \\ &= n\hat{\boldsymbol{\Sigma}} + n\mathbf{H}, \end{aligned} \quad (2.17)$$

em que: $\hat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^\top$ e $\mathbf{H} = (\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})^\top$.

Ao considerar as reparametrizações: $\lambda_1 = \sigma^2[1 + (p-1)\rho]$ e $\lambda_2 = \sigma^2(1-\rho)$ e o rearranjo da expressão (2.17) e substituí-los em $\sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}_0^{-1} (\mathbf{X}_i - \boldsymbol{\mu})$, obtém-se:

$$\begin{aligned} \sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}_0^{-1} (\mathbf{X}_i - \boldsymbol{\mu}) &= \frac{1}{\lambda_1} tr \left[\frac{\mathbf{J}(n\hat{\boldsymbol{\Sigma}} + n\mathbf{H})}{p} \right] + \frac{1}{\lambda_2} tr \left\{ \left(\mathbf{I} - \frac{\mathbf{J}}{p} \right) (n\hat{\boldsymbol{\Sigma}} + n\mathbf{H}) \right\} \\ &= \frac{n}{p\lambda_1} \left[\sum_{i=1}^p \hat{\sigma}_{ii} + 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} + tr(\mathbf{JH}) \right] + \\ &\quad + \frac{n}{p\lambda_2} \left[(p-1) \sum_{i=1}^p \hat{\sigma}_{ii} - 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} + ptr(\mathbf{H}) - tr(\mathbf{JH}) \right]. \end{aligned}$$

O determinante da matriz de covariâncias com simetria composta também será reparametrizado por: $|\boldsymbol{\Sigma}_0| = \lambda_1(\lambda_2)^{p-1}$. Após todas essas modificações realizadas de acordo com Ferreira (2018), a função de verossimilhança com espaço paramétrico restrito será:

$$\begin{aligned} L_{\Omega_0}(\mathbf{X}; \boldsymbol{\mu}; \boldsymbol{\Sigma}_0) &= (2\pi)^{-\frac{np}{2}} \lambda_1^{-\frac{n}{2}} (\lambda_2)^{-\frac{n(p-1)}{2}} \times \\ &\quad \times \exp \left\{ \frac{n}{2p\lambda_1} \left[\sum_{i=1}^p \hat{\sigma}_{ii} + 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} + tr(\mathbf{JH}) \right] - \right. \\ &\quad \left. - \frac{n}{2p\lambda_2} \left[(p-1) \sum_{i=1}^p \hat{\sigma}_{ii} - 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} + ptr(\mathbf{H}) - tr(\mathbf{JH}) \right] \right\}. \end{aligned} \quad (2.18)$$

A função suporte da função de verossimilhança com espaço paramétrico restrito será:

$$\begin{aligned}
g_{\Omega_0}(\mathbf{X}; \boldsymbol{\mu}; \boldsymbol{\Sigma}_0) = & -\frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln(\lambda_1) - \frac{n(p-1)}{2} \ln(\lambda_2) - \\
& -\frac{n}{2p\lambda_1} \left[\sum_{i=1}^p \hat{\sigma}_{ii} + 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} + \text{tr}(\mathbf{JH}) \right] - \\
& -\frac{n}{2p\lambda_2} \left[(p-1) \sum_{i=1}^p \hat{\sigma}_{ii} - 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} + p \text{tr}(\mathbf{H}) - \text{tr}(\mathbf{JH}) \right]. \quad (2.19)
\end{aligned}$$

Diante da função suporte, calcula-se o estimador de máxima verossimilhança para cada parâmetro da distribuição, por meio da derivada parcial de primeira ordem em relação ao parâmetro de interesse, em seguida iguala a zero a expressão obtida, para obter o estimador que maximiza a função de verossimilhança, como mostra-se a seguir:

- Estimador do vetor de médias é dado por:

$$\begin{aligned}
\frac{\partial g_{\Omega_0}(\mathbf{X}; \boldsymbol{\mu}; \boldsymbol{\Sigma}_0)}{\partial \boldsymbol{\mu}} = & \frac{n}{p\lambda_1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) + \frac{n}{p\lambda_2} [p(\bar{\mathbf{X}} - \boldsymbol{\mu}) - \mathbf{J}(\bar{\mathbf{X}} - \boldsymbol{\mu})] \\
& \frac{n}{p\lambda_1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) + \frac{n}{p\lambda_2} [p(\bar{\mathbf{X}} - \boldsymbol{\mu}) - \mathbf{J}(\bar{\mathbf{X}} - \boldsymbol{\mu})] = 0,
\end{aligned}$$

resultando em:

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}}.; \quad (2.20)$$

- Estimador do parâmetro λ_1 , é obtido substituindo-se o estimador de $\boldsymbol{\mu}$, o que faz com que $\hat{\mathbf{H}} = 0$. Logo,

$$\begin{aligned}
\frac{\partial g_{\Omega_0}(\mathbf{X}; \boldsymbol{\mu}; \boldsymbol{\Sigma}_0)}{\partial \lambda_1} = & -\frac{n}{2\lambda_1} + \frac{n}{2p\lambda_1^2} \left[\sum_{i=1}^p \hat{\sigma}_{ii} + 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} + \text{tr}(\mathbf{JH}) \right] \\
& -\frac{n}{2\hat{\lambda}_1} + \frac{n}{2p\hat{\lambda}_1^2} \left[\sum_{i=1}^p \hat{\sigma}_{ii} + 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} \right] = 0 \\
& -\hat{\lambda}_1 + \frac{1}{p} \left[\sum_{i=1}^p \hat{\sigma}_{ii} + 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} \right] = 0,
\end{aligned}$$

resultando em:

$$\hat{\lambda}_1 = \frac{1}{p} \left[\sum_{i=1}^p \hat{\sigma}_{ii} + 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} \right]; \quad (2.21)$$

- Estimador do parâmetro λ_2 , usa o estimador de $\boldsymbol{\mu}$, que faz com que $\hat{\mathbf{H}} = \mathbf{0}$. Logo,

$$\begin{aligned} \frac{\partial g_{\Omega_0}(\mathbf{X}; \boldsymbol{\mu}; \boldsymbol{\Sigma}_0)}{\partial \lambda_2} &= -\frac{n(p-1)}{2\lambda_2} + \frac{n}{2p\lambda_2^2} \left[(p-1) \sum_{i=1}^p \hat{\sigma}_{ii} - 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} + \text{ptr}(\mathbf{H}) - \text{tr}(\mathbf{JH}) \right] \\ &- \frac{n(p-1)}{2\hat{\lambda}_2} + \frac{n}{2p\hat{\lambda}_2^2} \left[(p-1) \sum_{i=1}^p \hat{\sigma}_{ii} - 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} \right] = 0 \\ &-(p-1)\hat{\lambda}_2 + \frac{1}{p} \left[(p-1) \sum_{i=1}^p \hat{\sigma}_{ii} - 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} \right] = 0, \end{aligned}$$

resultando em:

$$\hat{\lambda}_2 = -\frac{1}{p(p-1)} \left[(p-1) \sum_{i=1}^p \hat{\sigma}_{ii} - 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} \right]; \quad (2.22)$$

- O estimador de σ^2 é obtido da seguinte forma, usando-se o princípio da invariância dos estimadores de máxima verossimilhança:

$$\begin{aligned} \hat{\lambda}_1 + (p-1)\hat{\lambda}_2 &= \hat{\sigma}^2 [1 + (p-1)\hat{\rho}] + (p-1)\hat{\sigma}^2(1 - \hat{\rho}) \\ &= \hat{\sigma}^2 [1 + (p-1)\hat{\rho} + p - 1 - (p-1)\hat{\rho}] \\ &= p\hat{\sigma}^2, \end{aligned}$$

resultando em

$$\hat{\sigma}^2 = \frac{\hat{\lambda}_1 + (p-1)\hat{\lambda}_2}{p},$$

substituindo os estimadores de λ_1 e λ_2 apresentados respectivamente nas expressões (2.21) e (2.22), $\hat{\sigma}^2$ será a média dos estimadores viesados, como mostrado a seguir:

$$\begin{aligned} \hat{\sigma}^2 &= \frac{1}{p} \left\{ \frac{1}{p} \left[\sum_{i=1}^p \hat{\sigma}_{ii} + 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} \right] + \frac{(p-1)}{p(p-1)} \left[(p-1) \sum_{i=1}^p \hat{\sigma}_{ii} - 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} \right] \right\} \\ &= \frac{1}{p} \sum_{i=1}^p \hat{\sigma}_{ii}; \end{aligned} \quad (2.23)$$

- O estimador de ρ será obtido a partir da seguinte relação, usando-se o mesmo princípio da invariância dos estimadores de máxima verossimilhança como mostra-se a seguir:

$$\begin{aligned}
\hat{\sigma}^2(1 - \hat{\rho}) &= \hat{\lambda}_2 \\
\hat{\rho} &= 1 - \frac{\hat{\lambda}_2}{\hat{\sigma}^2} \\
&= 1 - \frac{(p-1) \sum_{i=1}^p \hat{\sigma}_{ii} - 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik}}{p(p-1) \hat{\sigma}^2} \\
&= \frac{(p-1)p \hat{\sigma}^2 - (p-1) \sum_{i=1}^p \hat{\sigma}_{ii} + 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik}}{p(1-p) \hat{\sigma}^2},
\end{aligned}$$

lembrando que $p \hat{\sigma}^2 = \sum_{i=1}^p \hat{\sigma}_{ii}$, obtém-se:

$$\hat{\rho} = \frac{(p-1) \sum_{i=1}^p \hat{\sigma}_{ii} - (p-1) \sum_{i=1}^p \hat{\sigma}_{ii} + 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik}}{p(p-1) \hat{\sigma}^2} = \frac{2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik}}{p(p-1) \hat{\sigma}^2}. \quad (2.24)$$

O máximo da função de verossimilhança com espaço paramétrico restrito será obtida ao substituir os estimadores de máxima verossimilhança dos respectivos parâmetros, além de utilizar as relações existentes entre λ_1 e λ_2 , como as de σ^2 e ρ . Ao realizar a substituição é necessário lembrar que $\hat{\mathbf{H}}$ é uma matriz nula quando $\boldsymbol{\mu}$ é trocado pelo seu estimador na matriz \mathbf{H} original. Assim,

$$\begin{aligned}
L_{\Omega_0}(\mathbf{X}; \hat{\boldsymbol{\mu}}; \hat{\boldsymbol{\Sigma}}_0) &= (2\pi)^{-\frac{np}{2}} \hat{\lambda}_1^{-\frac{n}{2}} (\hat{\lambda}_2)^{-\frac{n(p-1)}{2}} \times \\
&\times \exp \left\{ \frac{n}{2p\hat{\lambda}_1} \left[\sum_{i=1}^p \hat{\sigma}_{ii} + 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} + \text{tr}(\mathbf{J}\hat{\mathbf{H}}) \right] - \right. \\
&\left. - \frac{n}{2p\hat{\lambda}_2} \left[(p-1) \sum_{i=1}^p \hat{\sigma}_{ii} - 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} + \text{ptr}(\hat{\mathbf{H}}) - \text{tr}(\mathbf{J}\hat{\mathbf{H}}) \right] \right\}
\end{aligned}$$

$$\begin{aligned}
&= (2\pi)^{-\frac{np}{2}} (\hat{\sigma}^2)^{-\frac{np}{2}} (1 - \hat{\rho})^{-\frac{n(p-1)}{2}} [1 + (p-1)\hat{\rho}]^{-\frac{n}{2}} \times \\
&\quad \times \exp \left\{ \frac{n}{2p\hat{\lambda}_1} \left[\sum_{i=1}^p \hat{\sigma}_{ii} + 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} \right] - \right. \\
&\quad \left. - \frac{n}{2p\hat{\lambda}_2} \left[(p-1) \sum_{i=1}^p \hat{\sigma}_{ii} - 2 \sum_{i=1}^{p-1} \sum_{k=i+1}^p \hat{\sigma}_{ik} \right] \right\} \\
&= (2\pi)^{-\frac{np}{2}} (\hat{\sigma}^2)^{-\frac{np}{2}} (1 - \hat{\rho})^{-\frac{n(p-1)}{2}} [1 + (p-1)\hat{\rho}]^{-\frac{n}{2}} \times \\
&\quad \times \exp \left\{ \frac{np\hat{\lambda}_1}{2p\hat{\lambda}_1} - \frac{np(p-1)\hat{\lambda}_2}{2p\hat{\lambda}_2} \right\},
\end{aligned}$$

resultando em:

$$L_{\Omega_0}(\mathbf{X}; \hat{\boldsymbol{\mu}}; \hat{\boldsymbol{\Sigma}}_0) = (2\pi)^{-\frac{np}{2}} (\hat{\sigma}^2)^{-\frac{np}{2}} (1 - \hat{\rho})^{-\frac{n(p-1)}{2}} [1 + (p-1)\hat{\rho}]^{-\frac{n}{2}} \exp \left\{ -\frac{np}{2} \right\}.$$

2.3.3 Estatística do teste

Após obter o máximo das funções de verossimilhança com espaço paramétrico irrestrito e restrito respectivamente nas subseções 2.3.1 e 2.3.2, a estatística do teste é calculada por meio da razão destas quantidades:

$$\begin{aligned}
\Lambda &= \frac{L_{\Omega_0}(\mathbf{X}; \hat{\boldsymbol{\mu}}; \boldsymbol{\Sigma}_0)}{L_{\Omega}(\mathbf{X}; \hat{\boldsymbol{\mu}}; \hat{\boldsymbol{\Sigma}})} = \frac{(2\pi)^{-\frac{np}{2}} (\hat{\sigma}^2)^{-\frac{np}{2}} (1 - \hat{\rho})^{-\frac{n(p-1)}{2}} [1 + (p-1)\hat{\rho}]^{-\frac{n}{2}} \exp \left\{ -\frac{np}{2} \right\}}{(2\pi)^{-\frac{np}{2}} |\mathbf{S}_n|^{-\frac{n}{2}} \exp \left\{ -\frac{np}{2} \right\}} \\
&= \left\{ \frac{|\mathbf{S}_n|}{(\hat{\sigma}^2)^p (1 - \hat{\rho})^{p-1} [1 + (p-1)\hat{\rho}]} \right\}^{\frac{n}{2}},
\end{aligned}$$

em que \mathbf{S}_n é o estimador de máxima verossimilhança de $\boldsymbol{\Sigma}$, com espaço paramétrico irrestrito.

De acordo com Ferreira (2018), a distribuição de Λ é intratável, portanto, realiza-se a transformação $-2\ln(\Lambda)$, para obter uma aproximação assintótica da distribuição qui-quadrado, apresentada na expressão (2.25), cujo teste tem tamanho aproximado α , quando aplicado em grandes amostras. Logo,

$$\chi^2 = n[p\ln(\hat{\sigma}^2) + (p-1)\ln(1 - \hat{\rho}) + \ln[1 + (p-1)\hat{\rho}] - \ln|\mathbf{S}_n|], \quad (2.25)$$

com $\nu = p(p+1)/2 - 2$ graus de liberdade.

Com intuito de melhorar a convergência de $-2\ln(\Lambda)$ para distribuição qui-quadrado, substitui-se os estimadores de máxima verossimilhança viesados pelos seus respectivos estimadores não viesados e realiza-se a correção de Box (1949), como mostra a expressão (2.26):

$$\chi_c^2 = \left[(n-1) - \frac{p(p+1)^2(2p-3)}{6(p-1)(p^2+p+4)} \right] (\ln|\mathbf{S}_0| - \ln|\mathbf{S}|), \quad (2.26)$$

que sob a hipótese nula, a expressão (2.26) possui distribuição assintótica qui-quadrado, com $f = p(p+1)/2 - 2$ graus de liberdade, uma vez o modelo irrestrito contém $p + p(p+1)/2$ parâmetros estimados e o modelo restrito $p+2$ parâmetros estimados. Os estimadores de Σ não viesado com espaço paramétrico irrestrito e restrito já foram apresentados respectivamente nas expressões (2.1) e (2.3).

A hipótese nula será rejeitada se, e somente se, o resultado da estatística do teste for maior ou igual a $\chi_{\alpha,f}^2$, em que $\chi_{\alpha,f}^2$ representa o quantil superior 100 $\alpha\%$ da distribuição qui-quadrado assintótica com f graus de liberdade. O tamanho real do teste será aproximadamente α se $\Sigma_0 \in \Omega_0$ com uma amostra de tamanho grande (CASELLA; BERGER, 2014).

