



JOANA FILIPA MAGALHÃES DIOGO

Licenciada em Matemática

**TESTE DE INDEPENDÊNCIA DE VARIÁVEIS
PARA ESTRUTURAS
DE COVARIÂNCIA ESPECÍFICAS: UM
ESTUDO DE SIMULAÇÃO**

MESTRADO EM MATEMÁTICA E APLICAÇÕES

Universidade NOVA de Lisboa
setembro, 2022



TESTE DE INDEPENDÊNCIA DE VARIÁVEIS PARA ESTRUTURAS DE COVARIÂNCIA ESPECÍFICAS: UM ESTUDO DE SIMULAÇÃO

JOANA FILIPA MAGALHÃES DIOGO

Licenciada em Matemática

Orientador: Filipe José Gonçalves Pereira Marques
Professor Associado, NOVA School of Science & Technology

Júri:

- Presidente:** Doutora Marta Cristina Vieira Faias Mateus
Professora Associada, NOVA School of Science & Technology
- Arguente:** Doutor Pedro José dos Santos Palhinhas Mota
Professor Associado, NOVA School of Science & Technology
- Orientador:** Doutor Filipe José Gonçalves Pereira Marques
Professor Associado, NOVA School of Science & Technology

Teste de independência de variáveis para estruturas de covariância específicas: um estudo de simulação

Copyright © Joana Filipa Magalhães Diogo, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade NOVA de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

AGRADECIMENTOS

Para a realização desta dissertação foi fundamental o apoio e o incentivo de diversas pessoas, às quais me encontro eternamente grata.

Ao meu orientador, o Professor Doutor Filipe Marques, agradeço toda a sua disponibilidade, apoio e todo o conhecimento partilhado.

À minha família, em particular aos meus pais, agradeço todas as oportunidades que me proporcionaram, por me terem ensinado o caminho certo ao mesmo tempo que sempre me permitiram fazer as minhas próprias escolhas assim como por toda a aprendizagem e sabedoria que comigo sempre partilharam.

Quero ainda deixar um agradecimento especial ao meu namorado, Márcio, que me apoiou não só durante a realização desta dissertação, como desde o início da minha licenciatura. Esteve sempre presente, fez questão de me incentivar e ainda me transmitiu toda a sua calma em todas as fases mais complicadas de todo este ciclo.

Por último quero agradecer a todos os meus amigos, cujo apoio foi também incalculável, e cujo contributo maior foram todos os momentos de conversas, de partilha de experiências e até mesmo de alguma descontração que permitiu levar com maior simplicidade todo este processo.

*“You must do the thing you think you cannot do.” (Eleanor
Roosevelt)*

RESUMO

A independência de variáveis deve ser verificada e analisada quando se realizam estudos estatísticos, é por isso importante ter metodologias que permitam inferir sobre este assunto. Muitas vezes os métodos usados para testar a independência acabam por ser difíceis de aplicar, ou mesmo impossíveis para algumas situações, como é o caso do teste de razão de verossimilhanças cuja distribuição da estatística de teste é complexa de manusear e que tem a limitação de não ser possível de aplicar em cenários de alta dimensionalidade.

Assim, neste trabalho, é proposto um teste de hipóteses, que pode ser utilizado para testar a nulidade das covariâncias entre variáveis aleatórias e que, portanto, permite inferir sobre a independência de variáveis aleatórias de um vetor com distribuição Normal Multivariada quando nos encontramos sob determinadas condições específicas. Este teste tem a vantagem de reduzir um problema Multivariado a um problema Univariado, além de utilizar uma estatística com distribuição χ^2 de fácil utilização.

Para comprovar a eficácia do teste proposto foram realizadas simulações considerando diferentes cenários, utilizando como método de comparação os testes de razão de verossimilhanças, quando considerados casos onde é possível aplicá-lo, e também o teste de Schoot proposto em Schoot (2005) [19].

Palavras-chave: Alta dimensionalidade, teste do qui-quadrado, testes de razão de verossimilhanças, matrizes bloco diagonais, matrizes circulares, matrizes de igualdade de variâncias e igualdade de covariâncias.

ABSTRACT

The independence of variables must be verified and analyzed when carrying out statistical studies, it is therefore important to have methodologies that allow inferences on this subject. Often the methods used to test independence turn out to be difficult to apply, or even impossible for some situations, as is the case of the likelihood ratio test whose distribution of the test statistic is complex to handle and which has the limitation of not being possible to apply in high-dimensional scenarios.

Thus, in this work, a hypothesis test is proposed, which can be used to test the nullity of covariances between random variables and which, therefore, allows inferring about the independence of random variables of a vector with Multivariate Normal distribution when we are under certain specific conditions. This test has the advantage of reducing a Multivariate problem to a Univariate problem, in addition to using an easy-to-use χ^2 distribution statistic.

To prove the effectiveness of the proposed test, simulations were carried out considering different scenarios, using the likelihood ratio tests as a comparison method, when considering cases where it is possible to apply it, and also the Schoot test proposed in Schoot (2005) [19].

Keywords: High dimensional setting, chi-square test, likelihood ratio tests, diagonal block matrices, circular matrices, equality of variance and equality of covariance matrices.

ÍNDICE

| | |
|---|-------------|
| Índice de Figuras | xvii |
| Índice de Tabelas | xix |
| 1 Introdução | 1 |
| 2 Conceitos Base | 5 |
| 2.1 Testes de Hipóteses para a variância populacional | 5 |
| 2.2 Distribuição Normal Multivariada | 7 |
| 2.2.1 Propriedades | 7 |
| 2.2.2 Estimação | 9 |
| 2.2.3 Distribuições de amostragem de \bar{X} | 10 |
| 2.3 Distribuição Wishart | 10 |
| 2.3.1 Propriedades | 11 |
| 2.3.2 Distribuição de amostragem de S | 11 |
| 2.4 Testes de razão de verossimilhanças | 12 |
| 3 Métodos | 13 |
| 3.1 Testes de razão de verossimilhanças para independência de grupos de variáveis | 13 |
| 3.1.1 Testes de razão de verossimilhanças para independência completa de variáveis | 14 |
| 3.2 Testes de razão de verossimilhanças para independência completa em estruturas específicas | 14 |
| 3.2.1 Teste de Esfericidade vs. Equivariância-Equicorrelação | 15 |
| 3.2.2 Teste de Esfericidade vs Circularidade | 16 |
| 3.3 Teste de Schout | 17 |
| 3.4 Redução Univariada | 18 |
| 3.4.1 Metodologia do Teste de Redução Univariada | 20 |

| | | |
|----------|--|-----------|
| 4 | Simulações | 23 |
| 4.1 | Estimativas das potências dos testes | 26 |
| 4.2 | Níveis de significância e potências dos testes | 30 |
| 4.3 | Desvios da Normalidade | 33 |
| 4.4 | Testes de Esfericidade vs Equivariância-Equicorrelação | 36 |
| 4.5 | Testes de Esfericidade vs Circularidade | 38 |
| 5 | Conclusões | 39 |
| | Bibliografia | 41 |
| | Anexos | |
| I | Anexo 1: Código Python utilizado nas simulações | 45 |
| I.1 | Código usado para a estrutura de covariância simétrica composta | 45 |
| I.2 | Código usado para a estrutura de covariância autorregressiva | 48 |
| I.3 | Código usado para a estrutura de covariância circular | 50 |
| I.4 | Código usado para as simulações usando populações t-Student | 55 |
| I.5 | Código usado para as simulações do teste de esfericidade vs equivariância- equicorrelação | 58 |
| I.6 | Código usado para as simulações do teste de esfericidade vs circularidade | 61 |

ÍNDICE DE FIGURAS

| | | |
|-----|--|----|
| 4.1 | Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verosimilhanças quando $p < n$ para a matriz de covariâncias de estrutura simétrica composta | 27 |
| 4.2 | Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verosimilhanças quando $p > n$ para a matriz de covariâncias de estrutura simétrica composta | 27 |
| 4.3 | Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verosimilhanças com $p < n$ para a matriz de covariâncias de estrutura autorregressiva | 28 |
| 4.4 | Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verosimilhanças com $p > n$ para a matriz de covariâncias de estrutura autorregressiva | 29 |
| 4.5 | Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verosimilhanças com $p < n$ para a matriz de covariâncias de estrutura circular | 29 |
| 4.6 | Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verosimilhanças com $p > n$ para a matriz de covariâncias de estrutura circular | 30 |
| 4.7 | Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verosimilhanças para amostras de populações t-Student, com $p < n$ e para a matriz de covariâncias de estrutura simétrica composta | 35 |

| | | |
|------|---|----|
| 4.8 | Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verossimilhanças para amostras de populações t-Student, com $p < n$ e para a matriz de covariâncias de estrutura simétrica composta | 36 |
| 4.9 | Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas para a matriz de covariâncias de estrutura de simétrica composta e do teste de razão de verossimilhanças dado por Marques e Coelho (2015) [13] | 37 |
| 4.10 | Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas para a matriz de covariâncias de estrutura de circular e do teste de razão de verossimilhanças dado por Olkin e Press (1969) [17] | 38 |