Desta forma, a medida que o número de observações amostrais cresce, a distribuição da estatística do teste converge para uma distribuição qui-quadrado (CASELLA; BERGER, 2014). Além disso, é importante ressaltar que ao realizar o teste, considera-se o pressuposto de normalidade multivariada das populações amostradas, desta forma, desvios de normalidade afetam fortemente o seu desempenho, evidenciando um teste pouco robusto (FERREIRA, 2018). A próxima seção apresentará diversos estudos presentes na literatura que buscam aperfeiçoar o teste para matriz de covariâncias com simetria composta.

2.3.4 Estudos sobre o teste original da matriz de covariâncias com simetria composta (LRTO)

O teste de hipóteses para matriz de covariâncias com simetria composta foi inicialmente elaborado por Wilks et al. (1946). Assim, em uma distribuição populacional normal multivariada, o teste considerava na hipótese nula que as variâncias eram iguais entre si para as p variáveis e que todas as covariâncias entre as combinações duas a duas das p variáveis eram iguais. A inspiração para a criação deste teste, de acordo com os autores Wilks et al. (1946), foi criar uma extensão do teste desenvolvido por Mauchly, relativo à esfericidade de uma distribuição normal multivariada, que era projetado para verificar se todas as variâncias eram iguais

entre si e que todas as covariâncias eram iguais a zero. Vale a pena salientar que o teste de esfericidade de Mauchly (1940), é na verdade, também um teste de simetria composta, em que as covariâncias são todas nulas.

Para desenvolver o teste, Wilks et al. (1946) utilizaram o método de razão de verossimilhanças, obtendo distribuição exata para estatística do teste, sob a hipótese nula, ao considerar $p = 2$ e $p = 3$. No entanto, a derivação da distribuição exata para maiores valores de p é muito complexa para ser utilizada na prática. Portanto, os autores apresentaram também a distribuição assintótica qui-quadrado para a estatística do teste, cuja precisão está diretamente relacionada ao tamanho da amostra, além de assumir normalidade multivariada dos dados amostrados, tornando-se um teste pouco robusto sobre este aspecto.

Votaw Jr (1948) desenvolveu um teste para simetria composta em subconjuntos de variáveis, cujo estudo aborda certos problemas psicométricos e de pesquisa médica, nos quais a hipótese de simetria composta é relevante para testar se várias formas de um exame podem ser usadas de forma intercambiável. Votaw Jr (1948) utilizou o método da razão de verossimilhanças para desenvolver o teste, apresentando como resultados a distribuição exata da estatística do teste, como a distribuição do produto de variáveis betas independentes, e os momentos exatos dessa estatística, além de uma distribuição aproximada da estatística do teste para grandes amostras.

Diante da complexidade de calcular os pontos percentuais da distribuição da estatística do teste usando expressões exatas como Votaw Jr (1948) fez, foi realizado um estudo que aproxima essa distribuição com uma distribuição adequada pertencente à família das distribuições de Pearson. Os autores verificaram que a precisão dessa aproximação é suficiente para propósitos práticos, cujo estudo apresenta os pontos percentuais da distribuição da estatística do teste para a simetria composta, calculados de acordo com a aproximação proposta (LEE; KRIDHNAIAH; CHANG, 1976).

Marques e Coelho (2015) realizaram um estudo que combina os testes de esfericidade e simetria composta em um único teste, assumindo para a matriz de covariâncias na hipótese nula uma estrutura de esfericidade e na hipótese alternativa a estrutura de simetria composta. Neste estudo derivou-se a estatística do teste da razão de verossimilhanças para obter uma distribuição exata, por meio do produto de variáveis betas independentes. Em virtude das dificuldades existentes no manuseio da distribuição exata, foram realizadas aproximações assintóticas por meio

de misturas de distribuições gama, cujos estudos numéricos avaliaram ter ótima aproximação, mesmo em pequenas amostras.

Em um estudo realizado por Coelho, Marques e Oliveira (2016) foi abordado outra forma de testar a matriz de covariâncias com simetria composta por meio de dois testes com hipóteses simultâneas. A hipótese nula do primeiro teste afirmava a nulidade das médias e estrutura de simetria composta da matriz de covariâncias. No segundo teste, na hipótese nula considerava-se a igualdade entre todas as médias e estrutura de simetria composta da matriz de covariâncias.

Como as distribuições exatas para as estatísticas dos testes citados acima são muito elaboradas para serem usadas na prática, os autores Coelho, Marques e Oliveira (2016) propuseram o uso de distribuições quase exatas para estas estatísticas. Desta forma, as distribuições assintóticas foram obtidas utilizando um conceito diferente de aproximação de distribuições, realizando a decomposição do logaritmo da estatística para manter a maior parte da função característica inalterada e substituindo o restante (menor parte) por uma aproximação assintótica. Estudos numéricos mostraram qualidade e a precisão das distribuições quase exatas desenvolvidas, ao contrário do que acontece, por exemplo, com a aproximação qui-quadrado usual.

Em todos os estudos citados os autores buscaram melhorar os testes de hipóteses, da matriz de covariâncias com estrutura de simetria composta, em relação a distribuição assintótica da estatística do teste, todos eles abordaram a pressuposição que os dados amostrados para realização dos testes eram provenientes de uma população com distribuição normal multivariada.

Uma nova abordagem surge com os autores Morris, Payton e Santorico (2011), ao realizarem um teste não paramétrico para matriz de covariâncias com simetria composta, dado que um dos requisitos para realizar o teste da razão de verossimilhanças é que os dados amostrados sejam de uma população normal multivariada. Desta forma, esses autores propuseram o uso de permutação ou aleatorização, para desenvolver um novo teste robusto em relação aos desvios de normalidade dos dados, uma vez que o método de permutação reduz ou elimina as suposições distribucionais e permite o uso em praticamente qualquer estatística.

De acordo com Morris, Payton e Santorico (2011), o teste proposto foi nomeado de MCPT e a sua estatística foi calculada somando as diferenças absolutas entre os elementos da diagonal ou abaixo dela, correspondentes na matriz de covariâncias obtida dos dados observados e na matriz hipotética de covariâncias estimada a partir dos dados amostrais. Esta estatística é

dada por:

$$D = \mathbf{1}_{\frac{1}{2}p(p+1)}^\top |vec(\mathbf{M})| = \mathbf{1}_{\frac{1}{2}p(p+1)}^\top |vec(\boldsymbol{\Sigma}_{obs} - \hat{\boldsymbol{\Sigma}}_{CS})|,$$

em que $vec(M)$ é o vetor dos elementos na diagonal ou abaixo da diagonal de \mathbf{M} , $\boldsymbol{\Sigma}_{obs}$ e $\hat{\boldsymbol{\Sigma}}_{CS}$ são respectivamente a matriz de covariâncias dos dados observados e a matriz hipotética de covariâncias dos dados observados. A matriz hipotética foi obtida da seguinte forma:

$$\hat{\boldsymbol{\Sigma}}_{CS} = \bar{s}^2 [(1 - \bar{r})\mathbf{I}_p + \bar{r}\mathbf{J}],$$

em que \mathbf{J} é uma matriz com dimensões $(p \times p)$ composta por valores iguais a 1, \bar{s}^2 e \bar{r} são as médias das variâncias e correlações. Esta estatística do teste foi calculada para cada permutação possível dos dados. A proporção dos valores da estatística do teste maior que ou igual ao obtido a partir dos dados originais é o valor- p .

De acordo com os resultados, ao considerar a distribuição normal multivariada, o teste MCPT foi liberal em relação a taxa de erro tipo I quando o número de variáveis ($p = 3$) se aproximava do tamanho amostral ($n = 5$) e exato a medida que aumentava o tamanho amostral ($n > 25$) com número de variáveis fixo ($p = 3$). Ainda considerando a distribuição normal multivariada o teste MCPT exibiu baixo poder ($< 0,8$), principalmente para pequenos tamanhos amostrais. Ao considerar outras distribuições de probabilidade como exponencial de dois parâmetros, uniforme e exponencial dupla o teste MCPT foi robusto para não normalidade, mas exibiu baixo poder, particularmente em amostras pequenas.

2.4 Outliers Multivariados

Outliers podem ser definidos intuitivamente como observações que se afastam das demais observações analisadas em um conjunto de dados, a presença dos *outliers* são recorrentes nos conjuntos de dados em qualquer domínio de aplicação (HAWKINS, 1980).

Os *outliers* podem ser gerados ao escolher um modelo paramétrico errado ou quando parte das observações seguem outro modelo (HAMPEL, 1971). Os *outliers* também podem ser provenientes de uma distribuição com caudas pesadas como a t de Student, dado que algumas famílias de distribuições podem produzir *outliers* mais comumente (HAWKINS, 1980).

No entanto, de acordo com Hampel et al. (2011), a principal causa da geração de *outliers* é pela ocorrência de erros brutos, devido a uma fonte de desvios que agem apenas oca-

sionalmente. Algumas fontes de erros brutos são: erro de digitação, erro de cópia, falha de equipamentos e outros efeitos transitórios.

Quando a amostra contém *outliers* gerados por meio de erros brutos, claramente este valor discrepante pode ser muito prejudicial ao procedimento estatístico, como no caso da média aritmética, o ideal seria identificar estes erros brutos e removê-los da amostra (HAMPEL, 1971).

Em conjunto de dados multivariados, a detecção de *outliers* pode ser extremamente difícil, necessitando de uma estatística robusta, pois a maioria dos estimadores clássicos falham quando a fração de contaminação é maior que $1/(p + 1)$, em que p é a dimensão dos dados, ou seja, isso significa que em alta dimensão, uma pequena fração de *outliers* pode resultar em estimativas muito ruins (ROCKE; WOODRUFF, 1996).

2.5 Estatística robusta

Na estatística clássica, os principais métodos de inferências são realizados por meio de modelos paramétricos, condicionados a determinadas pressuposições, que quando violadas, comprometem a confiabilidade das inferências. A estatística robusta busca melhorar estes modelos estatísticos já existentes, deixando-os robustos quando os dados apresentam algum tipo de desvio que poderia comprometer as pressuposições do modelo (HAMPEL et al., 2011).

De acordo com Hampel et al. (2011), a estatística robusta é uma extensão da estatística paramétrica clássica, levando em consideração que os modelos paramétricos são apenas aproximações da realidade, pois pressuposições como normalidade multivariada e independência dos dados amostrados, nem sempre são obtidas.

Um das principais causas dos desvios de pressupostos dos modelos paramétricos é a ocorrência de *outliers* devidos a uma fonte de desvios que agem apenas ocasionalmente (HAMPEL et al., 2011). Para que o modelo estatístico assumido obtenha eficiência ótima ou perto de ótima, diante dos *outliers*, utiliza-se a estatística robusta por meio de estimadores robustos (CASELLA; BERGER, 2014).

Portanto, a estatística robusta aborda o conceito de ponto de ruptura, entendido como uma medida de confiabilidade global do estimador, que descreve até que fração de *outliers* o estimador ainda fornece alguma informação da distribuição original dentro do modelo paramétrico. Indicando assim, a maior fração de contaminação que o estimador possa suportar e ainda fornece informação confiável (HAMPEL et al., 2011).

O teste sobre a estrutura da matriz de covariâncias de uma população, possui o pressuposto de normalidade multivariada da população amostrada, portanto, desvios de normalidade afetam fortemente seu desempenho (FERREIRA, 2018). Este trabalho aborda o estimador robusto Comedian, para obter a matriz de covariâncias robusta, com intuito de aprimorar o teste já existente LRTO, deixando-o robusto quanto ao desvio de normalidade dos dados causado por *outliers*.

Na literatura existem diversos métodos equivalentes para detecção de *outliers* como: Mínimo Volume Elipsóide (MVE), o Mínimo Determinante da Covariância (MCD), o S-Estimador e o M-Estimador, os quais estimam a matriz de covariâncias de forma robusta a partir de dados multivariados (ROCKE; WOODRUFF, 1996). No entanto, o estimador robusto Comedian foi escolhido devido as conclusões do estudo realizado por Pereira e Ferreira (2017), que verificou que os testes que utilizam o estimador Comedian obtiveram os melhores resultados, controlando as taxas de erro do tipo I e obtendo alto poder.

2.5.1 Estimador robusto Comedian

Falk (1997) criou o Comedian (COM) ao propor uma medida de dependência robusta entre duas variáveis aleatórias, cuja medida pode ser abordada como uma alternativa robusta da covariação ou correlação entre duas variáveis, a qual é dada por:

$$COM = Med((X - Med(X))(Y - Med(Y))),$$

em que *Med* representa a mediana. Este resultado é proveniente de uma generalização do desvio absoluto mediano (*MAD*), ao realizar $\sigma^2 = MAD^2$, quando $X = Y$, dado que *MAD* é uma alternativa robusta do estimador clássico do desvio padrão (σ), sendo descrita da seguinte forma:

$$MAD(X) = Med(|X - Med(X)|).$$

A mediana amostral é considerada uma medida robusta devido à insensibilidade a observações extremas, possuindo um ponto de ruptura de 50% (CASELLA; BERGER, 2014). Sendo assim, o Comedian é na realidade uma mediana, e como a mediana, ele também possui o maior ponto de ruptura possível (FALK, 1997).

De acordo com Falk (1997) as principais propriedades e características do Comedian são:

- Comedian pode ser utilizado como uma medida alternativa de covariância robusta, portanto, $COM(X,Y)$ é correspondente a $COV(X,Y)$. Entretanto, $COV(X,Y)$ requer a existência dos dois primeiros momentos das variáveis X e Y , enquanto $COM(X,Y)$ sempre existirá, dado que seu cálculo é realizado por meio da mediana, que é uma medida de locação;
- Se X e Y são independentes, $COM(X,Y) = 0$;
- Comedian é simétrico, invariante em relação a translações e possui escala equivariante, isto é: $COM(X, aY + b) = aCOM(X,Y) = aCOM(Y,X)$;
- Se $Y = aX + b$, $COM(X, aY + b) = aMAD^2(X)$;
- Uma alternativa baseada na mediana natural para o coeficiente de correlação será a correlação mediana:

$$\delta = \frac{COM(X,Y)}{MAD(X)MAD(Y)},$$

sendo que, se X e Y são independentes, $\delta = 0$ e, em que $\delta \in [-1,1]$.

De acordo com Sajesh e Srinivasan (2012), na estatística multivariada utiliza-se o Comedian para obter a matriz de covariâncias robusta ($COM(\mathbf{X})$). Para isso, considera-se \mathbf{X} a matriz de dados com dimensões $n \times p$, composta por vetores p -dimensionais \mathbf{X}_i^\top formando suas linhas, ($i = 1, \dots, n$) e vetores n -dimensionais \mathbf{X}_k , ($k = 1, \dots, p$) formando suas colunas, cuja matriz $COM(\mathbf{X})$ será calculada por:

$$COM(\mathbf{X}) = \begin{bmatrix} MAD^2(X_1) & COM(X_1, X_2) & \dots & COM(X_1, X_p) \\ COM(X_2, X_1) & MAD^2(X_2) & \dots & COM(X_2, X_p) \\ \vdots & \vdots & \ddots & \vdots \\ COM(X_p, X_1) & COM(X_p, X_2) & \dots & MAD^2(X_p) \end{bmatrix}.$$

De forma similar é possível obter a matriz do coeficiente de correlação robusto da seguinte forma:

$$\delta(\mathbf{X}) = DCOM(\mathbf{X})D^\top,$$

em que \mathbf{D} é uma matriz diagonal com elementos diagonais $\frac{1}{MAD(\mathbf{X}_i)}$; ($i = 1, \dots, p$) (SAJESH; SRINIVASAN, 2012).

No entanto, a matriz Comedian $\mathbf{COM}(\mathbf{X})$, como alternativa robusta à matriz de covariâncias, em geral não é positiva definida. Com intuito de resolver esta questão, Maronna e Zamar (2002) propuseram um método geral para obter a matriz de covariâncias positiva definida e aproximadamente invariante em relação as transformações lineares, cuja transformação é realizada da seguinte forma:

i) Realiza-se a decomposição espectral $\boldsymbol{\delta}(\mathbf{X}) = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}^\top$, em que $\boldsymbol{\Lambda}_{(p \times p)}$ é uma matriz diagonal, cujos elementos são os autovalores λ_j e $\mathbf{P}_{(p \times p)}$ é uma matriz, cujas colunas são os autovetores \mathbf{e}_j , para $j = 1, \dots, p$, ambas dadas por:

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \end{bmatrix} \quad \text{e} \quad \mathbf{P} = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p];$$

ii) considera-se $\mathbf{Q} = \mathbf{P}\mathbf{D}$, em que \mathbf{D} é uma matriz diagonal com elementos diagonais $\frac{1}{MAD(\mathbf{x}_i)}$; ($i = 1, \dots, p$) e $\mathbf{z}_i = \mathbf{Q}^{-1}\mathbf{x}_i$, para $i = 1, \dots, n$;

iii) os estimadores robustos para locação e escala são respectivamente:

$$\mathbf{m}(\mathbf{X}) = \mathbf{Q}\mathbf{L} \quad \mathbf{S}(\mathbf{X}) = \mathbf{Q}\boldsymbol{\Gamma}\mathbf{Q}^\top,$$

em que $\boldsymbol{\Gamma}_{(p \times p)} = \text{diag}(MAD^2(Z_1), \dots, MAD^2(Z_p))$ e $\mathbf{L}_{(p \times 1)} = [Med(Z_1), \dots, Med(Z_p)]^\top$.

Um processo iterativo é aplicado para melhorar as estimativas, substituindo $\boldsymbol{\delta}(\mathbf{X})$ por $\mathbf{S}(\mathbf{X})$ e repetindo os passos (i),(ii) e (iii), até que as estimativas se estabilizem em um ponto qualquer do processo iterativo (MARONNA; ZAMAR, 2002). O algoritmo Comedian encontra-se disponível no pacote *robustbase* realizado por Maechler et al. (2019), no programa R (R Core Team, 2018).

2.6 Método Monte Carlo

O método Monte Carlo é um mecanismo computacionalmente intensivo que pode ser utilizado para avaliação das propriedades de um novo teste ou de um novo procedimento de estimação. Sendo útil para avaliar propriedades dos estimadores, determinar tamanhos amostrais, além de obter soluções de outros inúmeros problemas de inferência estatística nas mais variadas áreas do conhecimento (FERREIRA, 2013).

Segundo Ferreira (2013), o método Monte Carlo realiza simulações de experimentos em que pelo menos um componente aleatório esteja presente. Desta forma, deve-se assumir uma distribuição de probabilidade para o componente aleatório do modelo, para então gerar dados aleatórios, que são na verdade pseudoaleatórios. Neste estudo, o método Monte Carlo será abordado para avaliar as propriedades de testes de hipóteses para estrutura de simetria composta, assumindo-se como distribuições de probabilidade a normal multivariada e normal multivariada contaminada.

Este método de simulação está cada vez mais presente em estudos estatísticos inferenciais, devido a sua facilidade para obtenção de soluções e ao avanço tecnológico da informática. No entanto, para utilizá-lo é necessário assumir a distribuição de probabilidade da variável aleatória de forma correta ou aproximadamente correta, para que a inferência não fique comprometida. Nos casos em que não é possível assumir precisamente a distribuição de probabilidade para o componente aleatório, deve-se utilizar outro método computacional intensivo como *bootstrap* não paramétrico ou permutação (FERREIRA, 2013).

2.7 *Bootstrap*

O método *Bootstrap* foi desenvolvido por Efron (1992) e tem como intuito conhecer as características da população em estudo, por meio de uma amostra original que será reamostrada milhares de vezes. A partir de B reamostragens da amostra original é possível obter informações para inferir a respeito de característica ou parâmetros da população, uma vez que a amostra original possui a máxima informação possível da população, por ter características análogas (CASELLA; BERGER, 2014).

Desta forma, o *bootstrap* é um método computacionalmente intensivo de reamostragens com reposição, que realiza a estimação de erros padrões, estima parâmetros por intervalos de confiança e realiza teste de hipóteses. Possui a vantagem de realizar inferências quando a distribuição de probabilidade da população não é conhecida ou quando a condição de normalidade é violada (FERREIRA, 2013).

As reamostragens podem ser realizadas de acordo com os métodos denominados de *bootstrap* paramétrico e *bootstrap* não-paramétrico. O *bootstrap* paramétrico considera uma amostra aleatória multivariada $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ a partir de uma distribuição com função densidade de probabilidade $f(\mathbf{x}|\boldsymbol{\theta})$, em que $\boldsymbol{\theta}$ é um vetor de parâmetros. A partir da amostra original, estima-se $\boldsymbol{\theta}$ por meio do estimador $\hat{\boldsymbol{\theta}}$ e em seguida, obtêm-se as reamostragens a partir do

modelo ajustado

$$\mathbf{X}_1^*, \mathbf{X}_2^*, \dots, \mathbf{X}_n^* \sim f(\mathbf{x}|\hat{\boldsymbol{\theta}}), \quad (2.27)$$

que fica completamente determinado por meio do vetor de parâmetros estimados (CASELLA; BERGER, 2014).

Portanto, as B reamostragens com reposição de tamanho n são realizadas a partir do modelo ajustado $f(\mathbf{x}|\hat{\boldsymbol{\theta}})$. Em cada reamostragem calcula-se uma estimativa da estatística de interesse que tem a mesma forma funcional da estatística original. Diante da distribuição de *bootstrap* obtida é possível realizar inferências sobre os parâmetros populacionais (DAVISON; HINKLEY, 1997).

O *bootstrap* não-paramétrico se caracteriza por não assumir nenhuma forma funcional para a função densidade de probabilidade da população em estudo (CASELLA; BERGER, 2014). Logo são realizadas B reamostragens, de tamanho n , da amostra original com probabilidade $1/n$ para cada componente da amostra ser sorteado, pois considera-se que a amostra original é independente e identicamente distribuída. Em seguida, calcula-se a estimativa do parâmetro de interesse em cada uma das B reamostragens, obtendo assim a distribuição *bootstrap* deste parâmetro, que servirá para se inferir sobre característica ou parâmetros da população (DAVISON; HINKLEY, 1997).

2.8 Avaliação dos testes de hipóteses

Ao realizar um teste de hipóteses, dois tipos de erros podem ocorrer: o erro tipo I, quando o teste decide incorretamente rejeitar a hipótese nula, e o erro tipo II, quando o teste decide incorretamente aceitar a hipótese nula. Geralmente os testes são avaliados e comparados por meio de suas probabilidades de incorrer nestes erros (CASELLA; BERGER, 2014).

A probabilidade de ocorrer o erro tipo I é representado por α e pode ser previamente definida, cujo complemento $(1 - \alpha)$ representa o coeficiente de confiança do teste. A probabilidade de ocorrer o erro tipo II é representado por β , a qual não é conhecida e nem pode ser especificada *a priori*. O complemento da probabilidade de incorrer no erro tipo II $(1 - \beta)$ é denominado poder do teste, representando a probabilidade de rejeitar a hipótese nula, dado que ela realmente é falsa (FERREIRA, 2009).

Em geral, se obtém uma amostra aleatória de tamanho n e especifica-se o valor de α , sendo que $0 \leq \alpha \leq 1$. O valor de α tem relação inversa com o valor de β . Assim, uma baixa probabilidade de ocorrência do erro tipo I está associada a alta probabilidade de ocorrência do erro tipo II e a única forma de reduzir a probabilidade do erro tipo II, fixado o valor de α , é aumentando o tamanho da amostra (FERREIRA, 2009).

Geralmente assume-se $\alpha = 0,01, 0,05$ e $0,10$ para controlar o erro tipo I, fixando assim a sua probabilidade de ocorrência em um valor relativamente pequeno (CASELLA; BERGER, 2014). Desta forma, o teste pode ser classificado em três categorias de acordo com o nível nominal α pré-estabelecido: a) teste conservativo: aquele que apresenta um valor real de taxa de erro tipo I inferior ao valor nominal α ; b) teste exato: aquele que apresenta um valor real de taxa de erro tipo I igual ao valor nominal α ; e c) teste liberal: aquele que apresenta um valor real de taxa de erro tipo I maior que o valor nominal α .

Um bom teste de hipóteses seria aquele que tem baixa probabilidade de ocorrência do erro tipo I, além de possuir também baixa probabilidade de ocorrência do erro tipo II. Portanto, ao desenvolver um novo teste, a avaliação do poder do teste, $P(1 - \beta)$, é um critério largamente empregado para que se escolha o melhor teste entre aqueles que possuem probabilidades do erro tipo I iguais ou inferiores ao nível nominal de significância fixado (FERREIRA, 2009). A avaliação dos testes é feita analiticamente ou por simulação Monte Carlo. A primeira alternativa é muito árdua de se aplicar em muitas situações reais, sendo que as simulações Monte Carlo, portanto, assumem um papel preponderante nesta tarefa.

2.9 Distribuições multivariadas

De acordo com Ferreira (2018), as distribuições multivariadas referem-se às distribuições de um vetor aleatório $\mathbf{X} = [X_1, X_2, \dots, X_p]^T$, em que ($p \geq 2$) representa o número de variáveis e os elementos X_i são as variáveis aleatórias univariadas com função de distribuição marginais $F_{X_i}(x_i)$. A função de distribuição conjunta do vetor aleatório é:

$$F_{\mathbf{X}}(\mathbf{x}) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_p \leq x_p).$$

A função de probabilidade conjunta do vetor aleatório é obtida quando o vetor aleatório é composto por componentes, que são variáveis aleatórias unidimensionais, X_i 's discretas, com

$i = 1, 2, \dots, p$, sendo definida por:

$$p_{\mathbf{X}}(\mathbf{x}) = P(X_1 = x_1, X_2 = x_2, \dots, X_p = x_p).$$

Já a função densidade de probabilidade conjunta do vetor aleatório, quando for absolutamente contínuo, pode ser obtida por meio da função de distribuição conjunta da seguinte forma:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\partial^p F_{\mathbf{X}}(\mathbf{x})}{\partial x_1 \partial x_2 \dots \partial x_p},$$

dado que $F_{\mathbf{X}}(\mathbf{x})$ é contínua em $-\infty < x_i < \infty$, para todo $i = 1, 2, \dots, p$.

Ainda de acordo com Ferreira (2018), se os componentes X_1, X_2, \dots, X_p possuem independência estatística, então:

$$F_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^p F_{X_i}(x_i),$$

válido para ambos os casos, contínuo e discreto. Outra forma alternativa no caso de independência, para os casos contínuo e discreto, é dada, respectivamente, por:

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^p f_{X_i}(x_i) \quad (2.28)$$

$$p_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^p p_{X_i}(x_i). \quad (2.29)$$

Este estudo aborda a distribuição normal multivariada. De acordo com Casella e Berger (2014), esta distribuição apresenta um papel central de um grande conjunto de estatísticas, isso porque ela é uma distribuição muito tratável analiticamente, que possui simetria, além do fato da existência do Teorema do limite central, que mostra que em condições moderadas, a distribuição normal pode ser utilizada para aproximar uma grande variedade de distribuições com grandes amostras. Este estudo também utiliza a distribuição normal contaminada, podendo ser entendida como a junção de duas distribuições normais multivariadas.

2.9.1 Distribuição normal multivariada

De acordo com Ferreira (2018), pode-se obter uma demonstração intuitiva da normal multivariada ao considerar p variáveis normais independentes X_1, X_2, \dots, X_p , cada uma possuindo uma função densidade de probabilidade marginal dada por:

$$f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi\sigma_{ii}}} \exp \left\{ -\frac{(x_i - \mu_i)^2}{2\sigma_{ii}} \right\},$$

em que μ_i e σ_{ii} são respectivamente a média e a variância para a i -ésima variável. Conforme a expressão (2.28), para obter a função densidade de probabilidade conjunta das p variáveis independentes do vetor aleatório \mathbf{X} , realiza-se o produtório das funções densidade de probabilidade marginais, da seguinte forma:

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^p f(x_i) = (2\pi)^{-\frac{p}{2}} \left(\prod_{i=1}^p \sigma_{ii} \right)^{-\frac{1}{2}} \exp \left\{ -\sum_{i=1}^p \frac{(x_i - \mu_i)^2}{2\sigma_{ii}} \right\}. \quad (2.30)$$

Rearranjando a expressão (2.30) utilizando matrizes e vetores, obtém-se:

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &= (2\pi)^{-\frac{p}{2}} \begin{bmatrix} \sigma_{11} & 0 & \dots & 0 \\ 0 & \sigma_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{pp} \end{bmatrix}^{-\frac{1}{2}} \times \\ &\times \exp \left\{ -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \begin{bmatrix} \sigma_{11} & 0 & \dots & 0 \\ 0 & \sigma_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{pp} \end{bmatrix}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \\ &= (2\pi)^{-\frac{p}{2}} |\mathbf{V}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \mathbf{V}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}, \end{aligned}$$

em que $\mathbf{V} = \text{diag}(\sigma_{ii})$ e $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_p]^\top$ representando respectivamente a matriz de covariâncias e o vetor de médias.

De acordo com Ferreira (2018), ao utilizar uma matriz de covariâncias positiva definida mais geral do que \mathbf{V} , obtém-se a função densidade de probabilidade da normal multivariada, dada por:

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-\frac{p}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}, \quad (2.31)$$

em que

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{bmatrix}.$$

2.9.2 Distribuição normal multivariada contaminada

De acordo com Ferreira (2018), um vetor aleatório $\mathbf{X} = [X_1, X_2, \dots, X_p]^\top \in \mathbb{R}^p$ com distribuição normal multivariada contaminada possui a seguinte função densidade de probabilidade dada por:

$$f_{\mathbf{X}}(\mathbf{x}) = \delta (2\pi)^{-\frac{p}{2}} |\boldsymbol{\Sigma}_1|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_1)^\top |\boldsymbol{\Sigma}_1|^{-1} (\mathbf{x} - \boldsymbol{\mu}_1) \right\} + \\ + (1 - \delta) (2\pi)^{-\frac{p}{2}} |\boldsymbol{\Sigma}_2|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_2)^\top |\boldsymbol{\Sigma}_2|^{-1} (\mathbf{x} - \boldsymbol{\mu}_2) \right\},$$

em que $\boldsymbol{\Sigma}_1$ e $\boldsymbol{\Sigma}_2$ são as matrizes de covariâncias positivas definidas, $\boldsymbol{\mu}_1$ e $\boldsymbol{\mu}_2$ são os vetores de médias e δ está contido entre $[0,1]$, representado a parcela de dados contaminantes, de forma complementar, $(1 - \delta)$ são os dados gerados sem contaminação.

3 MÉTODOS

Esta seção apresentará todos os mecanismos estatísticos necessários para realização do teste de hipóteses original e dos testes propostos computacionalmente intensivos e assintóticos baseados em estimadores robustos e clássicos. Além de descrever como será a avaliação dos seus respectivos desempenhos, realizados por meio de simulações Monte Carlo para obter a taxa de erro tipo I e o poder de cada teste.

3.1 Teste de hipóteses

Considera-se uma amostra aleatória $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$, em que \mathbf{X}_i é uma variável aleatória p -dimensional, retirada de uma população com distribuição normal multivariada, com vetor de médias $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$, sendo que ambos parâmetros são desconhecidos.

O teste de hipóteses de interesse considera que a matriz populacional de covariâncias possui estrutura de simetria composta, ou seja, todas as correlações entre pares de diferentes variáveis são iguais (ρ) e todas as variâncias são homogêneas (σ^2). As hipóteses nula e alternativa dos testes são dadas por: $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ versus $H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0$, sendo que:

$$\boldsymbol{\Sigma}_0 = \sigma^2 \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{bmatrix} = \sigma^2[(1 - \rho)\mathbf{I} + \rho\mathbf{J}], \quad (3.1)$$

em que \mathbf{J} é uma matriz com dimensões $(p \times p)$ e composta por valores iguais a 1 e \mathbf{I} é uma matriz identidade com dimensões $(p \times p)$.

O teste da hipótese nula é realizado e por meio do teste da razão de verossimilhanças considerando a pressuposição de normalidade multivariada dos dados, cuja estatística possui distribuição assintótica qui-quadrado. Lembrando que o desempenho deste teste é fortemente afetado se houver desvios de normalidade nos dados, que podem ser causados por *outliers*.

Será descrito a seguir, o teste original da razão de verossimilhanças, LRTO, e os testes computacionalmente intensivos e assintóticos baseados em estimadores robustos e clássicos, LRTR, BLRTR e BLRTO, que este estudo propõe.

3.1.1 Teste da razão de verossimilhanças original (LRTO)

A estatística do teste original para a hipótese da estrutura de simetria composta para matriz de covariâncias é obtida a partir de uma amostra aleatória de uma população normal multivariada, calculando-se os estimadores \mathbf{S}_0 sob a hipótese nula e \mathbf{S} sob a hipótese alternativa da matriz de covariâncias, obtidos conforme as expressões (2.3) e (2.1), respectivamente. Com a correção de Box (1949), a estatística é dada da seguinte forma:

$$\chi_c^2 = \left[(n-1) - \frac{p(p+1)^2(2p-3)}{6(p-1)(p^2+p+4)} \right] (\ln|\mathbf{S}_0| - \ln|\mathbf{S}|), \quad (3.2)$$

que, sob a hipótese nula, possui distribuição assintótica qui-quadrado com $f = p(p+1)/2 - 2$ graus de liberdade.