ÍNDICE DE TABELAS

| | | |
|-----|---|----|
| 4.1 | Combinções de p variáveis e amostras de tamanho n consideradas nas simulações | 25 |
| 4.2 | Resultados das estimativas do nível de significância do teste de redução univariada para $\alpha = 0.05$ e variâncias desconhecidas | 31 |
| 4.3 | Resultados das estimativas do nível de significância do teste de redução univariada para $\alpha = 0.01$ e variâncias desconhecidas | 31 |
| 4.4 | Resultados das estimativas do nível de significância do teste de Schoot (Schoot, 2005 [19]) para $\alpha = 0.05$, valores retirados da Tabela 1 de Schoot (2005) [19] | 32 |
| 4.5 | Resultados das estimativas do nível de significância do teste de razão de verossimilhanças para $\alpha = 0.05$, valores retirados da Tabela 2 de Schoot (2005) [19] | 32 |
| 4.6 | Potências estimadas do teste de redução univariada para as variâncias desconhecidas com base em amostras Normais Multivariadas de vetor de médias $\underline{\mu} = \underline{0}$ e matriz de covariâncias de estrutura simétrica composta (4.1) com $\sigma^2 = 1$ e $\rho = 0.1$ e para $\alpha = 0.05$ | 33 |
| 4.7 | Potências estimadas do teste de Schoot (Schoot, 2005 [19]) com base em amostras Normais Multivariadas de vetor de médias $\underline{\mu} = \underline{0}$ e matriz de covariâncias de estrutura simétrica composta (4.1) com $\sigma^2 = 1$ e $\rho = 0.1$ e para $\alpha = 0.05$, resultados retirados da Tabela 3 de Schoot (2005) [19] | 33 |
| 4.8 | Potências estimadas do teste de razão de verossimilhanças com base em amostras Normais Multivariadas de vetor de médias $\underline{\mu} = \underline{0}$ e matriz de covariâncias de estrutura simétrica composta (4.1) com $\sigma^2 = 1$ e $\rho = 0.1$ e para $\alpha = 0.05$, resultados retirados da Tabela 4 de Schoot (2005) [19] | 34 |

INTRODUÇÃO

O teste de independência de várias variáveis é um teste muito importante em análises estatísticas e na aplicação de diferentes técnicas. Como é referido em Mudholkar et al. (1982) [16], as análises de dados multivariados podem ser muito simplificadas se pudermos assumir que todas as variáveis ou alguns grupos de variáveis são independentes, dado que se as variáveis forem independentes podemos simplificar processos usando, por exemplo, metodologias univariadas. Por outro lado, a não verificação da independência pode levar a conclusões erradas ou resultados enganadores. Assim, é fundamental ter metodologias, fáceis de implementar e de usar, que nos permitam testar a independência de variáveis com rigor e simplicidade.

Quando se pretende testar a independência de um conjunto de variáveis de uma população Normal Multivariada é suficiente estudar se a matriz de covariâncias dessas mesmas variáveis é uma matriz diagonal, pois se as covariâncias entre as várias variáveis forem nulas significa que as mesmas são independentes.

Um dos métodos mais usados para testar a independência de variáveis quando está garantido o pressuposto da Normalidade é o teste de razão de verosimilhanças, que usa a estatística de teste de razão de verosimilhanças, Λ (Anderson, 2003 [1]). Em Coelho e Marques (2010) [7] e em Marques et al. (2011) [14] os autores mostram que a distribuição da estatística de razão de verosimilhanças pode ser representada como a distribuição do produto de variáveis aleatórias Beta independentes, no entanto, na prática, esta distribuição torna-se complicada de manusear (Marques et al., 2011 [14] e Marques & Coelho, 2020 [12]). Para ultrapassar este entrave, Wilks (1938) [22] propôs uma aproximação a uma distribuição χ^2 , Box (1949) [2] propõe aproximações baseadas em misturas de distribuições χ^2 , Butler et al. (1993) [3] propõem as aproximações Ponto de Sela (ou 'Saddle Point') e Coelho (2004) [6] e Coelho e Marques (2010) [7] obtiveram as distribuições quase-exatas, sendo que, todas estas propostas, por se tratarem de aproximações, podem por vezes não ser precisas o suficiente.

O tradicional teste de razão de verosimilhanças pode não ser o mais adequado, em alguns casos, como por exemplo, quando as amostras são pequenas ou o número de variáveis é elevado, uma vez que a aproximação à distribuição χ^2 usualmente considerada

para a distribuição da estatística de teste não consegue fornecer aproximações precisas nestes casos. Outro obstáculo que surge na utilização desta estatística é o facto da mesma não poder ser usada quando o tamanho da amostra é inferior ao número de variáveis.

Quando a utilização do teste de razão de verosimilhanças não é a mais adequada, existem outros métodos que podem ser usados para testar a independência das variáveis. Por exemplo, é possível usar uma estatística dada pela soma dos quadrados das correlações amostrais, dada em Schoot (2005) [19] onde o autor mostra que a estatística tem distribuição normal assintótica. Esta estatística de teste é usada como meio de comparação nas simulações realizadas neste trabalho e os resultados obtidos pelo autor (Schoot, 2005 [19]) são também usados para comparar com os obtidos para o teste proposto neste trabalho. Os autores Ledoit e Wolf (2002) [9] estudaram a propriedade da consistência e a distribuição limite de várias estatísticas quando o número de variáveis e o tamanho da amostra tendem para o infinito e a sua razão tende para um valor finito diferente de zero. Li e Liu (2016) [10] utilizam um teste de permutação cuja estatística de teste é baseada no máximo entre o maior valor fora da diagonal principal da matriz de correlações amostral e o maior valor próprio dessa mesma matriz, ou seja, o máximo entre a maior correlação e o maior valor próprio da matriz de correlação. Além disso, os autores deduzem as distribuições limite desta estatística para cada um dos casos, que são respetivamente, uma distribuição de valores extremos tipo I , e a distribuição da Lei de Tracy-Widom tipo I . Outro teste que pode ser utilizado para testar a independência das variáveis, sob a hipótese de Normalidade Multivariada, é o teste proposto por Srivastava (2005) [20], que consiste na utilização de uma estatística baseada na soma dos quadrados das correlações amostrais, quando assumida uma condição específica no traço da matriz de covariâncias, e cujo autor mostra que a distribuição da mesma tem distribuição assintótica normal à medida que o tamanho da amostra e o número de variáveis tendem para o infinito.

Recentemente, com o aumento da capacidade de recolha e armazenamento de dados, os estatísticos foram confrontados com bases de dados em que a dimensão da amostra é inferior ao número de variáveis, cenários usualmente designados de alta dimensionalidade. Esta nova realidade coloca novos desafios que não podem ser resolvidos usando as metodologias tradicionais, como por exemplo os testes de razão de verosimilhanças. Neste trabalho, vamos procurar dar resposta aos problemas mencionados anteriormente, desenvolvendo uma metodologia que permite testar a independência de variáveis de uma forma simples, rigorosa e que também pode ser utilizada em cenários de alta dimensionalidade. Para a realização do teste proposto é necessário considerar que as covariâncias são todas não positivas ou todas não negativas. Embora este pressuposto possa parecer demasiado limitativo, é possível verificar que na literatura moderna existem muitas metodologias que são desenvolvidas assumindo pressupostos para as matrizes de covariância que verificam o pressuposto exigido para este teste. Por exemplo, em Coelho (2017, 2022) [5, 4] o autor analisa o teste de igualdade de vetores de médias de várias populações assumindo que as matrizes de covariância são todas iguais e que na sua estrutura apresentam *i)* as variâncias todas iguais e *ii)* as covariâncias todas iguais. Neste tipo de estruturas

temos garantido que as covariâncias são todas não positivas ou não negativas, pelo que é um tipo de estrutura que verifica o pressuposto exigido na metodologia proposta neste trabalho. Adicionalmente, o teste proposto neste trabalho pode ser usado para testar os testes de razão de verosimilhanças entre outras estruturas, por exemplo, *i*) esfericidade versus circularidade *ii*) esfericidade versus igualdade de variâncias e de covariâncias, com a vantagem de ser mais fácil de utilizar. De facto quando se testa, por exemplo, a esfericidade versus igualdade de variâncias e de covariâncias, não rejeitar H_0 é não rejeitar a independência das variáveis aleatórias.

Para testar a fiabilidade e a robustez do teste aqui proposto, foram realizadas simulações considerando diferentes casos para o tamanho da amostra, para o número de variáveis e para a estrutura de covariâncias. Também se estudou o comportamento do teste quando as populações se desviam da normalidade, e ainda foram realizadas simulações considerando diferentes estruturas para a hipótese alternativa.

CONCEITOS BASE

Neste capítulo são apresentados alguns conceitos que são importantes para compreender a metodologia proposta na Secção 3 e no desenvolvimento e análise das simulações apresentadas na Secção 4.

2.1 Testes de Hipóteses para a variância populacional

Para inferir sobre os parâmetros de uma determinada população X é possível recorrer aos testes de hipóteses que consistem em realizar um teste sobre uma conjectura que se faça sobre esse mesmo parâmetro. Às conjecturas feitas damos o nome de hipótese, e neste método estatístico é sempre necessário termos duas hipóteses: a hipótese nula, H_0 , e a hipótese alternativa, H_1 .