3.1.2 Teste da razão de verossimilhanças baseado em estimadores robustos (LRTR)

Nesta subseção apresenta-se o primeiro teste de hipóteses assintótico proposto, baseado em estimadores robustos. Sua estatística do teste é obtida ao substituir na estatística do LRTO, na expressão (3.2), o estimador da matriz de covariâncias sob a hipótese alternativa \mathbf{S} e o estimador da matriz de covariâncias sob a hipótese nula \mathbf{S}_0 , pelos seus respectivos estimadores robustos, \mathbf{S}^* e \mathbf{S}_0^* , obtidos pelo método Comedian, descrito na subseção 2.5.1. O estimador robusto da matriz de covariâncias é dado por:

$$\mathbf{S}_0^* = S_c^{*2} \begin{bmatrix} 1 & r^* & \dots & r^* \\ r^* & 1 & \dots & r^* \\ \vdots & \vdots & \ddots & \vdots \\ r^* & r^* & \dots & 1 \end{bmatrix} = S_c^{*2} [(1-r^*)\mathbf{I} + r^*\mathbf{J}], \quad (3.3)$$

em que, S_c^{*2} é o estimador não viesado da variância populacional σ^2 e r^* é o estimador comum não viesado da correlação populacional ρ , dados por:

$$S_c^{*2} = \frac{1}{p} \sum_{k=1}^p S_{kk}^* \quad (3.4)$$

e

$$r^* = \frac{2}{p(p-1)S_c^{*2}} \sum_{i=1}^{p-1} \sum_{k=i+1}^p S_{ik}^*. \quad (3.5)$$

Assim, a estatística do teste é

$$\chi_c^2 = \left[(n-1) - \frac{p(p+1)^2(2p-3)}{6(p-1)(p^2+p+4)} \right] (\ln|\mathbf{S}_0^*| - \ln|\mathbf{S}^*|), \quad (3.6)$$

que sob H_0 considerou-se que possui distribuição qui-quadrado assintótica, com $f = p(p+1)/2 - 2$ graus de liberdade.

3.1.3 Teste *bootstrap* paramétrico baseado em estimadores robustos (BLRTR)

Devido a incerteza que o primeiro teste proposto na seção 3.1.2, possua como distribuição nula de sua estatística a distribuição qui-quadrado assintótica com $f = p(p+1)/2 - 2$ graus de liberdade, esta proposta utilizou o método computacionalmente intensivo, *bootstrap* paramétrico, como uma forma de buscar uma solução para esta questão. Sendo assim, nesta proposta o teste utiliza a estatística na expressão (3.6), sendo que a distribuição nula passa a utilizar o método *bootstrap* paramétrico com estimadores robustos Comedian, conforme descrição apresentada na sequência.

Considere χ_{LTR}^2 o valor da estatística obtida na amostra original, utilizando os estimadores robustos Comedian das matrizes de covariâncias. A partir desta mesma amostra original, inicialmente, são estimados os parâmetros $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$, respectivamente, utilizando os parâmetros robustos $\bar{\mathbf{X}}^*$ e \mathbf{S}_0^* , em que este último estimador utiliza-se do estimador robusto \mathbf{S}^* e a expressão (3.4) para ser definido.

Os estimadores $\bar{\mathbf{X}}^*$ e \mathbf{S}_0^* são utilizados para especificar a função densidade da normal multivariada $N_p(\bar{\mathbf{X}}^*, \mathbf{S}_0^*)$, da qual serão geradas B reamostragens *bootstrap* paramétrico sob a hipótese nula. A especificação do estimador \mathbf{S}_0^* , como na expressão (3.3), é a forma que a hipótese nula é imposta na geração da distribuição nula da estatística do teste. Para cada reamostragem do modelo $N_p(\bar{\mathbf{X}}^*, \mathbf{S}_0^*)$, a estatística associada é computada.

Se χ_b^{*2} for o b -ésimo valor da estatística do teste computada na b -ésima reamostragem *bootstrap* paramétrico, incluindo entre elas o valor computado na amostra original, então o valor- p do teste é dado por

$$\text{valor-}p = \frac{\sum_{b=1}^{B+1} I(\chi_b^{*2} \geq \chi_{LTR}^2)}{B+1},$$

em que, $I(\chi_b^{*2} \geq \chi_{LRTO}^2)$, é a função indicadora que retorna 1 se a desigualdade $\chi_b^{*2} \geq \chi_{LRTO}^2$ for verdadeira e 0 em caso contrário, B é o número de reamostragens de *bootstrap*, cujo trabalho considerou $B = 2000$ e soma-se 1 devido a estatística realizada da amostra original.

3.1.4 Teste *bootstrap* paramétrico baseado em estimadores clássicos (BLRTO)

O terceiro teste proposto também utiliza-se do método *bootstrap* paramétrico para gerar a distribuição nula da estatística, sendo similar ao segundo teste proposto na subseção 3.1.3. A diferença, porém, reside na utilização dos estimadores não robustos. Portanto, a partir de uma amostra original, computa-se a estatística original χ_{LRTO}^2 usando a expressão (3.2) e obtêm-se os estimadores $\bar{\mathbf{X}}$ do vetor de médias e \mathbf{S}_0 da matriz de covariâncias utilizando-se para isso a expressão (2.3), que calcula os estimadores clássicos. Estes estimadores definem o modelo normal multivariado ajustado $N_p(\bar{\mathbf{X}}, \mathbf{S}_0)$, impondo H_0 , que irá fornecer as B reamostragens de *bootstrap*.

Para cada reamostragem a estatística associada é computada usando a expressão (3.2), com \mathbf{S}_0 e \mathbf{S} computados a partir da reamostragem *bootstrap* paramétrico gerada, obtendo-se o b -ésimo valor da distribuição *bootstrap* paramétrico χ_b^2 . O valor- p do teste é dado por:

$$\text{valor-}p = \frac{\sum_{b=1}^{B+1} I(\chi_b^2 \geq \chi_{LRTO}^2)}{B+1},$$

em que, $I(\chi_b^2 \geq \chi_{LRTO}^2)$ é a função indicadora que retorna 1 ou 0, conforme a desigualdade seja respectivamente verdadeira ou falsa, dividido por $B+1$, em que B é o número de reamostragens de *bootstrap*, cujo trabalho considerou $B = 2000$, mais 1, devido a estatística realizada da amostra original.

3.2 Avaliação do desempenho dos testes

Os cenários de simulação da taxa de erro tipo I e poder foram baseados nos estudos realizados por Morris, Payton e Santorico (2011), ao criarem o teste robusto, MCPT, para desvios de normalidade. É importante ressaltar que a proposta do teste MCPT era ser robusto a dados provenientes de outras distribuições, por isso considerou-se nas simulações outras distribuições de probabilidade além da normal multivariada.

Como o foco do presente estudo é criar testes robustos para desvios de normalidade causados por *outliers*, considerou-se apenas as distribuições normal multivariada, como no estudo de Morris, Payton e Santorico (2011), e a normal multivariada contaminada. Portanto, o desempenho dos testes LRTR, LRTO, BLRTR e BLTO só podem ser comparados com o desempenho do teste MCPT ao considerar dados normalmente distribuídos.

Diante da taxa de erro tipo I e poder, avaliou-se os três testes propostos, LRTR, BLRTR e BLTO, juntamente com o teste original LRTO, por meio das distribuições normal multivariada e normal multivariada contaminada. Além de verificar se os testes que abordaram os estimadores robustos Comedian obtiveram melhores desempenhos em comparação aos testes que abordaram estimadores clássicos.

3.2.1 Erro tipo I

Para avaliar a taxa de erro tipo I dos testes LRTO, LRTR, BLRTR e BLRTO realizou-se 1.000 simulações Monte Carlo, no R Core Team (2018), das distribuições normal multivariada e normal multivariada contaminada, as quais foram abordadas em dois cenários sob a hipótese nula. O primeiro cenário considerou as seguintes combinações a serem simuladas: $n = 5, 10, 15, 25, 50, 100$; $p = 3$; $\sigma^2 = 9$ e $\rho = 0,6$, já o segundo cenário considerou: $n = 25$; $p = 3, 5, 10, 20$; $\sigma^2 = 9$ e $\rho = 0,6$, representando respectivamente tamanho amostral, número de variáveis, variância e correlação. A escolha destes cenários se baseou nos estudos realizados por Morris, Payton e Santorico (2011).

Desta forma, avaliou-se a taxa de erro tipo I nestes dois cenários considerando inicialmente a distribuição normal multivariada, sob H_0 como:

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-\frac{p}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\},$$

em que o $\boldsymbol{\mu} = [0, 0, \dots, 0]^\top$ e $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$. Posteriormente para avaliar a taxa de erro tipo I na presença de desvios de normalidade, os dados foram simulados da distribuição normal multivariada contaminada, ainda sob H_0 , como:

$$f_{\mathbf{X}}(\mathbf{x}) = \delta (2\pi)^{-\frac{p}{2}} |\boldsymbol{\Sigma}_1|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_1)^\top |\boldsymbol{\Sigma}_1|^{-1} (\mathbf{x} - \boldsymbol{\mu}_1) \right\} + \\ + (1 - \delta) (2\pi)^{-\frac{p}{2}} |\boldsymbol{\Sigma}_2|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_2)^\top |\boldsymbol{\Sigma}_2|^{-1} (\mathbf{x} - \boldsymbol{\mu}_2) \right\},$$

em que $\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_0$ e $\boldsymbol{\Sigma}_2 = 2\boldsymbol{\Sigma}_0$, cuja proporção de contaminação foi de 5%, 10% e 30%.

A proporção de vezes que a hipótese nula é rejeitada incorretamente foi calculada para cada teste da seguinte forma:

$$\hat{\alpha} = \frac{\sum_{i=1}^N \mathbf{I}(\text{valor-}p_i \leq \alpha)}{N},$$

em que α é o nível nominal de significância, escolhido como: 0,01, 0,05 e 0,10 e N é o número de simulações Monte Carlo. No entanto, como o cômputo da taxa de erro tipo I por meio de simulações Monte Carlo não é isento de erro, utilizou-se o teste binomial exato para avaliar este erro, que utiliza uma relação entre a distribuição F e a distribuição binomial. Se a probabilidade de sucesso é $p = \alpha$ e w é o número de rejeições da hipótese H_0 nas simulações Monte Carlo em um dado cenário, a estatística do teste binomial é computada por:

$$F_c = \binom{w+1}{N-w} \left(\frac{1-\alpha}{\alpha} \right),$$

que, sob a hipótese nula, segue uma distribuição F com $v_1 = (N-w)$ e $v_2 = 2(w+1)$. Desta forma, rejeita-se H_0 quando o $F_c \leq F_{\alpha/2}$ ou $F_c \geq F_{(1-\alpha)/2}$, com v_1 e v_2 graus de liberdade (OLIVEIRA; FERREIRA, 2010). Após obter a taxa de erro tipo I de todos os testes, eles foram classificados como conservativo, exato ou liberal, de acordo com os resultados do teste binomial.

3.2.2 Poder do teste

Os valores do poder dos testes LRTO, LRTR, BLRTO e BLRTR em detectar matrizes com estruturas distintas a de simetria composta foram estimados por meio de 1.000 simulações Monte Carlo no R Core Team (2018). Baseado no estudo realizado por Morris, Payton e Santorico (2011), considerou-se dois cenários para simular o poder de cada teste. No primeiro cenário, o poder foi simulado considerando dados multivariados gerados a partir de distribui-

ções com a matriz de covariâncias com estrutura de heterocedasticidade, como:

$$\mathbf{\Sigma}_H = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 + d/(p-1) & 0 & \dots & 0 \\ 0 & 0 & 1 + 2d/(p-1) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 + d \end{bmatrix}, \quad (3.7)$$

em que d representa a diferença entre o primeiro e o último (ou menor e maior) elementos da diagonal principal, considerando $n = 5, 10, 25, 50$, $p = 3$, $d = 4$ e $\rho = 0$.

Portanto, as simulações foram realizadas por meio de duas distribuições, a distribuição normal multivariada, sob H_1 , com parâmetros $\boldsymbol{\mu} = [0, 0, \dots, 0]^\top$ e $\boldsymbol{\Sigma} = \mathbf{\Sigma}_H$, e a distribuição normal multivariada contaminada também sob H_1 , com parâmetros $\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \mathbf{\Sigma}_H$ e $\boldsymbol{\Sigma}_2 = 2\mathbf{\Sigma}_H$, contaminando pelo vetor de médias e pela matriz de covariâncias, cuja proporção de contaminação foi de 5%, 10% e 30%.

No segundo cenário o poder dos testes é simulado utilizando uma matriz autoregressiva de primeira ordem, dada da seguinte forma:

$$\mathbf{\Sigma}_{AP} = \frac{\sigma^2}{1 - \rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{p-1} \\ \rho & 1 & \rho & \dots & \rho^{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \dots & 1 \end{bmatrix}, \quad (3.8)$$

em que $\sigma^2/(1 - \rho^2)$ é a variância comum das p variáveis, ρ é a correlação entre sucessivas observações das variáveis, considerando $n = 10, 25, 50, 75$, $p = 5$, $\sigma^2 = 1$ e $\rho = 0,6$.

Neste cenário, as simulações também foram realizadas por dados multivariados gerados pela distribuição normal multivariada, sob H_1 , com parâmetros $\boldsymbol{\mu} = [0, 0, \dots, 0]^\top$ e $\boldsymbol{\Sigma} = \mathbf{\Sigma}_{AP}$ e pela distribuição normal multivariada contaminada, também sob H_1 , com parâmetros $\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \mathbf{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = 2\mathbf{\Sigma}_{AP}$ contaminando pelo vetor de médias e pela matriz de covariâncias, cuja proporção de contaminação foi de 5%, 10% e 30%.

Com intuito de avaliar melhor o desempenho dos testes, realizou-se simulações adicionais sobre seus poderes, considerando as matrizes $\mathbf{\Sigma}_H$ e $\mathbf{\Sigma}_{AP}$ do estudo de Morris, Payton e Santorico (2011), porém, com diferentes configurações de simulações. Com a matriz $\mathbf{\Sigma}_H$, os poderes dos testes foram simulados considerando $n = 25$, $p = 20$, $d = 4$, $\sigma^2 = 1$ e $\alpha = 0,05$,

com as distribuições normal multivariada contaminada ($\delta = 0,10$ e $\delta = 0,30$) e normal multivariada. Ainda com a matriz de covariâncias $\mathbf{\Sigma}_H$, simulou-se o poder dos testes com a distribuição normal multivariada, considerando $n = 25$, $p = 20$, $d = 10$ e 20 , $\sigma^2 = 1$ e $\alpha = 0,05$. Sendo, $\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$ e $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_H$ os parâmetros da distribuição normal multivariada e $\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_H$ e $\mathbf{\Sigma}_2 = 2\mathbf{\Sigma}_H$ os parâmetros das distribuições normal multivariada contaminada ($\delta = 0,10$ e $\delta = 0,30$).

Já no segundo cenário, com a matriz $\mathbf{\Sigma}_{AP}$, simulou-se o poder dos testes considerando $n = 25$, $p = 20$, $\sigma^2 = 9$, $\rho = 0,6$ e $\alpha = 0,05$ com as distribuições normal multivariada contaminada ($\delta = 0,10$ e $\delta = 0,30$), com parâmetros $\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_{AP}$ e $\mathbf{\Sigma}_2 = 2\mathbf{\Sigma}_{AP}$ e normal multivariada, com parâmetros $\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$ e $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_{AP}$. Ainda no segundo cenário, simulou-se o poder dos testes considerando $n = 10, 25, 50, 75$, $p = 5$, $\sigma^2 = 1$ e $\rho = 0,6$ com as distribuições normal multivariada contaminada ($\delta = 0,05$, $\delta = 0,10$ e $\delta = 0,30$), com parâmetros $\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_{AP}$ e $\mathbf{\Sigma}_2 = \mathbf{\Sigma}_{AP}$, contaminando apenas pelo vetor de médias.

4 RESULTADOS E DISCUSSÃO

Nas próximas subseções são apresentados os resultados das taxas de erro tipo I e poder, para os testes LRTO, LRTR, BLRTO e BLRTR, em dois cenários de simulação. É válido lembrar que o teste da razão de verossimilhanças para matriz de covariâncias com estrutura de simetria composta possui o pressuposto de normalidade multivariada dos dados amostrais, tornando-se pouco robusto na presença de *outliers*. Portanto, espera-se que os testes propostos baseados em estimadores robustos obtenham bons desempenhos na presença de *outliers*, como também no cenário ideal de dados normalmente distribuídos.