Este método consiste na recolha de informação de uma amostra da população X para calcular a região de rejeição que dá as condições para as quais a hipótese nula, H_0 , é rejeitada e também a região de não rejeição que corresponde às condições para as quais não rejeitamos a hipótese nula, H_0 . Dadas estas duas regiões, que são complementares, é possível inferir sobre a rejeição ou não rejeição da hipótese nula. Está aqui implícito, por só existirem duas hipóteses, que ao não rejeitarmos a hipótese nula, H_0 , estamos a rejeitar a hipótese alternativa, H_1 , e ao rejeitarmos a hipótese nula, H_0 , estamos a não rejeitar a hipótese alternativa, H_1 .

Por se tratar de um método que tira conclusões a partir de uma amostra podemos cometer erros na decisão, sendo esses, o erro do tipo *I* e o erro do tipo *II*. O erro do tipo *I* é cometido quando a decisão é rejeitar a hipótese nula, H_0 , mas esta é verdadeira e o erro do tipo *II* é cometido quando decidimos não rejeitar a hipótese nula, H_0 , mas esta é falsa. A probabilidade de ocorrer o erro do tipo *I* chamamos de nível de significância, α , ou seja

$$\alpha = P[\text{Erro tipo I}] = P[\text{Rejeitar } H_0 \mid H_0 \text{ verdadeira}]$$

enquanto que, para

$$\beta = P[\text{Erro tipo II}] = P[\text{Não rejeitar } H_0 \mid H_0 \text{ falsa}],$$

a $Q = 1 - \beta$ chamamos de potência do teste.

Na realização de um teste de hipóteses o ideal seria diminuir estas probabilidades, α e β , em simultâneo, no entanto é impossível por estas serem inversamente proporcionais, ou seja, quando α aumenta β diminui e vice-versa.

O mais usual na realização destes testes é estabelecermos α , isto é, escolhermos o nível de significância para o qual pretendemos realizar o teste, e procurarmos maximizar a potência, Q , do mesmo.

Um caso particular dos testes de hipóteses, é o teste de hipóteses bilateral para a variância, cuja conjectura é feita relativamente à variância da distribuição de uma determinada população X .

Esta técnica permite testar se a variância de uma população é igual a um determinado valor, especificado na hipótese nula, H_0 , versus a hipótese alternativa, H_1 , da variância ser diferente desse mesmo valor, ou seja,

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{vs.} \quad H_1 : \sigma^2 \neq \sigma_0^2.$$

Para a realização deste teste assumimos que a população X tem distribuição Normal de parâmetro de média μ desconhecida e parâmetro de variância σ^2 também desconhecido.

Assim, para uma amostra aleatória de tamanho n , (X_1, \dots, X_n) , da população X , a estatística de teste usada neste caso é dada por

$$\frac{(n-1)S^2}{\sigma_0^2} \sim \chi_{n-1}^2$$

onde n é o tamanho da amostra, $S^2 = \frac{1}{n-1} \left(\sum_{i=1}^n (X_i - \bar{X})^2 \right)$ é o estimador centrado da variância, σ_0^2 é o valor de σ^2 sob H_0 , $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ é o estimador centrado da média da população e χ_{n-1}^2 representa a distribuição qui-quadrado com $n-1$ graus de liberdade.

Por se tratar de um teste de hipóteses bilateral, a região de rejeição é

$$]0, \chi_{n-1, 1-\frac{\alpha}{2}}^2 [\cup] \chi_{n-1, \frac{\alpha}{2}}^2, +\infty [$$

onde $\chi_{n-1, \alpha}^2$ denota o quantil de ordem $1-\alpha$.

O nível de significância é definido por

$$\begin{aligned} \alpha &= P[\text{Rejeitar } H_0 \mid H_0 \text{ é verdadeiro}] \\ &= P\left[\frac{(n-1)S^2}{\sigma_0^2} < \chi_{n-1, 1-\frac{\alpha}{2}}^2\right] + P\left[\frac{(n-1)S^2}{\sigma_0^2} > \chi_{n-1, \frac{\alpha}{2}}^2\right]. \end{aligned}$$

Assim, rejeitamos a hipótese H_0 para o nível de significância α quando

$$\frac{(n-1)S^2}{\sigma_0^2} < \chi_{n-1, 1-\frac{\alpha}{2}}^2 \quad \text{ou} \quad \frac{(n-1)S^2}{\sigma_0^2} > \chi_{n-1, \frac{\alpha}{2}}^2.$$

Por outro lado, a potência do teste é

$$\begin{aligned} Q &= P[\text{Rejeitar } H_0 \mid H_0 \text{ é falsa}] \\ &= 1 - P[\text{Não rejeitar } H_0 \mid H_1 \text{ é verdadeiro}] \\ &= 1 - P\left[\chi_{n-1, 1-\frac{\alpha}{2}}^2 \leq \frac{(n-1)S^2}{\sigma_0^2} \leq \chi_{n-1, \frac{\alpha}{2}}^2 \mid \sigma^2 < \sigma_0^2 \vee \sigma^2 > \sigma_0^2\right]. \end{aligned}$$

2.2 Distribuição Normal Multivariada

A distribuição Normal ou Gaussiana é a distribuição contínua mais utilizada por se ajustar nos mais diversos cenários.

Dado um vector $\underline{X} \in \mathbb{R}^p$, dizemos que a sua distribuição é normal multivariada se a sua função de densidade de probabilidade (f.d.p.) é dada por

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu})' \Sigma^{-1} (\underline{x}-\underline{\mu})},$$

podendo ser representado por $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$, sendo $\underline{\mu}$ o vetor de médias e Σ a matriz de covariâncias de \underline{X} .

Em seguida apresentamos as principais propriedades desta distribuição.

2.2.1 Propriedades

Suponhamos que $\underline{X}' = (X_1, \dots, X_p) \in \mathbb{R}^p$ é tal que $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$, ou seja, \underline{X} tem distribuição Normal Multivariada de vetor de médias $\underline{\mu}$ e matriz de covariâncias Σ . São válidas as seguintes propriedades:

1. Normalidade de combinações lineares
 - a) A combinação linear $\underline{a}'\underline{X} = a_1X_1 + \dots + a_pX_p$, com $\underline{a}' = (a_1, \dots, a_p)$ o vetor de constantes tal que $a_i \in \mathbb{R} \forall i = 1, 2, \dots, p$, tem distribuição Normal Univariada de parâmetro de média $\underline{a}'\underline{\mu}$ e parâmetro de variância $\underline{a}'\Sigma\underline{a}$, ou seja, $\underline{a}'\underline{X} \sim N(\underline{a}'\underline{\mu}, \underline{a}'\Sigma\underline{a})$;
 - b) As q combinações lineares $\mathbf{A}\underline{X}$, para \mathbf{A} a matriz de constantes $p \times q$, têm distribuição Normal Multivariada de vetor de médias $\mathbf{A}\underline{\mu}$ e matriz de covariâncias $\mathbf{A}\Sigma\mathbf{A}'$, ou seja, $\mathbf{A}\underline{X} \sim N_q(\mathbf{A}\underline{\mu}, \mathbf{A}\Sigma\mathbf{A}')$.

2. Distribuição Normal Multivariada Padrão

Sendo $\underline{Z} = (\Sigma)^{-\frac{1}{2}}(\underline{X} - \underline{\mu})$, temos que \underline{Z} tem distribuição Normal Multivariada Padrão, isto é, $\underline{Z} \sim N_p(\underline{0}, \mathbf{I})$, com $\underline{0}$ o vector de zeros de dimensão p e \mathbf{I} a matriz identidade de ordem p .

3. Normalidade das margens

- a) Para $q < p$, seja \underline{Y} um subconjunto de \underline{X} . Temos que $\underline{Y} \sim N_q(\underline{\mu}_Y, \Sigma_Y)$ sendo $\underline{\mu}_Y$ o vetor de médias de \underline{Y} e Σ_Y a matriz de covariâncias de \underline{Y} . Assim, é possível provar que um vetor só tem distribuição Normal Multivariada se todas as suas partições derem origem a subvetores com distribuição Normal Multivariada;
- b) Dados dois subvetores de \underline{X} , $\underline{Y}'_1 = (X_1, X_2, \dots, X_q)$ e $\underline{Y}'_2 = (X_{q+1}, X_{q+2}, \dots, X_p)$ com q e $p - q$ variáveis respectivamente, a partição constituída por estes subvetores tem distribuição Normal Multivariada (ou Univariada caso $q = 1$ ou $p - q = 1$), isto é, $\underline{Y}_1 \sim N_q(\underline{\mu}_1, \Sigma_{11})$ e $\underline{Y}_2 \sim N_{p-q}(\underline{\mu}_2, \Sigma_{22})$. Ou ainda, considerando p partições de \underline{X} , temos que $X_j \sim N(\mu_j, \sigma_j^2)$ com $j = 1, \dots, p$. É importante notar que se $X_j \sim N(\mu_j, \sigma_j^2)$, então não se tem necessariamente que $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$;
- c) No que diz respeito à distribuição conjunta de dois, ou mais, subvetores de $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$, digamos $\underline{Y}_1 \sim N_q(\underline{\mu}_1, \Sigma_{11})$ e $\underline{Y}_2 \sim N_{p-q}(\underline{\mu}_2, \Sigma_{22})$, esta é também Normal Multivariada de parâmetros

$$\underline{\mu} = \begin{bmatrix} \underline{\mu}_1 \\ \underline{\mu}_2 \end{bmatrix} \text{ e } \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

4. Independência

Sejam $\underline{Y}'_1 = (X_1, X_2, \dots, X_q)$ e $\underline{Y}'_2 = (X_{q+1}, X_{q+2}, \dots, X_p)$ dois subvetores resultantes da partição de \underline{X} em 2 conjuntos com q e $p - q$ elementos, respectivamente. Então temos que

- a) Dois subvetores são independentes se, e só se, $\Sigma_{\underline{Y}_1 \underline{Y}_2} = \mathbf{0}$;
- b) Duas variáveis X_i e X_j , $i \neq j$, são independentes se, e só se, $\sigma_{ij} = 0$.

Notemos que para esta propriedade é fundamental que as populações sejam normais.

5. Distribuição Condicional

Se os subvetores \underline{Y}_1 e \underline{Y}_2 de \underline{X} não forem independentes, então $\Sigma_{\underline{Y}_1 \underline{Y}_2} \neq \mathbf{0}$ e a distribuição condicionada de \underline{Y}_2 dado \underline{Y}_1 é Normal Multivariada com vetor de médias

$$E(\underline{Y}_2 | \underline{Y}_1 = \underline{y}_1) = \underline{\mu}_{\underline{Y}_2} + \Sigma_{\underline{Y}_2 \underline{Y}_1} \Sigma_{\underline{Y}_1 \underline{Y}_1}^{-1} (\underline{y}_1 - \underline{\mu}_{\underline{Y}_1})$$

e matriz de covariâncias

$$Cov(\underline{Y}_2 | \underline{Y}_1 = \underline{y}_1) = \Sigma_{\underline{Y}_2 \underline{Y}_2} - \Sigma_{\underline{Y}_2 \underline{Y}_1} \Sigma_{\underline{Y}_1 \underline{Y}_1}^{-1} \Sigma_{\underline{Y}_1 \underline{Y}_2},$$

onde \underline{y}_1 é o vector amostral aleatório observado de \underline{Y}_1 .

6. Distribuição da soma

Consideremos que \underline{Y} tal que $\underline{Y} \sim N_p(\underline{\mu}_Y, \Sigma_Y)$ e ainda que \underline{X} e \underline{Y} são independentes. Então

$$\underline{X} \pm \underline{Y} \sim N_p(\underline{\mu} \pm \underline{\mu}_Y, \Sigma + \Sigma_Y).$$

2.2.2 Estimação

Seja $(\underline{X}_1, \dots, \underline{X}_n)$ uma amostra aleatória de uma população Normal Multivariada com média $\underline{\mu}$, matriz de covariâncias Σ e $\underline{X}_i, i \in (1, \dots, n)$, mutuamente independentes e identicamente distribuídas (i.i.d.), ou seja,

$$(\underline{X}_1, \dots, \underline{X}_n) \stackrel{iid}{\sim} N_p(\underline{\mu}, \Sigma).$$

A função de densidade conjunta é determinada tendo em conta que $\underline{X}_i, i \in (1, \dots, n)$ são i.i.d. e, por isso, é dada por

$$f_{(\underline{X}_1, \dots, \underline{X}_n)}(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n) = \prod_{i=1}^n \frac{1}{2\pi^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(\underline{x}_i - \underline{\mu})' \Sigma^{-1} (\underline{x}_i - \underline{\mu})}. \quad (2.1)$$

Quando na função dada em (2.1) se considera os vetores observados $(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)$ do vetor aleatório \underline{X} fixo, então passa a ser uma função dos parâmetros $\underline{\mu}$ e Σ denominada de função de verosimilhança, L , sobre a qual falaremos mais à frente na Secção 2.4. Conhecida a função de verosimilhança é possível calcular o estimador de máxima verosimilhança (EMV) para os parâmetros da distribuição, $\underline{\mu}$ e Σ , sendo que os estimadores encontrados por esse método são os que maximizam a função de verosimilhança. Os estimadores de máxima verosimilhança de $\underline{\mu}$ e Σ são dados respetivamente por (Anderson, 2003 [1]):

$$\hat{\underline{\mu}} = \bar{\underline{X}} \quad \text{e} \quad \hat{\Sigma} = \frac{(n-1)}{n} \mathbf{S} \quad (2.2)$$

onde $\bar{\underline{X}}$ é a média amostral dada por $\bar{\underline{X}} = (\frac{1}{n} \sum_{i=1}^n x_{1i}, \frac{1}{n} \sum_{i=1}^n x_{2i}, \dots, \frac{1}{n} \sum_{i=1}^n x_{pi})$ com x_{ij} a entrada (i, j) da matriz observada

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

e \mathbf{S} é a matriz de covariâncias amostrais dada por

$$\mathbf{S} = \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1p} \\ s_{21} & s_{22} & \dots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \dots & s_{np} \end{bmatrix},$$

com s_{ij} para $i \neq j$ e $i, j = 1, \dots, n$ a covariância amostral entre as variáveis i e j , e com s_{ii} para $i = 1, \dots, n$ a variância amostral da variável i .

Os valores observados de (2.2) são as estimativas de máxima verosimilhança de $\underline{\mu}$ e Σ respetivamente. Por $\hat{\Sigma}$ ser um estimador enviesado de Σ é usual recorrer a \mathbf{S} para estimar Σ .

2.2.3 Distribuições de amostragem de $\bar{\underline{X}}$

Em muitos casos torna-se útil conhecer não só a distribuição da população como também a distribuição das suas estatísticas amostrais, ou seja, a distribuição de amostragem.

Considerando, em particular, $(\underline{X}_1, \dots, \underline{X}_n)$ uma amostra aleatória de uma população Normal Multivariada tal que $\underline{X}_i \sim N_p(\underline{\mu}, \Sigma)$ para todo $i = 1, \dots, n$, vem que $\bar{\underline{X}}$ tem distribuição de amostragem normal de vetor de médias $\underline{\mu}$ e matriz de covariâncias $\frac{\Sigma}{n}$, isto é,

$$\bar{\underline{X}} \sim N_p\left(\underline{\mu}, \frac{\Sigma}{n}\right).$$

É fácil mostrar que,

$$n(\bar{\underline{X}} - \underline{\mu})' \Sigma^{-1} (\bar{\underline{X}} - \underline{\mu}) \sim \chi_p^2.$$

No caso em que a amostra aleatória não é proveniente de uma distribuição normal, é conhecido o seguinte teorema que representa a adaptação do Teorema do Limite Central (TLC) ao caso multivariado:

Teorema 1. *Sejam $(\underline{X}_1, \dots, \underline{X}_n)$ observações independentes de uma qualquer população $\underline{X} = (X_1, \dots, X_p)^T$ com valor médio $\underline{\mu}$ e com covariância finita Σ . Para amostras de grande dimensão (n deve ser grande relativamente a p) temos que*

$$\bar{\underline{X}} \stackrel{a}{\sim} N_p\left(\underline{\mu}, \frac{\Sigma}{n}\right).$$

Logo, pelo Teorema 1, que pode ser consultado em Rencher e Christensen (2012) [18], para uma amostra aleatória $(\underline{X}_1, \dots, \underline{X}_n)$ de variáveis independentes e identicamente distribuídas (i.i.d.) provenientes de uma distribuição multivariada qualquer, temos, para n suficientemente grande, que

$$\bar{\underline{X}} \stackrel{a}{\sim} N_p\left(\underline{\mu}, \frac{\Sigma}{n}\right) \iff \sqrt{n}(\bar{\underline{X}} - \underline{\mu}) \stackrel{a}{\sim} N_p(\underline{0}, \Sigma) \iff \sqrt{n}\Sigma^{-\frac{1}{2}}(\bar{\underline{X}} - \underline{\mu}) \stackrel{a}{\sim} N_p(\underline{0}, \mathbf{I}).$$

Assim, podemos concluir que

$$n(\bar{\underline{X}} - \underline{\mu})' \Sigma^{-1} (\bar{\underline{X}} - \underline{\mu}) \stackrel{a}{\sim} \chi_p^2.$$

2.3 Distribuição Wishart

Definição 1. *Para $i = 1, \dots, n$, sejam \underline{W}_i vetores aleatórios i.i.d. tais que $\underline{W}_i \sim N_p(\underline{0}, \Sigma)$. Para $n \geq p$, a matriz \mathbf{W} , de dimensão $p \times p$, dada por*

$$\mathbf{W} = \sum_{i=1}^n \underline{W}_i \underline{W}_i'$$

diz-se ter Distribuição de Wishart com parâmetros n (graus de liberdade) e Σ . Neste caso, representa-se por $\mathbf{W} \sim W_p(n, \Sigma)$.