Diante de uma avaliação dos resultados sobre a taxa de erro tipo I, no primeiro e no segundo cenário, é possível notar padrões de comportamentos semelhantes dos testes LRTO, LRTR, BLRTO e BLRTR, quando considera-se a distribuição normal multivariada com níveis de contaminação de 10% e 5%. Optou-se então, por apresentar os resultados das simulações com nível de contaminação de 10% e 30%, e os resultados das simulações com contaminação 5% encontram-se nos apêndices, nas Tabelas A.1 e A.2.

Em relação ao poder de cada teste, optou-se por apresentar no primeiro e no segundo cenário os resultados das simulações considerando 5% de significância, pois os resultados das simulações para os três níveis de significância $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$, apresentaram comportamentos semelhantes. Os demais resultados de 10% e 1% encontram-se nos apêndices, nas Tabelas A.3, A.4, A.5, A.6, A.7, A.8, A.9 e A.10.

Os resultados das simulações adicionais sobre o poder de cada teste, também serão apresentados a seguir, exceto os que encontram-se nos apêndices nas Tabelas A.11, A.12 e A.13, referentes as simulações realizadas com a matriz Σ_{AP} , do segundo cenário, em que os dados foram contaminados apenas pelo vetor de médias. Pois estes resultados foram semelhantes aos resultados dos dados contaminados pelo vetor de médias e pela matriz de covariâncias.

4.1 Taxa de erro tipo I dos testes

Os resultados para avaliação do erro tipo I foram gerados por meio das distribuições normal multivariada e normal multivariada contaminada, 1.000 simulações Monte Carlo no R (R Core Team, 2018). A taxa de erro tipo I dos testes assintóticos e computacionalmente intensivos, com estimadores clássicos ou robustos, considerando os valores nominais de significância de α iguais a 0,10, 0,05 e 0,01, foram computadas e são apresentadas na sequência.

Os cenários de simulação foram estabelecidos conforme Morris, Payton e Santorico (2011), avaliaram a taxa de erro tipo I do teste robusto MCPT, para simetria composta. Este teste é uma proposta robusta para desvios de normalidade causados por dados são provenientes de outras distribuições de probabilidade. O objetivo do presente estudo é diferente em relação a este trabalho, propondo-se testes robustos em relação a desvios de normalidades dos dados causados por *outliers*.

Portanto, abordou-se a distribuição normal multivariada, como no estudo de Morris, Payton e Santorico (2011), podendo assim, comparar os testes LRTO, LRTR, BLTO e BLRTR com o teste MCPT. Abordou-se também a distribuição normal multivariada contaminada com intuito de gerar *outliers* no conjunto de dados, neste contexto os testes LRTO, LRTR, BLTO e BLRTR não serão comparados ao teste MCPT.

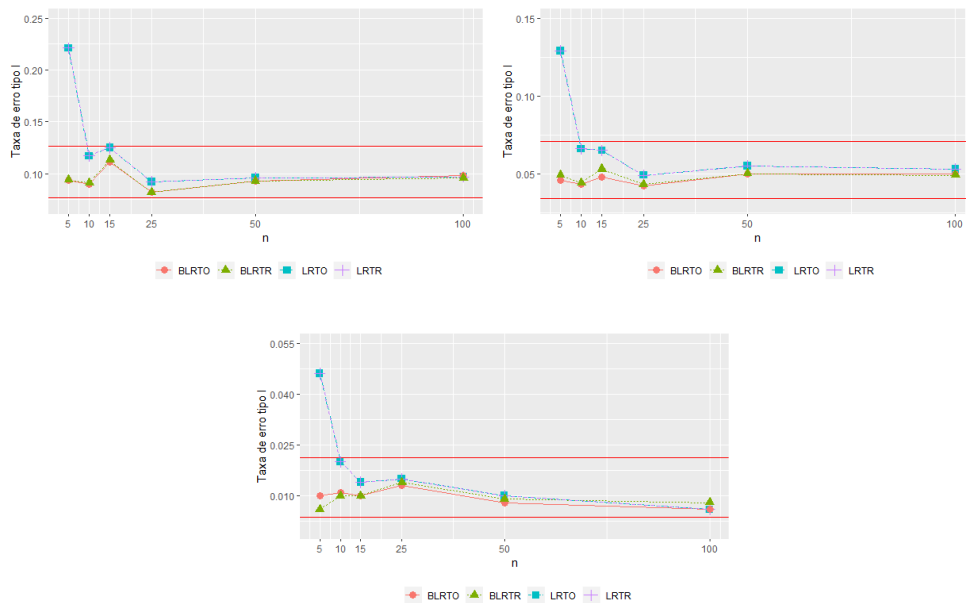
Conforme Hampel et al. (2011), mesmo que os dados sejam de alta qualidade, ou seja, gerados por meio da distribuição normal multivariada, os métodos robustos podem fornecer uma melhoria notável em relação ao modelo estatístico clássico. Portanto, é interessante analisar os resultados das taxas de erro tipo I do teste MCPT, e dos testes LRTO, LRTR, BLRTR e BLRTO, diante de circunstâncias ideais, isto é, sem a presença de *outliers* e com dados distribuídos normalmente. Por isso, o primeiro e o segundo cenários de simulação, representados, respectivamente nas Figuras 4.1 e 4.2, foram realizadas por meio de dados gerados pela distribuição normal multivariada.

De acordo com Morris, Payton e Santorico (2011), nas mesmas configurações de simulações da Figura 4.1, o teste MCPT apresentou taxas reais de erro tipo I maiores que o valor nominal $\alpha = 0,05$ para pequenos valores amostrais e taxas reais de erro tipo I iguais ao valor nominal $\alpha = 0,05$ à medida que aumenta o tamanho amostral.

Na Figura 4.1, as taxas de erro tipo I em função do n , dos testes LRTO e LRTR estão sobrepostas, como também acontece para as taxas de erro tipo I dos testes BLRTR e BLRTO. Verifica-se que os testes LRTO e LRTR apresentaram taxas reais de erro tipo I maiores que todos valores nominais α para $n = 5$. Para os demais valores de n , estes testes foram exatos. Já os testes BLRTO e BLRTR apresentaram valores reais das taxas de erro tipo I iguais aos valores nominais α para todos os tamanhos de amostras.

Na Figura 4.2, que representa o segundo cenário de simulação, verifica-se que os testes LRTO e LRTR apresentaram taxas reais de erro tipo I maiores que os valores nominais $\alpha = 0,10$, $0,05$ e $0,01$ para $p = 20$. Estes mesmos testes também foram liberais para $p = 10$ ao considerar

Figura 4.1 – Primeiro cenário: Simulação do erro tipo I para os testes de simetria composta ($p = 3$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada, considerando respectivamente $\alpha = 0,1$, $\alpha = 0,05$ e $\alpha = 0,01$



Fonte: Da autora (2020).

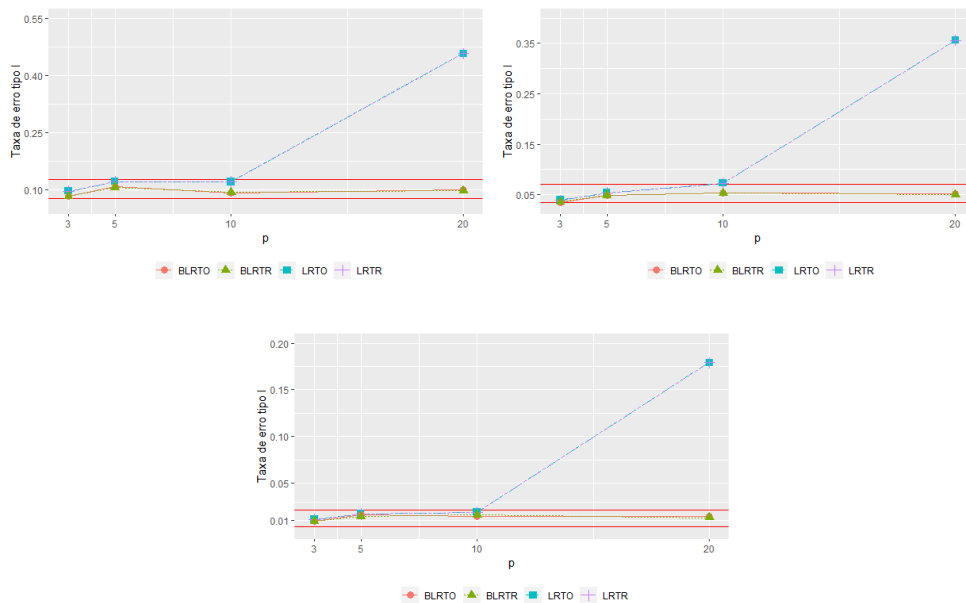
o valor nominal $\alpha = 0,05$. Os testes computacionalmente intensivos, BLRTO e BLRTR, foram exatos, com valores reais da taxa de erro tipo I iguais aos valores nominais $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$, para todos os valores de p .

Os resultados e Morris, Payton e Santorico (2011), provenientes no mesmo contexto de simulação da Figura 4.2, indicam que o teste MCPT obteve taxas reais de erro tipo I iguais ao valor nominal de $\alpha = 0,05$ para $p = 3, 5$ e 10 , já para $p = 20$, apresentou taxa real de erro tipo I superior ao valor nominal de $\alpha = 0,05$.

Diante de uma avaliação conjunta dos resultados sobre as taxas de erro tipo I, nas Figuras 4.1 e 4.2, verifica-se que os testes assintóticos, foram liberais para pequenos tamanhos amostrais ou para grande número de variáveis. Os mesmos testes em ambos os cenários controlaram a taxa de erro tipo I para pequenos valores de p , com maiores tamanhos de n . Já os testes computacionalmente intensivos foram exatos tanto no primeiro como no segundo cenário, controlando a taxa de erro tipo I mesmo com pequenos tamanhos amostrais ou grandes números de variáveis no modelo.

As taxas de erro tipo I dos testes assintóticos baseados em estimadores clássicos e robustos foram praticamente idênticas, como também acontece para a taxa de erro tipo I dos testes computacionalmente intensivos baseados em estimadores robustos e clássicos. Desta forma,

Figura 4.2 – Segundo cenário: Simulação do erro tipo I para os testes de simetria composta ($n = 25$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada, considerando respectivamente $\alpha = 0,1$, $\alpha = 0,05$ e $\alpha = 0,01$



Fonte: Da autora (2020).

não verificou-se melhorias ao abordar o estimador robusto em comparação ao estimador clássico em circunstâncias ideais. Estes resultados vão de encontro ao que afirmam Hapel et al. (1986), que alegam que métodos robustos poderiam fornecer melhorias notáveis em relação ao modelo estatístico clássico mesmo com dados de alta qualidade.

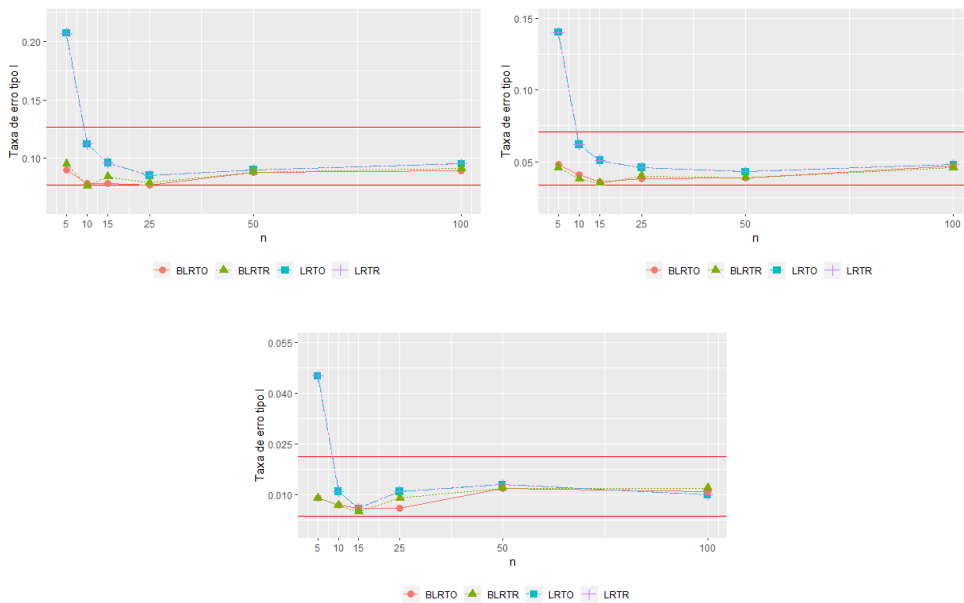
Pode-se verificar também, que ao comparar os resultados da taxa de erro tipo I do teste MCPT com os testes LRTR, LRTRO, BLRTR e BLRTO, os testes computacionalmente intensivos superam o teste MCPT. Pois eles são exatos no primeiro no segundo cenários, para todos os valores de n e p , enquanto o MCPT é liberal para $n = 5, 10$ e 15 , no primeiro cenário, e para $p = 20$, no segundo cenário.

Os resultados nas Figuras 4.3 e 4.4, são referentes a taxa de erro tipo I de cada um dos testes, considerando $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$ respectivamente. As simulações foram realizadas por meio de dados com a presença de *outliers*, gerados pela distribuição normal multivariada contaminada, com $\delta = 0,10$, representando o nível de contaminação gerado pelo vetor de médias e pela matriz de covariâncias.

Os testes assintóticos LRTO e LRTR, na Figura 4.3, obtiveram taxas reais de erro tipo I maiores que os valores nominais de α para $n = 5$, apresentado ser liberais para pequenas amostras, e para os demais valores de n os mesmos testes foram exatos, em todos os níveis nominais

de α considerados. Ainda na Figura 4.3, que representa o primeiro cenário de simulação, na presença de *outliers*, os testes BLRTR e BLRTO foram exatos para todos os tamanhos amostrais em todos os níveis nominais α , pois os resultados de suas taxas reais de erro tipo I foram iguais aos valores nominais de α pré-definidos.

Figura 4.3 – Primeiro cenário: Simulação do erro tipo I para os testes de simetria composta ($p = 3$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\delta = 0,10$), considerando respectivamente $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$

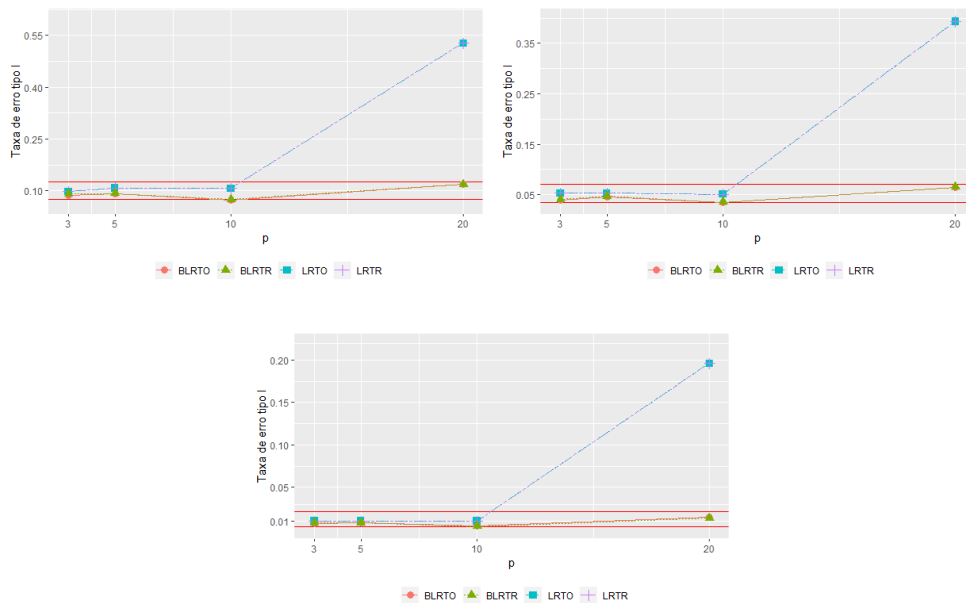


Fonte: Da autora (2020).

De acordo com resultados na Figura 4.4, no segundo cenário, na presença de *outliers*, os testes computacionalmente intensivos, BLRTR e BLRTO, foram exatos para todos valores de p , nos níveis nominais de $\alpha = 0,05$ e $\alpha = 0,01$. Considerando o nível nominal de significância $\alpha = 0,10$ com $p = 10$, os testes BLRTR e BLRTO foram conservativos, para os demais valores de p exatos. Os testes LRTO e LRTR, apresentam taxas reais de erro tipo I maiores que os valores nominais α para $p = 20$, sendo considerados liberais quando há um aumento do número de variáveis no modelo em estudo, mantido fixo o tamanho amostral n .