Quando $p = 1$ vem $\mathbf{W} = \sum_{i=1}^n \underline{W}_i \underline{W}'_i \sim W_1(n, \sigma^2)$, onde $W_i \sim N(0, \sigma^2)$, $i = 1, \dots, p$, são i.i.d.. Logo

$$z_i = \frac{W_i}{\sigma} \sim N(0, 1)$$

e portanto

$$z_i^2 = \frac{W_i^2}{\sigma^2} \sim \chi_1^2.$$

Assim concluímos que

$$\sum_{i=1}^n z_i^2 = \sum_{i=1}^n \frac{W_i^2}{\sigma^2} = \frac{\sum_{i=1}^n W_i^2}{\sigma^2} = \frac{\mathbf{W}}{\sigma^2} \sim \chi_n^2.$$

2.3.1 Propriedades

Três das propriedades mais importantes da distribuição de Wishart são:

1. Dadas $\mathbf{W}_1, \dots, \mathbf{W}_k$ matrizes aleatórias independentes com $\mathbf{W}_i \sim W_p(n_i, \Sigma)$, $i = 1, \dots, k$ vem que

$$\sum_{i=1}^k \mathbf{W}_i \sim W_p\left(\sum_{i=1}^k n_i, \Sigma\right).$$

2. Se $\mathbf{W} \sim W_p(n, \Sigma)$ e $\underline{a}_{p \times 1}$ é um vetor de dimensão p , então

$$\frac{\underline{a}' \mathbf{W} \underline{a}}{\underline{a}' \Sigma \underline{a}} \sim \chi_n^2, \text{ para } \underline{a}' \Sigma \underline{a} \neq 0.$$

3. Dados $\mathbf{W} \sim W_p(n, \Sigma)$ com $n > p$ e \mathbf{M} uma matriz $k \times p$ não aleatória de característica k , temos

$$\mathbf{M} \mathbf{W} \mathbf{M}' \sim W_k(n, \mathbf{M} \Sigma \mathbf{M}').$$

2.3.2 Distribuição de amostragem de S

Dado $\underline{X}_i \sim N_p(\underline{\mu}, \Sigma)$, temos que $\underline{W}_i = (\underline{X}_i - \underline{\mu})(\underline{X}_i - \underline{\mu})' \sim N_p(0, \Sigma)$, $i = 1, \dots, n$. Logo, a matriz dada por $\sum_{i=1}^n \underline{W}_i \underline{W}'_i$ tem distribuição de Wishart, ou seja,

$$\sum_{i=1}^n \underline{W}_i \underline{W}'_i = \sum_{i=1}^n (\underline{X}_i - \underline{\mu})(\underline{X}_i - \underline{\mu})' \sim W_p(n, \Sigma).$$

Fazendo a substituição de $\underline{\mu}$ por $\bar{\underline{X}}$, o seu estimador de máxima verosimilhança, vem que

$$\sum_{i=1}^n (\underline{X}_i - \bar{\underline{X}})(\underline{X}_i - \bar{\underline{X}})' = (n-1)\mathbf{S},$$

e daqui mostramos que

$$(n-1)\mathbf{S} \sim W_p(n-1, \Sigma)$$

é a distribuição de amostragem da matriz de covariâncias amostrais \mathbf{S} .

2.4 Testes de razão de verosimilhanças

Dado um vetor aleatório $\underline{X}' = (X_1, \dots, X_p)$ e uma amostra observada (x_1, \dots, x_n) do mesmo, é possível obter a função de verosimilhança $L(\theta)$. Aqui, θ representa o conjunto dos parâmetros da distribuição de \underline{X} e o conjunto dos valores que os parâmetros podem tomar podem ser representados por Θ .

Para testar se um parâmetro da distribuição do vetor aleatório \underline{X} tem um determinado valor fixo e conhecido, isto é, para testar

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \notin \Theta_0$$

é possível utilizar o método de razão de verosimilhanças, no qual a estatística de teste é

$$\Lambda = \frac{\max_{\theta \in \Theta_0} L(\theta)}{\max_{\theta \in \Theta} L(\theta)}. \quad (2.3)$$

Este método tem por base que a hipótese nula é um caso particular da hipótese alternativa e pretende testar qual das duas hipóteses se ajusta melhor aos dados da amostra. Aqui a decisão de rejeição ou não da hipótese nula assenta na teoria de quanto maior a verosimilhança, melhor é o ajuste dos dados, visto que este valor representa o valor do parâmetro para o qual é mais provável obter a amostra usada.

Assim, é fácil de concluir que se na expressão (2.3) o numerador for muito menor que o denominador, então rejeitamos H_0 , pois nesse caso Θ_0 não contém valores que sejam possíveis para θ . Por outro lado, tomamos como regra de rejeição de H_0 ,

$$\text{Rejeitar } H_0 \text{ para } \Lambda < c \quad (2.4)$$

onde c é uma constante escolhida que faça sentido. Usualmente, nos testes de hipótese é escolhido um nível de significância α tal que

$$\alpha = P[\text{Rejeitar } H_0 | H_0 \text{ verdadeira}]$$

e a constante c em (2.4) é escolhida de forma a produzir um teste de nível de significância α , o que é possível quando conhecida a distribuição de Λ . A distribuição de Λ nem sempre é fácil de encontrar, no entanto existem algumas aproximações estudadas. Por exemplo, Wilks (1938) [22] e Rencher e Christensen (2012) [18] propõe aproximações a uma distribuição χ^2 . Coelho (2004) [6] e Coelho e Marques (2010) [7] propõem as distribuições quase exatas. Para algumas estatísticas de razão de verosimilhanças Box (1949) [2] propõe aproximações baseadas em misturas de distribuições χ^2 e Butler et al. (1993) [3] são propostas as aproximações ponto de sela ou "Saddle Point".

3.1 Testes de razão de verosimilhanças para independência de grupos de variáveis

Existem várias aplicações possíveis para os testes de razão de verosimilhanças, sendo uma delas testar a independência de grupos de variáveis em populações Normais Multivariadas.

Consideremos o vetor aleatório $\underline{X}' = (X_1, \dots, X_p)$ tal que $\underline{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$ e k subconjuntos de \underline{X} , ou seja, $\underline{X} = (\underline{Y}_1, \dots, \underline{Y}_k)$ tal que \underline{Y}_i é constituído por p_i variáveis de tal maneira que $p = p_1 + \dots + p_k$, com $i = 1, \dots, k$. A matriz de covariâncias de \underline{X} pode ser escrita como

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \dots & \Sigma_{1k} \\ \Sigma_{21} & \Sigma_{22} & \dots & \Sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{k1} & \Sigma_{k2} & \dots & \Sigma_{kk} \end{bmatrix},$$

onde Σ_{ij} representa a matriz de covariâncias entre \underline{Y}_i e \underline{Y}_j .

Pretendendo-se testar a independência dos k vetores aleatórios e por se tratar de um vetor aleatório com distribuição Normal Multivariada, a hipótese de independência pode ser escrita como

$$H_0 : \Sigma_{ij} = \mathbf{0}, \quad \forall i \neq j. \quad (3.1)$$

A estatística de razão de verosimilhanças para testar a hipótese nula em (3.1) é

$$\Lambda = \left(\frac{|\hat{\Sigma}|}{\prod_{i=1}^k |\hat{\Sigma}_{ii}|} \right)^{\frac{n}{2}}$$

com $\hat{\Sigma}$ o estimador de máxima verosimilhança para Σ e $|\hat{\Sigma}_{ii}|$ o i -ésimo elemento da diagonal de $\hat{\Sigma}$.

Uma vez que a distribuição exata de Λ é de difícil manipulação (Rencher & Christensen, 2012 [18]), para realizar o teste podemos utilizar a estatística

$$U^* = -(n-1)c \ln(\Lambda)$$

com

$$c = 1 - \frac{1}{12 \frac{p^2 - \sum_{i=1}^k p_i^2}{2} (n-1)} \left(2(p^3 - \sum_{i=1}^k p_i^3) + 3(p^2 - \sum_{i=1}^k p_i^2) \right).$$

Como \underline{X} tem distribuição Normal Multivariada, então Rencher e Christensen (2012) [18] mostram que U^* tem distribuição aproximada χ^2 com $m = \frac{p^2 - \sum_{i=1}^k p_i^2}{2}$ graus de liberdade. Assim, para um nível de significância α rejeitamos H_0 para

$$U^* > \chi_{m,1-\alpha}^2,$$

onde $\chi_{m,\alpha}^2$ denota o quantil de ordem α .

Notemos que este teste pode ser ainda aplicado no caso particular de $k = p$, ou seja, em que cada subvetor \underline{Y}_i é composto por apenas uma variável aleatória X_i . É este caso particular que este trabalho vai analisar com mais detalhe.