Nas Figuras 4.5 e 4.6, as simulações foram realizadas com dados gerados sob a hipótese nula, por meio da distribuição normal multivariada contaminada, com $\delta = 0,3$, cujos *outliers* são provenientes do vetor de médias e da matriz de covariâncias. Portanto, esta configuração de simulação foi a que gerou mais *outliers* no conjunto de dados. Na Figura 4.5, pôde-se verificar que os testes LRTO e LRTR apresentaram taxas reais de erro tipo I maiores que os valores nominais de significância α para $n = 5$. Os mesmos testes foram liberais para $\alpha = 0,10$ com

Figura 4.4 – Segundo cenário: Simulação do erro tipo I para os testes de simetria composta ($n = 25$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\delta = 0,1$), considerando respectivamente $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$



Fonte: Da autora (2020).

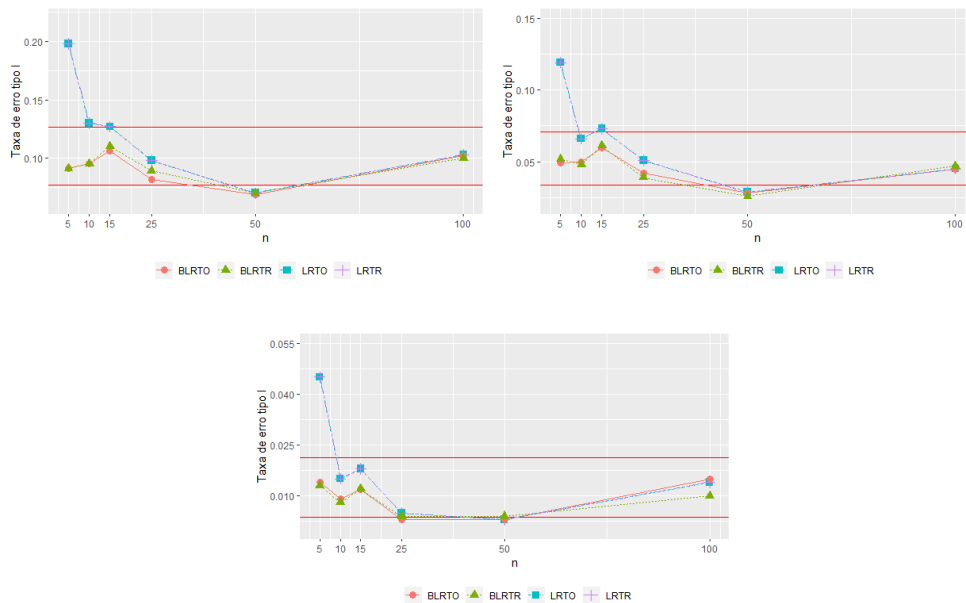
$n = 10$ e para $\alpha = 0,05$ com $n = 15$. Já para $n = 100$, em todos níveis nominais de significância α , ambos os testes foram exatos.

Ainda sobre os resultados na Figura 4.5, verificou-se que os quatro testes, LRTO, LRTR, BLRTR e BLRTO, apresentaram valores reais de taxa de erro tipo I inferiores aos valores nominais α quando $n = 50$. O teste BLRTO foi conservativo quando $n = 25$ para $\alpha = 0,01$ e para os demais valores de n , em todos os níveis nominais α , os testes BLRTR e BLRTO foram exatos.

De acordo com os resultados do segundo cenário de simulação, com alta contaminação ($\delta = 0,3$), na Figura 4.6, os testes BLRTO e BLRTR apresentaram taxas reais de erro tipo I maiores que os valores nominais α para $p = 20$, os mesmos testes também foram liberais para $\alpha = 0,10$ e $0,05$ com $p = 10$. Os testes assintóticos, em todos os níveis nominais de significância α , foram liberais com $n = 10$, e ainda mais liberais que os testes computacionalmente intensivos, quando $n = 20$.

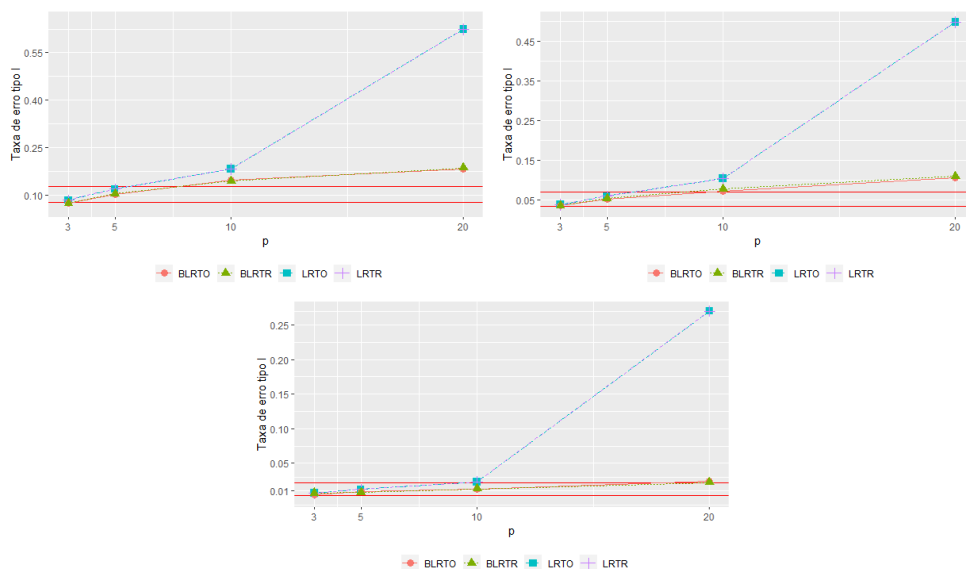
Os testes computacionalmente intensivos foram, praticamente exatos para todos os níveis nominais de significância α , ao considerar 10% de contaminação nos dados. Quando os dados foram contaminados 30%, os mesmos testes foram conservativos no primeiro cenário para $n = 50$ e liberais para $p = 10$ e 20 com $\alpha = 0,10$, no segundo cenário. Os testes assintóticos também foram conservativos para $n = 50$, no primeiro cenário, porém, foram expressivamente

Figura 4.5 – Primeiro cenário: Simulação do erro tipo I para os testes de simetria composta ($p = 3; \sigma^2 = 9; \rho = 0,6$) com a distribuição normal multivariada contaminada ($\delta = 0,3$), considerando respectivamente $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$



Fonte: Da autora (2020).

Figura 4.6 – Segundo cenário: Simulação do erro tipo I para os testes de simetria composta ($n = 25; \sigma^2 = 9; \rho = 0,6$) com a distribuição normal multivariada contaminada ($\delta = 0,3$), considerando respectivamente $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$



Fonte: Da autora (2020).

liberais, no segundo cenário, em comparação com os testes BLRTR e BLRTR. Desta forma, os testes computacionalmente intensivos controlaram mais a taxa de erro tipo I, do que os testes assintóticos, tanto para os dados distribuídos normalmente quanto na presença de *outliers*.

4.2 Poder dos testes

Os resultados do poder dos testes assintóticos (LRTR e LRTO) e computacionalmente intensivos (BLRTR e BLRTO), baseados em estimadores clássicos e robustos, foram obtidos em 1.000 simulações Monte Carlo no R (R Core Team, 2018) das distribuições normal multivariada e normal multivariada contaminada. Isso permitiu que dados fossem gerados sob a hipótese alternativa e o poder dos testes em detectar matrizes de covariâncias com estruturas distintas a de simetria composta pudesse ser determinado, usando-se estimadores clássicos e robustos.

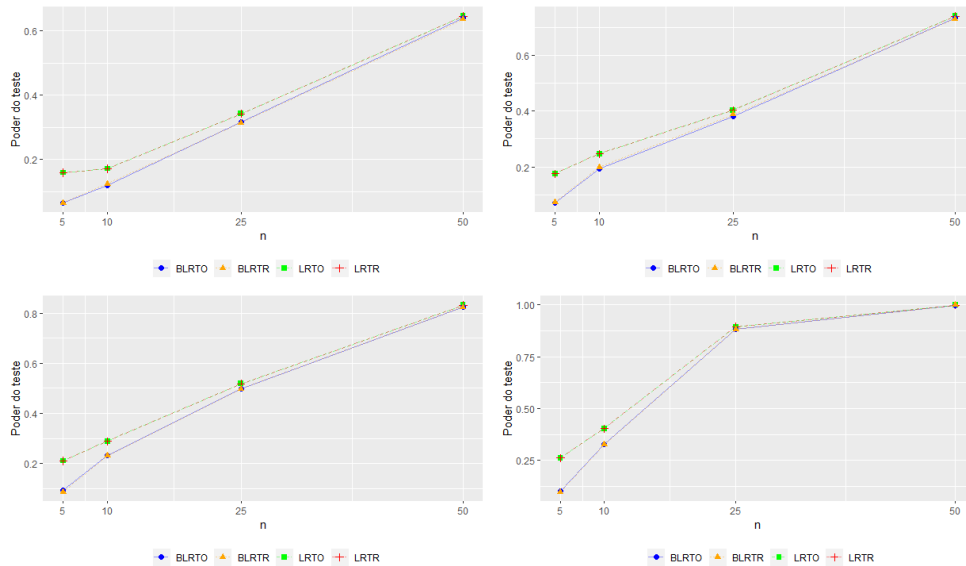
Nas Figuras 4.7 e 4.8 o poder dos testes são obtidos considerando respectivamente dados gerados pelas distribuições normal multivariada contaminada, com níveis de contaminação de $\delta = 0,3$, $\delta = 0,1$ e $\delta = 0,05$, e normal multivariada. A utilização da distribuição normal multivariada contaminada para gerar dados sob a hipótese alternativa, é uma forma de desarranjar ainda mais os dados, desviando-os da hipótese nula de simetria composta. Uma vez que a matriz de covariâncias da distribuição é formada em função das duas covariâncias e dos dois vetores de médias das distribuições envolvidas na mistura.

Na Figura 4.7 encontra-se os resultados do poder dos testes LRTR, LRTO, BLRTR e BLRTO, que foram simulados sob a hipótese alternativa, considerando a matriz de covariâncias Σ_H definida da expressão (3.7). Já na Figura 4.8, os resultados do poder dos testes, foram simulados sob a hipótese alternativa, considerando a matriz de covariâncias Σ_{AP} definida da expressão (3.8).

No estudo realizado por Morris, Payton e Santorico (2011), dados foram gerados pela distribuição normal multivariada, sob a hipótese alternativa, considerando as mesmas matrizes de covariâncias Σ_H e Σ_{AP} para obter o poder do teste MCPT. No primeiro cenário de simulação, o MCPT apresentou excelente desempenho com poder de 0,983 quando $n = 50$, porém apresentou baixo poder ($\leq 0,2$) para pequenas amostras. No segundo cenário, o teste MCPT foi menos poderoso na detecção da estrutura da matriz Σ_{AP} em comparação ao primeiro cenário, com a matriz Σ_H .

De acordo com os resultados na Figura 4.7, o poder dos testes LRTO e LRTR foram praticamente idênticos em todas as distribuições consideradas para todos os n . Os testes BLRTR e BLRTO também obtiveram os valores de poder muito parecidos em todas as distribuições consideradas para todos os n . Pode-se perceber também que os testes assintóticos obtiveram maior poder em comparação aos testes computacionalmente intensivos.

Figura 4.7 – Primeiro cenário (Σ_H): Simulação do poder para os testes de simetria composta ($p = 3$; $d = 4$; $\rho = 0$) respectivamente para as distribuições normal multivariada contaminada ($\delta = 0,30$), normal multivariada contaminada ($\delta = 0,10$), normal multivariada contaminada ($\delta = 0,05$) e normal multivariada, considerando $\alpha = 0,05$



Fonte: Da autora (2020).

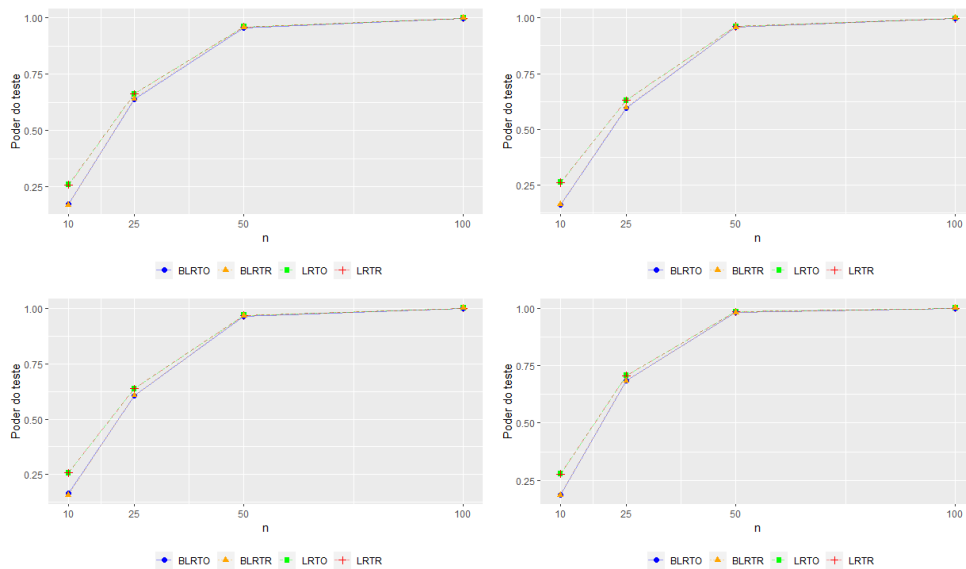
Ainda sobre os resultados na Figura 4.7, verifica-se que o poder de cada um dos testes aumentam à medida que os dados são gerados pelas distribuições com menos contaminação até chegar na normal multivariada. Os testes assintóticos foram mais poderosos dos que os testes computacionalmente intensivos para pequenos tamanhos amostrais. Nos dados contaminados, os poderes dos testes assintóticos se igualam aos poderes dos testes computacionalmente intensivos, quando $n = 50$. Porém, quando os dados são gerados sem contaminação, os poderes dos testes assintóticos se igualam mais rapidamente aos poderes dos testes computacionalmente intensivos, com $n = 25$.

Na Figura 4.8, os testes assintóticos e computacionalmente intensivos apresentaram alto poder para $n = 50$ e 100 em todas as distribuições consideradas, cujos valores são muito parecidos. Para $n = 10$ e 25 todos os testes apresentaram baixo poder, no entanto os testes LRTR e LRTO obtiveram maior poder em comparação aos testes BLRTR e BLRTO.

Ao comparar os resultados do poder dos testes LRTR, LRTO, BLRTR e BLRTO com o poder do teste MCPT, no primeiro cenário com dados gerados pela distribuição normal multivariada. Verifica-se que todos os testes assintóticos e computacionalmente intensivos obtiveram poder igual a 0,999 para $n = 50$, superando o teste MCPT de Morris, Payton e Santorico (2011), que obteve poder 0,983 quando $n = 50$. No segundo cenário, com dados gerados pela distribui-

ção normal multivariada, os testes assintóticos e computacionalmente intensivos obtiveram alto poder (0,983) a partir de $n \geq 50$, enquanto o teste MCPT tinha 0,65 de poder para $n = 50$.

Figura 4.8 – Segundo cenário (Σ_{AP}): Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) respectivamente para as distribuições normal multivariada contaminada ($\delta = 0,30$), normal multivariada contaminada ($\delta = 0,10$), normal multivariada contaminada ($\delta = 0,05$) e normal multivariada, considerando $\alpha = 0,05$



Fonte: Da autora (2020).

Diante dos resultados da taxa de erro tipo I nas Figuras 4.2, 4.4 e 4.6, com $\alpha = 0,05$, os testes assintóticos foram liberais para $n = 25$ e $p = 20$, já os testes computacionalmente intensivos foram exatos ou aproximadamente exatos. Por isso, optou-se por realizar simulações adicionais em relação ao poder de cada teste considerando $n = 25$ e $p = 20$, com as matrizes de covariâncias, sob a hipótese alternativa, já consideradas anteriormente (Σ_H e Σ_{AP}).

Portanto, os resultados na Tabela 4.1, são referentes ao poder de cada teste considerando a matriz de covariâncias Σ_{AP} , do segundo cenário de simulação, com $\alpha = 0,05$, $n = 25$, $p = 20$ e as distribuições normal multivariada e normal multivariada contaminada ($\delta = 0,10$ e $\delta = 0,30$). Os testes assintóticos apresentaram poder igual a 1,000, porém, suas taxas de erro tipo I, apresentadas respectivamente nas Figuras 4.2, 4.4 e 4.6, foram liberais. Desta forma, de nada adianta que os testes assintóticos tenham alto poder, se são liberais em relação a taxa de erro tipo I.

Ainda sobre os resultados na Tabela 4.1, os testes computacionalmente intensivos obtiveram alto poderes, atingindo 0,995 ao considerar a distribuição normal multivariada contaminada ($\delta = 0,30$). Os mesmos testes controlaram a taxa de erro tipo I, ao considerar a mesma

configuração de simulação, cujos resultados encontram-se nas Figuras 4.2, 4.4 e 4.6. Portanto, os testes BLRTR e BLRTO além de controlarem a taxa de erro tipo I, possuem alto poder ($> 0,991$).

Tabela 4.1 – Segundo cenário (Σ_{AP}): Simulação do poder para os testes de simetria composta ($p = 20$; $n = 25$; $\sigma^2 = 9$ e $\rho = 0,6$) para as distribuições normal multivariada, normal multivariada contaminada ($\delta = 0,10$) e normal multivariada contaminada ($\delta = 0,30$), considerando $\alpha = 0,05$

Grau de contaminação (δ)	LRTO	LRTR	BLRTO	BLRTR
$\delta = 0$	1,000	1,000	0,992	0,993
$\delta = 0,10$	1,000	1,000	0,991	0,991
$\delta = 0,30$	1,000	1,000	0,995	0,995

Fonte: Da autora (2020).

Os resultados na Tabela 4.2, também são referentes as simulações adicionais do poder de cada teste, considerando a matriz de covariâncias Σ_H , do primeiro cenário de simulação, com $\alpha = 0,05$, $p = 20$, $n = 25$ e distribuições normal multivariada e normal multivariada contaminada ($\delta = 0,10$ e $\delta = 0,30$). Os testes assintóticos obtiveram alto poder, alcançando 0,854 ao considerar a distribuição normal multivariada contaminada com $\delta = 0,30$. No entanto, é importante ressaltar novamente que os mesmos testes foram liberais em relação a taxa de erro tipo I, ao considerar $p = 20$ e $n = 25$ como mostra os resultados nas Figuras 4.2, 4.4 e 4.6.

Já os testes computacionalmente intensivos foram pouco poderosos, em que 0,399 e 0,401 foram os maiores valores obtidos, respectivamente para BLRTO e BLRTR, com dados contaminados 30%. A matriz de covariâncias Σ_H , definida na expressão (3.7), com $d = 4$ e $p = 20$ possui dimensões (20x20) e é escrita da seguinte forma:

$$\Sigma_H = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1,210526 & 0 & \dots & 0 \\ 0 & 0 & 1,421053 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 5 \end{bmatrix},$$

ao considerar $d = 4$, a estrutura da matriz Σ_H é semelhante a estrutura da matriz de covariâncias de simetria composta, em que todas as correlações entre pares de diferentes variáveis são iguais e as variâncias são homogêneas. Por isso, optou-se por realizar simulações do poder de cada

testes, considerando $d = 10$ e $d = 20$, com intuito de diferenciar mais a estrutura da matriz Σ_H com a de simetria composta, cujos resultados são apresentados na Tabela 4.3.

Tabela 4.2 – Primeiro cenário (Σ_H): Simulação do poder para os testes de simetria composta ($p = 20$; $n = 25$; $d = 4$ e $\rho = 0$) respectivamente para as distribuições normal multivariada, normal multivariada contaminada ($\delta = 0,10$) e normal multivariada contaminada ($\delta = 0,30$), considerando $\alpha = 0,05$

Grau de contaminação (δ)	LRTO	LRTR	BLRTO	BLRTR
$\delta = 0$	0,832	0,832	0,351	0,352
$\delta = 0,10$	0,763	0,763	0,271	0,277
$\delta = 0,30$	0,854	0,854	0,399	0,401

Fonte: Da autora (2020).

De acordo com os resultados na Tabela 4.3, os quatro testes obtiveram maiores poderes ao considerar $d = 10$ e $d = 20$ em comparação quando se considerou $d = 4$ na Tabela 4.2. Os testes assintóticos apresentaram elevado poder (0,999) para $d = 20$ e os testes computacionalmente intensivos também apresentaram elevado poder (0,922) para $d = 20$. Portanto, a medida que aumenta o valor de d , a matriz Σ_H difere mais sua estrutura em comparação a estrutura de simetria composta, conseqüentemente os testes aumentam seus poderes. Em todos os resultados apresentados nas Tabelas 4.3, 4.2 e 4.1, o poder dos testes LRTO e LRTR foram praticamente idênticos, e os testes BLRTR e BLRTO também obtiveram os valores de poder muito parecidos em todas as configurações de simulações.

Tabela 4.3 – Primeiro cenário (Σ_H): Simulação do poder para os testes de simetria composta ($p = 20$; $n = 25$; $d = 10$ e 20 e $\rho = 0$) para a distribuição normal multivariada, considerando $\alpha = 0,05$

d	LRTO	LRTR	BLRTO	BLRTR
10	0,976	0,976	0,722	0,722
20	0,999	0,999	0,922	0,922

Fonte: Da autora (2020).

5 CONCLUSÕES

Diante de uma análise sobre o desempenho dos testes assintóticos, LRTO e LRTR, e dos testes computacionalmente intensivos, BLRTO e BLRTR, conclui-se que os testes que abordaram estimadores robustos Comedian foram equivalentes aos testes que abordaram estimadores clássicos. Portanto, os estimadores robustos não agregaram melhorias aos testes assintóticos nem aos testes computacionalmente intensivos, quando os dados possuíam desvios de normalidade, causados por *outliers*, ou até mesmo quando eram normalmente distribuídos.

Os testes computacionalmente intensivos obtiveram melhores desempenhos em relação aos testes assintóticos. Pois, ainda que os testes LRTO e LRTR, obtiveram maiores poderes em comparação aos testes BLRTO e BLRTR, os testes BLRTO e BLRTR foram exatos ou praticamente exatos em relação a taxa de erro tipo I, em todas as configurações de simulações. Já os testes LRTO e LRTR, foram liberais ou expressivamente liberais, quando o número de variáveis (p) se aproximou do tamanho amostral (n), mesmo no cenário ideal de dados normalmente distribuídos.

Os testes computacionalmente intensivos também obtiveram melhores desempenhos em comparação ao teste robusto MCPT do estudo de Morris, Payton e Santorico (2011), dado que o teste MCPT foi liberal para pequenos tamanhos amostrais, e os testes BLRTO e BLRTR foram exatos neste mesmo cenário. Além de que os testes BLRTO e BLRTR obtiveram maiores poderes do que o teste MCPT.

Conclui-se que, ao considerar desvios de normalidade dos dados causados por *outliers* ou dados normalmente distribuídos, os testes que utilizaram o método *bootstrap* paramétrico foram os que tiveram melhores desempenhos, principalmente quando o número de variáveis (p) aproxima-se do tamanho amostral (n). Portanto, os testes BLRTR e BLRTO superaram o teste original da razão de verossimilhanças, LRTO, em dois aspectos: controlam a taxa de erro tipo I quando há presença de *outliers* nos dados e não utilizam distribuição assintótica para estatística do teste.

REFERÊNCIAS BIBLIOGRÁFICAS

- BOX, G. E. A general distribution theory for a class of likelihood criteria. **Biometrika**, Oxford, v. 36, n. 3/4, p. 317–346, dez. 1949. Disponível em: <<http://dx.doi.org/10.2307/2332671>>. Acesso em: mai. 2018.
- CASELLA, G.; BERGER, R. L. **Inferência estatística-tradução da 2ª edição norte-americana**. São Paulo: Centage Learning, 2014.
- COELHO, C. A.; MARQUES, F. J.; OLIVEIRA, S. Near-exact distributions for likelihood ratio statistics used in the simultaneous test of conditions on mean vectors and patterns of covariance matrices. **Mathematical Problems in Engineering**, London, v. 2016, p. 1–25, 2016. Disponível em: <<http://dx.doi.org/10.1155/2016/8975902>>. Acesso em: jun. 2018.
- DAVISON, A. C.; HINKLEY, D. V. **Bootstrap methods and their application**. Cambridge: Cambridge University Press, 1997. v. 1.
- EFRON, B. Bootstrap methods: another look at the jackknife. In: **Breakthroughs in statistics**. New York: Springer, 1992. v. 2, p. 569–593.
- FALK, M. On mad and comedians. **Annals of the Institute of Statistical Mathematics**, New York, v. 49, n. 4, p. 615–644, dez. 1997. Disponível em: <<http://dx.doi.org/10.1023/a:1003258024248>>. Acesso em: nov. 2018.
- FERREIRA, D. F. **Estatística básica**. 2. ed. Lavras: Editora UFLA, 2009.
- FERREIRA, D. F. **Estatística computacional em Java**. 1. ed. Lavras: Editora UFLA, 2013.
- FERREIRA, D. F. **Estatística multivariada**. 3. ed. Lavras: Editora UFLA, 2018.
- HAMPEL, F. R. A general qualitative definition of robustness. **The Annals of Mathematical Statistics**, New York, v. 42, n. 6, p. 1887–1896, 1971. Disponível em: <<https://www.jstor.org/stable/2240114>>. Acesso em: jul. 2018.
- HAMPEL, F. R. et al. **Robust statistics: the approach based on influence functions**. New York: John Wiley & Sons, 2011. v. 196.
- HAWKINS, D. M. **Identification of outliers**. New York: Springer, 1980. v. 11.
- HUYNH, H.; FELDT, L. S. Conditions under which mean square ratios in repeated measurements designs have exact f-distributions. **Journal of the American Statistical Association**, Alexandria, v. 65, n. 332, p. 1582–1589, dez. 1970. Disponível em: <<http://dx.doi.org/10.1080/01621459.1970.10481187>>. Acesso em: jul. 2019.
- LEE, J. C.; KRIDHNAIAH, P.; CHANG, T. On the distribution of the likelihood ratio test statistic for compound symmetry. **South African Statistical Journal**, Centurion, v. 10, n. 1, p. 49–62, jan. 1976.
- MAECHLER, M. et al. **robustbase: Basic Robust Statistics**. [S.l.], 2019. R package version 0.93-5. Disponível em: <<http://robustbase.r-forge.r-project.org/>>.
- MARONNA, R. A.; ZAMAR, R. H. Robust estimates of location and dispersion for high-dimensional datasets. **Technometrics**, Alexandria, v. 44, n. 4, p. 307–317, nov. 2002. Disponível em: <<http://dx.doi.org/10.1198/004017002188618509>>. Acesso em: jun. 2018.

MARQUES, F. J.; COELHO, C. A. The sphericity versus equivariance-euicorrelation test. In: AIP PUBLISHING. **AIP Conference Proceedings**. New York, 2015. v. 1648, n. 1, p. 540009.

MAUCHLY, J. W. Significance test for sphericity of a normal n-variate distribution. **The Annals of Mathematical Statistics**, New York, v. 11, n. 2, p. 204–209, jun. 1940. Disponível em: <<https://www.jstor.org/stable/2235878>>. Acesso em: jun. 2018.

MITTELHAMMER, R. C. **Mathematical statistics for economics and business**. 2. ed. New York: Springer, 2013.

MORRIS, T. L.; PAYTON, M. E.; SANTORICO, S. A. A permutation test for compound symmetry with application to gene expression data. **Journal of Modern Applied Statistical Methods**, Detroit, v. 10, n. 2, p. 447–461, nov. 2011. Disponível em: <<http://dx.doi.org/10.22237/jmasm/1320120300>>. Acesso em: abr. 2019.

OLIVEIRA, I. Cardoso de; FERREIRA, D. Multivariate extension of chi-squared univariate normality test. **Journal of Statistical Computation and Simulation**, Abingdon, v. 80, n. 5, p. 513–526, mai. 2010. Disponível em: <<http://dx.doi.org/10.1080/00949650902731377>>. Acesso em: nov. 2018.

PEREIRA, É. A.; FERREIRA, D. F. **Proposição de testes de normalidade multivariada baseados em distâncias robustas**. 2017. 105 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) — Universidade Federal de Lavras, Lavras, MG, 2017.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2018. Disponível em: <<https://www.R-project.org/>>.

ROCKE, D. M.; WOODRUFF, D. L. Identification of outliers in multivariate data. **Journal of the American Statistical Association**, Abingdon, v. 91, n. 435, p. 1047–1061, set. 1996. Disponível em: <<http://dx.doi.org/10.1080/01621459.1996.10476975>>. Acesso em: jul. 2018.

SAJESH, T.; SRINIVASAN, M. Outlier detection for high dimensional data using the comedian approach. **Journal of Statistical Computation and Simulation**, Abingdon, v. 82, n. 5, p. 745–757, mai. 2012. Disponível em: <<http://dx.doi.org/10.1080/00949655.2011.552504>>. Acesso em: jul. 2018.

UEDA, C. M. et al. **Modelos não lineares com diferentes estruturas de covariância em curvas de crescimento: uma aplicação no estudo da severidade da doença Late blight**. 2003. 111 p. Dissertação (Mestrado em Engenharia de Produção) — Universidade Federal de Santa Catarina, Florianópolis, SC, 2003.

VOTAW Jr, D. F. Testing compound symmetry in a normal multivariate distribution. **The Annals of Mathematical Statistics**, New York, v. 19, n. 4, p. 447–473, 1948. Disponível em: <<https://www.jstor.org/stable/2236016>>. Acesso em: mai. 2018.

WILKS, S. S. et al. Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution. **The Annals of Mathematical Statistics**, New York, v. 17, n. 3, p. 257–281, 1946. Disponível em: <<https://www.jstor.org/stable/2236125>>. Acesso em: mai. 2018.

A APÊNDICE - RESULTADOS DA SIMULAÇÃO

Tabela A.1 – Primeiro cenário: Simulação do erro tipo I para o teste de simetria composta ($p = 3$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\delta = 0,05$), considerando respectivamente $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$

n	α	LRTO	LRTR	BLRTO	BLRTR
5	0,10	0,215	0,215	0,107	0,109
	0,05	0,144	0,144	0,043	0,041
	0,01	0,039	0,039	0,010	0,012
10	0,10	0,125	0,125	0,092	0,090
	0,05	0,063	0,063	0,041	0,042
	0,01	0,015	0,015	0,010	0,009
15	0,10	0,102	0,102	0,082	0,082
	0,05	0,047	0,047	0,041	0,042
	0,01	0,021	0,021	0,013	0,013
25	0,10	0,085	0,085	0,078	0,078
	0,05	0,049	0,049	0,041	0,046
	0,01	0,006	0,006	0,007	0,004
50	0,10	0,090	0,090	0,089	0,086
	0,05	0,044	0,044	0,041	0,042
	0,01	0,008	0,008	0,008	0,006
100	0,10	0,083	0,083	0,083	0,081
	0,05	0,044	0,044	0,045	0,046
	0,01	0,011	0,011	0,011	0,010

Fonte: Da autora (2020).

Tabela A.2 – Segundo cenário: Simulação do erro tipo I para o teste de simetria composta ($n = 25$; $\sigma^2 = 9$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\delta = 0,05$), considerando respectivamente $\alpha = 0,10$, $\alpha = 0,05$ e $\alpha = 0,01$

p	α	LRTO	LRTR	BLRTO	BLRTR
3	0,10	0,094	0,094	0,085	0,088
	0,05	0,052	0,052	0,042	0,043
	0,01	0,007	0,007	0,006	0,006
5	0,10	0,102	0,102	0,089	0,089
	0,05	0,048	0,048	0,040	0,038
	0,01	0,010	0,010	0,006	0,006
10	0,10	0,119	0,119	0,084	0,087
	0,05	0,062	0,062	0,045	0,044
	0,01	0,011	0,011	0,009	0,008
20	0,10	0,483	0,483	0,093	0,101
	0,05	0,348	0,348	0,036	0,040
	0,01	0,168	0,168	0,004	0,005

Fonte: Da autora (2020).

Tabela A.3 – Primeiro cenário: Simulação do poder para os testes de simetria composta ($p = 3$; $d = 4$; $\rho = 0$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_H$ e $\boldsymbol{\Sigma}_2 = 2\boldsymbol{\Sigma}_H$), considerando $\alpha = 0,10$, $\alpha = 0,01$ e $\delta = 0,30$

n	α	LRTO	LRTR	BLRTO	BLRTR
5	0,10	0,248	0,248	0,118	0,120
	0,01	0,059	0,059	0,012	0,014
10	0,10	0,270	0,270	0,216	0,217
	0,01	0,047	0,047	0,031	0,034
25	0,10	0,464	0,464	0,442	0,441
	0,01	0,146	0,146	0,128	0,129
50	0,10	0,766	0,766	0,757	0,760
	0,01	0,400	0,400	0,387	0,393

Fonte: Da autora (2020).