3.1.1 Testes de razão de verosimilhanças para independência completa de variáveis

No caso de estarmos a trabalhar com populações normais multivariadas, as variáveis são independentes se, e só se, a matriz de covariâncias for diagonal. Ou seja, podemos testar a independência das variáveis recorrendo a um teste cuja hipótese nula é dada por:

$$H_0: \sigma_{ij}^2 = 0, \quad \forall i, j \text{ com } i \neq j. \quad (3.2)$$

A estatística do teste de razão de verosimilhanças usado para testar esta hipótese nula é:

$$\Lambda = \left(\frac{|\hat{\Sigma}|}{\prod_{i=1}^p s_{ii}} \right)^{\frac{n}{2}} \quad (3.3)$$

onde $\hat{\Sigma}$ é o estimador de máxima verosimilhança para Σ e s_{ii} é o i -ésimo elemento da diagonal de $\hat{\Sigma}$.

3.2 Testes de razão de verosimilhanças para independência completa em estruturas específicas

Os testes de razão de verosimilhanças podem ser utilizados em testes cujo tanto a hipótese nula como a hipótese alternativa correspondem à hipótese de que a matriz de covariâncias assume determinada estrutura específica.

No momento da escolha das estruturas a considerar é possível ainda garantir que além de estruturas específicas se teste simultaneamente a independência das variáveis

3.2. TESTES DE RAZÃO DE VEROSIMILHANÇAS PARA INDEPENDÊNCIA COMPLETA EM ESTRUTURAS ESPECÍFICAS

no caso das matrizes consideradas nas hipóteses representarem matrizes de covariância de um determinado vetor aleatório proveniente de uma população Normal Multivariada, bastando para isso garantir que a estrutura considerada na hipótese nula é uma matriz diagonal.

Outro ponto a ter aqui em conta, é o facto de, como já vimos na Secção 2.4 ser necessário garantir que a estrutura considerada sob a hipótese nula é um caso particular da estrutura considerada sob a hipótese alternativa.

3.2.1 Teste de Esfericidade vs. Equivariância-Equicorrelação

Para testar a igualdade das variâncias e a independência completa das variáveis para um determinado vetor $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$, é comumente usado o teste de esfericidade (Mauchly, 1940 [15]; Anderson, 2003 [1]; Marques & Coelho, 2008 [11]; Coelho & Marques, 2010 [7]). A hipótese nula do teste de esfericidade é dada por

$$H_0 : \Sigma = \sigma^2 I_p = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix}$$

onde I_p é a matriz identidade de ordem p e a estatística de razão de verosimilhança é bem conhecida e dada em Marques e Coelho (2015) [13] e Anderson (2003) [1].

Se, por outro lado, se pretender estudar se a variância das variáveis aleatórias que compõe o vetor são iguais e a suas covariâncias também o são, então podemos recorrer ao teste de equivariância-equicorrelação, apresentado por Wilks (1946) [21] e Olkin e Press (1969) [17], sendo este teste também conhecido por teste simétrico composto. A hipótese nula do mesmo é

$$H_0 = \Sigma = \Sigma_{CS} = \begin{bmatrix} \sigma^2 & \sigma^2 \rho & \dots & \sigma^2 \rho \\ \sigma^2 \rho & \sigma^2 & \dots & \sigma^2 \rho \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2 \rho & \sigma^2 \rho & \dots & \sigma^2 \end{bmatrix}$$

onde ρ é um número real tal que $\frac{-1}{p-1} < \rho < 1$ e a estatística para este teste é dada em Wilks (1946) [21] e Olkin e Press (1969) [17].

Assim, Marques e Coelho (2015) [13] propõem o teste de esfericidade vs equivariância-equicorrelação que consiste em testar

$$H_0 : \Sigma = \sigma^2 I_p = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix} \text{ vs } H_1 = \Sigma = \Sigma_{CS} = \begin{bmatrix} \sigma^2 & \sigma^2 \rho & \dots & \sigma^2 \rho \\ \sigma^2 \rho & \sigma^2 & \dots & \sigma^2 \rho \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2 \rho & \sigma^2 \rho & \dots & \sigma^2 \end{bmatrix}, \quad (3.4)$$

onde I_p é a matriz identidade de ordem p e ρ é um número real tal que $\frac{-1}{p-1} < \rho < 1$.

Notemos que está garantido que a estrutura considerada em H_0 é um caso particular da estrutura considerada em H_1 pois $\Sigma_{CS} = \sigma^2((1 - \rho)I_p + \rho J_p)$, com J_p a matriz de ordem p cujos elementos são todos iguais a um.

Marques e Coelho (2015) [13] obtiveram uma estatística de teste para testar as hipóteses em (3.4), sendo dada por:

$$\Lambda = \frac{\max_{\underline{\mu}, \sigma^2} L(\underline{\mu}, \sigma^2 I_p)}{\max_{\underline{\mu}, \sigma^2, \rho} L(\underline{\mu}, \Sigma_C)} = \frac{\frac{\max_{\underline{\mu}, \sigma^2} L(\underline{\mu}, \sigma^2 I_p)}{\max_{\underline{\mu}, \Sigma} L(\underline{\mu}, \Sigma)}}{\frac{\max_{\underline{\mu}, \sigma^2, \rho} L(\underline{\mu}, \Sigma_C)}{\max_{\underline{\mu}, \Sigma} L(\underline{\mu}, \Sigma)}} = \frac{\Lambda_S}{\Lambda_C} = \left(\frac{p^p}{(p-1)^{p-1}} \frac{a_{11}^* \text{tr}(A_2^*)^{p-1}}{\text{tr}(A^*)^p} \right)^{\frac{n}{2}} \quad (3.5)$$

onde, para a amostra observada (x_1, \dots, x_n) de \underline{X} a função de verosimilhança é dada por $L(\underline{\mu}, \Sigma)$; Λ_S é a estatística de razão de verosimilhanças para testar a esfericidade e Λ_C é a estatística de razão de verosimilhanças para testar a equivariância-euicorrelação sendo que em ambos os casos a hipótese alternativa é a matriz de covariância ter qualquer outra estrutura que não a estrutura considerada na hipótese nula; $A^* = \Gamma A \Gamma$ com A o estimador de máxima verosimilhança da matriz de covariâncias, Γ uma matriz de Helmert de ordem p que não é função de Σ ; a_{11}^* é o elemento da primeira linha e primeira coluna de A e A_2^* é o bloco diagonal de A^* que se segue a a_{11}^* . Notemos que sob H_0 os elementos de A^* são independentes.

Além disso, os autores ainda obtiveram aproximações assintóticas para a estatística de teste $\Lambda^* = \Lambda^{\frac{n-1}{n}}$ que são de fácil utilização sem comprometer a precisão dos resultados.

Notemos que, a não rejeição de H_0 permite concluir para um nível de significância α que as variáveis são independentes.

3.2.2 Teste de Esfericidade vs Circularidade

As matrizes de estrutura circular são matrizes do tipo

$$\Sigma_C = \begin{bmatrix} \sigma^2 & b_1 & \dots & b_{p-1} \\ b_{p-1} & \sigma^2 & \dots & b_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ b_1 & b_2 & \dots & \sigma^2 \end{bmatrix},$$

com $b_j = b_{p-j}$ com $j = 1, 2, \dots, \lfloor \frac{p}{2} \rfloor$, onde $\lfloor x \rfloor$ representa o maior número inteiro menor que x . Por exemplo, duas matrizes de estrutura circular de ordem 4 e 5 respetivamente são

$$\begin{bmatrix} 1 & b_1 & b_2 & b_1 \\ b_1 & 1 & b_1 & b_2 \\ b_2 & b_1 & 1 & b_1 \\ b_1 & b_2 & b_1 & 1 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 1 & b_1 & b_2 & b_2 & b_1 \\ b_1 & 1 & b_1 & b_2 & b_2 \\ b_2 & b_1 & 1 & b_1 & b_2 \\ b_2 & b_2 & b_1 & 1 & b_1 \\ b_1 & b_2 & b_2 & b_1 & 1 \end{bmatrix}. \quad (3.6)$$

Esta estrutura tem particular interesse em algumas áreas, nomeadamente, em engenharia, Gray (2006) [8], análise de séries temporais, estatísticas espaciais e mais algumas.

A hipótese nula do teste de circularidade é dada por:

$$H_0 : \Sigma = \Sigma_C = \begin{bmatrix} \sigma^2 & b_1 & \dots & b_{p-1} \\ b_{p-1} & \sigma^2 & \dots & b_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ b_1 & b_2 & \dots & \sigma^2 \end{bmatrix} \quad (3.7)$$

e a estatística do teste de razão de verosimilhanças para este teste é dada em Olkin e Press (1969) [17].

Em Olkin e Press (1969) [17] os autores propuseram o teste

$$H_0 : \Sigma = \sigma^2 I_p = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix} \quad \text{vs} \quad H_1 : \Sigma = \Sigma_C = \begin{bmatrix} \sigma^2 & b_1 & \dots & b_{p-1} \\ b_{p-1} & \sigma^2 & \dots & b_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ b_1 & b_2 & \dots & \sigma^2 \end{bmatrix} \quad (3.8)$$

onde I_p e $b_j = b_{p-j}$ são como definidos para (3.4). Os autores deduziram a estatística do teste de razão de verosimilhanças, que é dada por

$$\Lambda^{\frac{2}{n-1}} = \frac{p^p \prod_{j=1}^p v_j}{2^{2(p-m-1)} (\sum_{i=1}^{m+1} v_i)^p} \quad (3.9)$$

onde, $m = \lfloor \frac{p}{2} \rfloor$ é o maior inteiro menor que $\frac{p}{2}$, e para v_j dados pelas equações (2.5a) e (2.5b) em Olkin e Press (1969) [17].

Além da dedução da estatística de teste os autores também provaram que a distribuição de (3.9) pode ser representada como o produto de distribuições Betas independentes, sendo uma boa aproximação dada por uma mistura de distribuições χ^2 .

3.3 Teste de Schoot

Em Schoot (2005) [19] o autor apresenta um teste que pode ser utilizado para testar a independência completa das variáveis de um vetor aleatório $\underline{X}' = (X_1, \dots, X_p)$ proveniente de uma distribuição Normal Multivariada de média $\underline{\mu}$ e matriz de covariâncias Σ , ou seja, $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$. O autor propõe que a hipótese de independência das variáveis seja escrita de forma equivalente

$$H_0 : \rho_{ij} = 0 \quad (i > j), \quad (3.10)$$

onde ρ_{ij} é o (i, j) -ésimo elemento da matriz de correlação P de \underline{X} .

A estatística de teste mais simples para testar (3.10) é

$$\sum_{i=2}^p \sum_{j=1}^{i-1} r_{ij}^2$$

onde r_{ij} é o (i, j) -ésimo elemento da matriz de correlação amostral R obtida através de um vetor amostral de \underline{X} . No entanto, o autor prova, utilizando as propriedades do valor esperado, que é possível encontrar a estatística de teste de média nula dada por

$$t_{(n-1)p} = \sum_{i=2}^p \sum_{j=1}^{i-1} r_{ij}^2 - \frac{p(p-1)}{2(n-1)},$$

onde $E\left(\sum_{i=2}^p \sum_{j=1}^{i-1} r_{ij}^2\right) = \frac{p(p-1)}{2(n-1)}$.

De forma análoga, Schoot (2005) [19] também prova que a variância de $t_{(n-1)p}$ é dada por

$$\sigma_{t_{(n-1)p}}^2 = \frac{p(p-1)(n-2)}{(n-1)^2(n+1)}$$

e que portanto $\frac{t_{(n-1)p}}{\sigma_{t_{(n-1)p}}}$ é a estatística de teste que sob H_0 tem média 0 e variância 1. Notemos que a normalidade assintótica de $t_{(n-1)p}$ é garantida pois quando $n, p \rightarrow \infty$ então

$$\lim \frac{p}{n-1} = \gamma \text{ com } \gamma \in [0, \infty[. \quad (3.11)$$

De Schoot (2005) [19] é ainda importante de referir o seguinte teorema:

Teorema 2. *Suponhamos que a matriz de correlação amostral R foi obtida de uma amostra aleatória de uma população Normal Multivariada com matriz de correlação P . Se $P = I_p$ e for válida a condição (3.11) então $t_{(n-1)p}$ converge em distribuição para uma variável aleatória normal de média 0 e variância*

$$\gamma^2 = \lim \sigma_{t_{(n-1)p}}^2 = \lim \frac{p(p-1)(n-2)}{(n-1)^2(n+1)}.$$

3.4 Redução Univariada

Dadas X_1, \dots, X_p variáveis aleatórias, sabemos que

$$X_i \text{ independentes } \forall i = 1, \dots, p \Rightarrow \text{Var}\left(\sum_{i=1}^p X_i\right) = \sum_{i=1}^p \text{Var}(X_i). \quad (3.12)$$

No entanto, o recíproco de (3.12) nem sempre é válido, isto é, a variância da soma das variáveis X_1, \dots, X_p ser igual à somas das variâncias dessas mesmas variáveis não implica que as mesmas sejam independentes. O recíproco falha quando existe covariância entre variáveis mas estas covariâncias quando todas somadas se anulam, isto é, a soma do valor absoluto de todas as covariâncias positivas for igual à soma do valor absoluto das covariâncias negativas. Assim, se tomarmos o pressuposto de que as covariâncias entre as variáveis quando existem são todas não negativas ou não positivas, então obtemos a veracidade do recíproco de (3.12). Ou seja, suponhamos que

$$\text{Var}\left(\sum_{i=1}^p X_i\right) = \sum_{i=1}^p \text{Var}(X_i),$$

e que

$$Cov(X_i, X_j) \geq 0 \text{ (ou } Cov(X_i, X_j) \leq 0)$$

para $i, j = 1, \dots, p$ com $i \neq j$. Por outro lado, como

$$Var\left(\sum_{i=1}^p X_i\right) = \sum_{i=1}^p Var(X_i) + 2 \sum_{0 \leq i < j \leq p} Cov(X_i, X_j),$$

com $Cov(X_i, X_j)$ a covariância entre X_i e X_j , então daqui vem que

$$2 \sum_{0 \leq i < j \leq p} Cov(X_i, X_j) = 0 \Leftrightarrow \sum_{0 \leq i < j \leq p} Cov(X_i, X_j) = 0$$

mas como

$$Cov(X_i, X_j) \geq 0 \text{ (ou } Cov(X_i, X_j) \leq 0)$$

então sai que

$$Cov(X_i, X_j) = 0 \forall i, j = 1, \dots, p \text{ com } i \neq j.$$

Assim, sob a suposição da normalidade do vetor aleatório \underline{X} e das covariâncias serem todas não negativas ou não positivas, com $W = \sum_{i=1}^p X_i$, para testar a independência das variáveis torna-se suficiente testar

$$H_0 : Var(W) = Var\left(\sum_{i=1}^p X_i\right) = \sum_{i=1}^p Var(X_i). \quad (3.13)$$

Consideremos então a amostra $\mathbf{X} = (x_1, \dots, x_n)$ de tamanho n da população $\underline{X}' = (X_1, X_2, \dots, X_p) \sim N_p(\underline{\mu}, \Sigma)$, onde $\underline{\mu}' = (\mu_1, \mu_2, \dots, \mu_p)$ é o vetor de médias e Σ é a matriz de covariâncias e pode ser representada por

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \cdots & \sigma_{1p}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \cdots & \sigma_{2p}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1}^2 & \sigma_{p2}^2 & \cdots & \sigma_{pp}^2 \end{bmatrix}.$$

A matriz amostral por sua vez pode ser representada como

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

Considerando esta notação de seguida propomos a metodologia a usar para testar a independência das variáveis.

3.4.1 Metodologia do Teste de Redução Univariada

Passo 1:

Como o cenário mais real é as variâncias das variáveis aleatórias X_1, X_2, \dots, X_p não serem conhecidas o primeiro passo para esta metodologia é calcular uma estimativa para as mesmas, usando o estimador de máxima verosimilhança.

Passo 2:

Após serem obtidas todas as estimativas das variâncias, o passo seguinte consiste em obter uma amostra (w_1, \dots, w_n) de $W = \sum_{i=1}^p X_i$. Esta amostra obtém-se fazendo $w_j = \sum_{i=1}^p x_{ji}$, onde (x_{j1}, \dots, x_{jp}) , $i = 1, \dots, p$ é a j -ésima observação do vetor \underline{X} . Esta nova amostra é adicionada numa nova coluna à matriz amostral \mathbf{X} de forma a obtermos

$$\mathbf{X}^* = \left[\begin{array}{cccc|c} x_{11} & x_{12} & \dots & x_{1p} & w_1 \\ x_{21} & x_{22} & \dots & x_{2p} & w_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} & w_n \end{array} \right].$$

Passo 3:

Repetimos o passo 1, mas agora pretendemos encontrar a estimativa de máxima verosimilhança para W .

Passo 4:

Como já vimos, usando os pressupostos de que as variâncias quando existem são todas não negativas ou não positivas e de que o vetor aleatório $\underline{X}' = (X_1, X_2, \dots, X_p)$ é proveniente de uma população normal multivariada, então para testar a independência das variáveis é suficiente aplicar o teste (3.13), que podemos rescrever como,

$$H_0 : \sigma_W^2 = \sigma_0^2 \quad \text{vs} \quad \sigma_W^2 \neq \sigma_0^2.$$

Quando as variâncias são conhecidas, temos $\sigma_0^2 = \sum_{i=1}^p \sigma_{ii}^2$ e quando as mesmas não são conhecidas, $\sigma_0^2 = \sum_{i=1}^p s_{ii}^2$ com s_{ii}^2 a estimativa de máxima verosimilhança calculada no passo 1.

A estatística deste teste quando as variâncias são conhecidas é

$$\frac{(n-1)S_W^2}{\sum_{i=1}^p \sigma_{ii}^2} \sim \chi_{n-1}^2,$$

e quando as variâncias não são conhecidas é possível utilizar a distribuição assintótica dada por

$$\frac{(n-1)S_W^2}{\sum_{i=1}^p s_{ii}^2} \sim \chi_{n-1}^2.$$

Notemos que S_W^2 é o estimador de máxima verosimilhança obtido no passo 3.