Tabela A.4 – Primeiro cenário: Simulação do poder para os testes de simetria composta ($p = 3$; $d = 4$; $\rho = 0$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_H$ e $\boldsymbol{\Sigma}_2 = 2\boldsymbol{\Sigma}_H$), considerando $\alpha = 0,10$, $\alpha = 0,01$ e $\delta = 0,10$

n	α	LRTO	LRTR	BLRTO	BLRTR
5	0,10	0,282	0,282	0,130	0,130
	0,01	0,069	0,069	0,018	0,019
10	0,10	0,339	0,339	0,292	0,296
	0,01	0,091	0,091	0,066	0,068
25	0,10	0,542	0,542	0,518	0,523
	0,01	0,190	0,190	0,174	0,173
50	0,10	0,835	0,835	0,827	0,827
	0,01	0,519	0,519	0,502	0,510

Fonte: Da autora (2020).

Tabela A.5 – Primeiro cenário: Simulação do poder para os testes de simetria composta ($p = 3$; $d = 4$; $\rho = 0$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_H$ e $\boldsymbol{\Sigma}_2 = 2\boldsymbol{\Sigma}_H$), considerando $\alpha = 0,10$, $\alpha = 0,01$ e $\delta = 0,05$

n	α	LRTO	LRTR	BLRTO	BLRTR
5	0,10	0,317	0,317	0,164	0,165
	0,01	0,080	0,080	0,018	0,016
10	0,10	0,404	0,404	0,338	0,340
	0,01	0,116	0,116	0,075	0,069
25	0,10	0,636	0,636	0,618	0,620
	0,01	0,324	0,324	0,310	0,309
50	0,10	0,908	0,908	0,905	0,902
	0,01	0,644	0,644	0,632	0,631

Fonte: Da autora (2020).

Tabela A.6 – Primeiro cenário: Simulação do poder para os testes de simetria composta ($p = 3$; $d = 4$; $\rho = 0$) com a distribuição normal multivariada ($\boldsymbol{\mu} = [0, 0, \dots, 0]^\top$, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_H$), considerando $\alpha = 0,10$ e $\alpha = 0,01$

n	α	LRTO	LRTR	BLRTO	BLRTR
5	0,10	0,391	0,391	0,195	0,197
	0,01	0,094	0,094	0,022	0,019
10	0,10	0,542	0,542	0,474	0,473
	0,01	0,184	0,184	0,133	0,136
25	0,10	0,954	0,954	0,946	0,948
	0,01	0,737	0,737	0,703	0,704
50	0,10	1,000	1,000	1,000	1,000
	0,01	0,985	0,985	0,985	0,984

Fonte: Da autora (2020).

Tabela A.7 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = 2\boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$, $\alpha = 0,01$ e $\delta = 0,30$

n	α	LRTO	LRTR	BLRTO	BLRTR
10	0,10	0,378	0,378	0,276	0,280
	0,01	0,101	0,101	0,058	0,060
25	0,10	0,776	0,776	0,774	0,752
	0,01	0,433	0,433	0,399	0,409
50	0,10	0,976	0,976	0,975	0,975
	0,01	0,889	0,889	0,881	0,880
75	0,10	1,000	1,000	1,000	1,000
	0,01	0,991	0,991	0,990	0,987

Fonte: Da autora (2020).

Tabela A.8 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = 2\boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$, $\alpha = 0,01$ e $\delta = 0,10$

n	α	LRTO	LRTR	BLRTO	BLRTR
10	0,10	0,369	0,369	0,272	0,279
	0,01	0,083	0,083	0,034	0,041
25	0,10	0,735	0,735	0,714	0,715
	0,01	0,378	0,378	0,345	0,346
50	0,10	0,976	0,976	0,975	0,975
	0,01	0,883	0,883	0,875	0,872
75	0,10	1,000	1,000	1,000	1,000
	0,01	0,994	0,994	0,991	0,981

Fonte: Da autora (2020).

Tabela A.9 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = 2\boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$, $\alpha = 0,01$ e $\delta = 0,05$

n	α	LRTO	LRTR	BLRTO	BLRTR
10	0,10	0,364	0,364	0,270	0,269
	0,01	0,106	0,106	0,054	0,055
25	0,10	0,756	0,756	0,738	0,742
	0,01	0,409	0,409	0,381	0,384
50	0,10	0,985	0,985	0,984	0,982
	0,01	0,893	0,893	0,887	0,886
75	0,10	1,000	1,000	1,000	1,000
	0,01	0,993	0,993	0,993	0,992

Fonte: Da autora (2020).

Tabela A.10 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada ($\boldsymbol{\mu} = [0, 0, \dots, 0]^\top$ e $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$ e $\alpha = 0,01$

n	α	LRTO	LRTR	BLRTO	BLRTR
10	0,10	0,412	0,412	0,300	0,300
	0,01	0,108	0,108	0,061	0,060
25	0,10	0,816	0,816	0,791	0,789
	0,01	0,480	0,480	0,452	0,450
50	0,10	0,995	0,995	0,995	0,994
	0,01	0,932	0,932	0,929	0,924
75	0,10	1,000	1,000	1,000	1,000
	0,01	0,994	0,994	0,994	0,995

Fonte: Da autora (2020).

Tabela A.11 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$, $0,01$, $0,05$ e $\delta = 0,30$

n	α	LRTO	LRTR	BLRTO	BLRTR
10	0,10	0,392	0,392	0,291	0,290
	0,05	0,270	0,270	0,172	0,171
	0,01	0,110	0,110	0,057	0,054
25	0,10	0,795	0,795	0,773	0,773
	0,05	0,695	0,695	0,664	0,658
	0,01	0,440	0,440	0,405	0,398
50	0,10	0,990	0,990	0,989	0,990
	0,05	0,971	0,971	0,965	0,969
	0,01	0,898	0,898	0,884	0,886
75	0,10	0,999	0,999	0,999	0,999
	0,05	0,999	0,999	0,998	0,999
	0,01	0,991	0,991	0,990	0,990

Fonte: Da autora (2020).

Tabela A.12 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$, $0,01$, $0,05$ e $\delta = 0,10$

n	α	LRTO	LRTR	BLRTO	BLRTR
10	0,10	0,395	0,395	0,299	0,299
	0,05	0,279	0,279	0,182	0,183
	0,01	0,112	0,112	0,054	0,054
25	0,10	0,802	0,802	0,781	0,786
	0,05	0,693	0,693	0,669	0,663
	0,01	0,475	0,475	0,432	0,334
50	0,10	0,980	0,980	0,979	0,978
	0,05	0,960	0,960	0,959	0,957
	0,01	0,894	0,894	0,891	0,886
75	0,10	1,000	1,000	1,000	1,000
	0,05	0,999	0,999	0,999	0,998
	0,01	0,996	0,996	0,996	0,996

Fonte: Da autora (2020).

Tabela A.13 – Segundo cenário: Simulação do poder para os testes de simetria composta ($p = 5$; $\sigma^2 = 1$; $\rho = 0,6$) com a distribuição normal multivariada contaminada ($\boldsymbol{\mu}_1 = [0, 0, \dots, 0]^\top$, $\boldsymbol{\mu}_2 = [10, 10, \dots, 10]^\top$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{AP}$ e $\boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_{AP}$), considerando $\alpha = 0,10$, $0,01$, $0,05$ e $\delta = 0,05$

n	α	LRTO	LRTR	BLRTO	BLRTR
10	0,10	0,395	0,395	0,285	0,285
	0,05	0,261	0,261	0,179	0,175
	0,01	0,112	0,112	0,059	0,056
25	0,10	0,804	0,804	0,783	0,776
	0,05	0,681	0,681	0,662	0,656
	0,01	0,426	0,426	0,397	0,393
50	0,10	0,991	0,991	0,991	0,989
	0,05	0,970	0,970	0,969	0,967
	0,01	0,894	0,894	0,889	0,885
75	0,10	0,998	0,998	0,998	0,998
	0,05	0,998	0,998	0,998	0,998
	0,01	0,989	0,989	0,989	0,989

Fonte: Da autora (2020).

B APÊNDICE - COMANDOS USADOS NO R

```

# Função para fazer o teste LRT de simetria composta com
# correção de Box (1949) e estimador clássico de Sigma
LRTCSANR <- function(X)
{
  n <- nrow(X)
  p <- ncol(X)
  S <- var(X)
  S2c <- mean(diag(S))
  r <- sum(S - diag(diag(S))) / (p * (p - 1) * S2c)
  chi2c <- p * log(S2c) + (p - 1) * log(1 - r) +
    log(1 + (p - 1) * r) - log(det(S))
  chi2c <- ((n - 1) - p * (p + 1)^2 * (2 * p - 3) / (6 * (p - 1) *
    (p^2 + p + 4))) * chi2c
  nu <- p * (p + 1) / 2 - 2
  p.val <- 1.0 - pchisq(chi2c, nu)
  return(list(chi2c = chi2c, p.val = p.val))
}

# Função para fazer o teste LRT de simetria composta com
# correção de Box (1949) e estimador robusto de Sigma
# Depende do pacote robustbase
library(robustbase)
LRTCSAR <- function(X)
{
  n <- nrow(X)
  p <- ncol(X)
  tryCatch(
    {
      S <- covComed(X, n.iter = 0, reweight = TRUE)$cov
    },
    error = function(e)

```

```

{
  e #print(paste("Error! Invalid covMed! Don't worry!"))
},
finally = {
  S <- var(X)
})
S2c <- mean(diag(S))
r <- sum(S - diag(diag(S))) / (p * (p - 1) * S2c)
chi2c <- p * log(S2c) + (p - 1) * log(1 - r) +
  log(1 + (p - 1) * r) - log(det(S))
chi2c <- ((n - 1) - p * (p + 1)^2 * (2 * p - 3) / (6 * (p - 1) *
  (p^2 + p + 4))) * chi2c
nu <- p * (p + 1) / 2 - 2
p.val <- 1.0 - pchisq(chi2c, nu)
return(list(chi2c = chi2c, p.val = p.val))
}

# Testes bootstrap robustos e não robusto: baseados no
# estimador robusto e não robusto de Sigma de simetria
# composta, impondo H_0. Assim LRTCSANR e LRTCSAR
BP.tests <- function(X, B = 2000)
{
  n <- nrow(X)
  p <- ncol(X)
  LRTO <- LRTCSANR(X)$chi2c
  LRTR <- LRTCSAR(X)$chi2c
  tryCatch(
    {
      robust <- covComed(X, n.iter = 0, reweight = TRUE)
      Sest <- robust$cov
      XestR <- robust$center
    },

```

```

error = function(e)
{
  e #print(paste("Error! Invalid covMed! Don't worry!"))
},
finally = {
  Sest <- var(X)
  XestR <- apply(X, 2, mean)
})
S2cR <- mean(diag(Sest))
rR <- sum(Sest - diag(diag(Sest))) / (p * (p - 1) * S2cR)
S <- var(X)
S2c <- mean(diag(S))
Xbar <- apply(X, 2, mean)
r <- sum(S - diag(diag(S))) / (p * (p - 1) * S2c)
LRTB <- LRTO
LRTRB <- LRTR
ct <- 1 / (B + 1)
pvalue.LRTO <- ct
pvalue.LRTR <- ct
SigR <- S2cR * ((1 - rR) * diag(p) + rR * matrix(1, p, p))
SigNR <- S2c * ((1 - r) * diag(p) + r * matrix(1, p, p))
for (i in 1:B)
{
  XbR <- mvrnorm(n, XestR, SigR)
  XbNR <- mvrnorm(n, Xbar, SigNR)
  S <- var(XbNR)
  S2c <- mean(diag(S))
  r <- sum(S - diag(diag(S))) / (p * (p - 1) * S2c)
  chi2c <- p * log(S2c) + (p - 1) * log(1 - r) +
    log(1 + (p - 1) * r) - log(det(S))
  chi2c <- ((n - 1) - p * (p + 1)^2 * (2 * p - 3) / (6 * (p - 1) *
    (p^2 + p + 4))) * chi2c

```

```

LRTB <- c(LRTB, chi2c)
tryCatch(
  {
    S <- covComed(XbR, n.iter = 0, reweight = TRUE)$cov
  },
  error = function(e)
  {
    e #print(paste("Error! Invalid covMed! Don't worry!"))
  },
  finally = {
    S <- var(XbR)
  })
S2c <- mean(diag(S))
r <- sum(S - diag(diag(S))) / (p * (p - 1) * S2c)
chi2cb <- p * log(S2c) + (p - 1) * log(1 - r) +
  log(1 + (p - 1) * r) - log(det(S))
chi2cb <- ((n - 1) - p * (p + 1)^2 * (2 * p - 3) / (6 * (p - 1) *
  (p^2 + p + 4))) * chi2cb
LRTRB <- c(LRTRB, chi2cb)
if (chi2c >= LRT0) pvalue.LRT0 <- pvalue.LRT0 + ct
if (chi2cb >= LRTR) pvalue.LRTR <- pvalue.LRTR + ct
}
return(list(LRT0 = LRT0, pvalue.LRT0 = pvalue.LRT0,
  LRTR = LRTR, pvalue.LRTR = pvalue.LRTR))
}

#similar amostras normal multivariada contaminada
library(MASS)
rNCM <- function(n, delta , mu, mu2, Sigma, Sigma2)
{
  p <- nrow(Sigma)
  u <- runif(n)

```

```

n1 <- length(u[u <= delta])
n2 <- n - n1
X <- matrix(0, n, p)
if (n1 > 0)
X[u <= delta, ] <- mvrnorm(n1, mu, Sigma)
if (n2 > 0) X[u > delta, ] <- mvrnorm(n2, mu2, Sigma2)
return(X)
}

#Função para criar uma matriz autorregressiva de primeira ordem
criaSigmaAR1 <- function(p, sig2, rho)
{
times <- 1:p
H <- abs(outer(times, times, "-"))
V <- sig2 * rho^H
return(V)
}

#Função para criar uma matriz heterogenea d = 4
criaSigmaHet <- function(p, d = 4)
{
times <- 1 + (0:(p - 1)) * d / (p - 1)
V <- diag(times)
return(V)
}

criaSigmaHet(20)
# Programa para realizar simulações Monte Carlo
# Simular as populações 1: Normal, 2: Normal Contaminada
# Sob H_0: simetria composta e
# Sob H_1: sem restrições (tem dois casos)
# optH1 = 1: Sob H_1 situação 1 - AR1

```

```

# optH1 = 2: Sob H_1 situação 2 - heterocedasticidade diag. d=4 default
# Normal Contaminada: delta = 0.9 (contaminação é 1 - delta)

library(MASS)

library(mvtnorm)

Sim_CSTest_Desemp <- function(dist = 1, h0 = TRUE, N = 1000, B = 2000,
                              n, p, rho = 0.2, sig2 = 10, delta = 0.9,
                              optH1 = 1, d = 4)

{
  rej <- matrix(0, 4, 3)
  rownames(rej) <- c("LRTCSANR", "LRTCSAR", "LRTCSBNR", "LRTCSBR")
  colnames(rej) <- c("0,10", "0,05", "0,01")
  N1R <- 1.0 / N
  mu <- rep(0, times = p)
  if (h0 == TRUE)
  {
    Sigma <- sig2 * ((1 - rho) * diag(1, p, p) + rho * matrix(1, p, p))
  } else {

    if (optH1 == 1) Sigma <- criaSigmaAR1(p, sig2, rho) else
      if (optH1 == 2) Sigma <- criaSigmaHet(p, d)
  }
  if (dist == 2) # Normal Contaminada - Pensar Parâmetros
  {
    mu2 <- rep(10, times = p)
    Sigma2 <- 2 * Sigma
  }
  #lixo <- NULL
  for (i in 1:N)
  {
    if (dist == 1) X <- mvrnorm(n, mu, Sigma) else
      if (dist == 2) X <- rNCM(n, delta, mu, mu2, Sigma, Sigma2) else
      res1 <- LRTCSANR(X)
  }
}

```

```
res2 <- LRTCSAR(X)
res3 <- BP.tests(X, B)
if (res1$p.val <= 0.05) rej[1,2] <- rej[1,2] + N1R
if (res2$p.val <= 0.05) rej[2,2] <- rej[2,2] + N1R
if (res3$pvalue.LRTO <= 0.05) rej[3,2] <- rej[3,2] + N1R
if (res3$pvalue.LRTR <= 0.05) rej[4,2] <- rej[4,2] + N1R
}
return(rej)
}

# Configurar simulações
dist <- 1
h0 <- FALSE # TRUE or FALSE
B <- 2000
N <- 1000
n <- 25
p <- 20
rho <- 0
delta <- 0 # Caso da normal contaminada
sig2 <- 1
optH1 <- 2
d <- 10
Result <- Sim_CSTest_Desemp(dist, h0, N, B, n, p, rho, sig2, delta, optH1, d)
Result
```