O pressuposto da normalidade do vetor aleatório é fundamental para a dedução da distribuição das estatísticas dos testes, no entanto quando o mesmo não se verifica é possível considerar resultados assintóticos.

1. Quando $X_i, i = 1, \dots, p$, não são provenientes de populações normais, pelo Teorema do Limite Central, se as variáveis aleatórias forem independentes e identicamente distribuídas com variâncias $\sigma_{ii}^2 = \sigma^2$ finitas e médias $\mu_{ii} = \mu$ também finitas, então

$$W = \sum_{i=1}^p X_i \stackrel{a}{\sim} N \left(\sum_{i=1}^p \mu_{ii} = p\mu, \sum_{i=1}^p \sigma_{ii}^2 = p\sigma^2 \right). \quad (3.14)$$

2. Quando $X_i, i = 1, \dots, p$, não são provenientes de populações normais e as variáveis não são identicamente distribuídas com variâncias $\sigma_{ii}^2 = \sigma^2$, mas $|X_i|$ têm momentos finitos de alguma ordem $2 + \delta, \delta > 0$, com a taxa de crescimento limitada pela condição de Lyapunov, então se as variáveis forem independentes, pelo teorema de Lyapunov vem

$$W = \sum_{i=1}^p X_i \stackrel{a}{\sim} N \left(\sum_{i=1}^p \mu_i, \sum_{i=1}^p \sigma_{ii}^2 \right). \quad (3.15)$$

Com a perda da normalidade do vetor aleatório também se perde a garantia de que se as correlações forem nulas, então as variáveis são independentes. No entanto, diferentes cenários foram simulados na Secção 4 para garantir a robustez do teste de redução univariada, quando existe um pequeno desvio da normalidade.

Para garantir os resultados precisos na utilização do teste é ainda necessário ter em consideração algumas situações:

- Para garantir boas aproximações em (3.14) e (3.15), quando a população se desvia da normalidade, é necessário garantir que o número de variáveis p é grande o suficiente;
- Se as variâncias forem desconhecidas é importante que o tamanho da amostra n seja grande de forma a permitir uma estimativa o mais precisa possível;
- Esta metodologia pode ser usada mesmo quando estão em falta alguns valores;
- Uma das mais valias deste teste é o facto de permitir que o mesmo seja utilizado mesmo quando $p > n$, contrariamente ao que acontece por exemplo com o teste de razão de verosimilhanças (Anderson (2003 [1]));
- Este teste tem a característica de utilizar um teste univariado para inferir sobre os parâmetros de uma distribuição multivariada, o que é um benefício por se tornar mais simples que qualquer procedimento que envolva cálculos com distribuições multivariadas.

SIMULAÇÕES

Com o intuito de avaliar a qualidade do teste proposto na Secção 3.4, o teste de redução univariada para a independência de variáveis, foram realizadas simulações gerando aleatoriamente amostras de populações Normais Multivariadas. Para perceber o comportamento do teste em diferentes situações, foram simuladas várias amostras de diferentes tamanhos n e com diferentes número de variáveis p e para cada uma das combinações n e p foram testadas 3 tipos de estruturas diferentes para a matriz de covariâncias, sendo que se considerou em qualquer um dos casos que a amostra é proveniente de uma população Normal Multivariada de vetor de médias $\underline{\mu} = \underline{0}$. Além disso, para cada um dos casos considerados, como o pretendido é estudar os resultados obtidos quer as amostras sejam provenientes de variáveis aleatórias independentes ou dependentes, foram simuladas amostras para cada umas das 3 estruturas de covariâncias considerando diferentes valores de ρ que, como podemos consultar mais à frente nesta secção, é um valor do qual dependem as estruturas de covariâncias consideradas, tendo sido considerados os valores do vetor $\underline{\rho} = (\rho_1, \dots, \rho_k)$, com $\rho_1 < \dots < \rho_k$ e $\rho_i \in \mathbb{R} \forall i = 1, \dots, k$ que foi calculado garantindo que ρ_1 é o mínimo valor para o qual a matriz é definida positiva, que $\exists j$ com $1 < j < k$ tal que $\rho_j = 0$ e que ρ_k é tal que a potência do teste de redução univariada é próxima de 1. Para as estruturas 4.1 e 4.3 foi ainda garantido que $\underline{\rho}$ começa com valores negativos para as covariâncias, a menos que para valores negativos de ρ a matriz não seja definida positiva, e termina com valores positivos para as covariâncias.

As estruturas consideradas para as simulações foram então:

- Estrutura de equivariância-equicorrelação ou simétrica composta: consiste numa matriz cujas variâncias das p variáveis é a mesma e cujas covariâncias entre as variáveis são também as mesmas, ou seja, é a estrutura

$$\Sigma_{CS} = \sigma^2 \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{bmatrix}. \quad (4.1)$$

- Estrutura autorregressiva: é uma matriz cujas variáveis têm todas a mesma variância e cuja covariância é $\sigma^2 \rho^{|j-i|}$ onde $|j-i|$ representa o valor absoluto de $j-i$ para $i, j = 1, \dots, p$, ou seja,

$$\Sigma_{AR} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{p-1} \\ \rho & 1 & \rho & \dots & \rho^{p-2} \\ \vdots & \vdots & \ddots & \vdots & \\ \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \dots & 1 \end{bmatrix}. \quad (4.2)$$

- Estrutura circular: aqui a variância das variáveis é também a mesma e a covariância segue uma regra onde para cada linha cada entrada é movida apenas uma posição para a direita relativamente à linha anterior mantendo no entanto os valores da diagonal principal sempre σ^2 , ou seja, tem a seguinte estrutura:

$$\Sigma_C = \begin{bmatrix} \sigma^2 & b_1 & \dots & b_{p-1} \\ b_{p-1} & \sigma^2 & \dots & b_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ b_1 & b_2 & \dots & \sigma^2 \end{bmatrix}. \quad (4.3)$$

Na Secção 3.2.2 em (3.6) é possível consultar dois exemplos de matrizes circulares de ordem 4 e 5.

Notemos que para a matriz de estrutura circular (4.3) não é visível a dependência de ρ no entanto quando p é par

$$(b_1, \dots, b_{p-1}) = \left(\frac{-1}{200} + \rho, \frac{-1}{300} + \rho, \dots, \frac{-1}{100(r-1)} + \rho, \frac{-1}{100r} + \rho, \right. \\ \left. \frac{-1}{100(r+1)} + \rho, \frac{-1}{100r} + \rho, \frac{-1}{100(r-1)} + \rho, \dots, \frac{-1}{300} + \rho, \frac{-1}{200} + \rho \right)$$

e quando p é ímpar

$$(b_1, \dots, b_{p-1}) = \left(\frac{-1}{200} + \rho, \frac{-1}{300} + \rho, \dots, \frac{-1}{100(r-1)} + \rho, \frac{-1}{100r} + \rho, \frac{-1}{100(r+1)} + \rho, \right. \\ \left. \frac{-1}{100(r+1)} + \rho, \frac{-1}{100r} + \rho, \frac{-1}{100(r-1)} + \rho, \dots, \frac{-1}{300} + \rho, \frac{-1}{200} + \rho \right).$$

Como meio de comparação, para cada simulação é também aplicado o teste de razão de verosimilhanças (Anderson, 2003 [1]) e o teste de Schout (Schout, 2005 [19]).

Para as simulações a metodologia considerada foi a seguinte:

Passo 1: Encontramos o vetor $\underline{\rho}$ tendo em conta os pontos referidos anteriormente;

Passo 2: Calculamos a matriz de covariâncias para o primeiro valor;

Passo 3: Simulamos 2000 amostras, cada uma de tamanho n , para uma população Normal Multivariada com p variáveis de vetor de médias $\underline{\mu} = \underline{0}$ e matriz de covariâncias calculada no **Passo 2**;

Passo 4: Para cada uma das 2000 amostras calculamos a matriz de covariâncias amostrais;

Passo 5: Calculamos, para cada amostra, a estatística de teste de redução univariada para as variâncias desconhecidas (com base na matriz obtida no **Passo 4**) e para as variâncias conhecidas (com base na matriz obtida no **Passo 2**), a estatística de teste da razão de verossimilhanças (Anderson (2003 [1]), sendo que esta apenas pode ser calculada nos casos em que $n > p$, e a estatística do teste de Schoot (Schoot, 2005 [19]);

Passo 6: Somando o número de vezes que se rejeita a hipótese nula para cada um dos testes obtemos as estimativas para a potência dos testes dividindo este valor por 2000, o número de amostras simuladas;

Passo 7: Repetimos os passos de 2 a 5 para cada $\rho \in \underline{\rho}$;

O procedimento é repetido para cada uma das 3 estruturas de covariâncias e para cada combinação de p variáveis e amostra de tamanho n consideradas, que podem ser consultadas na Tabela 4.1.

| p : número de variáveis | n : tamanho das amostras |
|---------------------------|----------------------------|
| 5 | 50 |
| 5 | 200 |
| 15 | 50 |
| 15 | 200 |
| 25 | 20 |
| 45 | 50 |
| 50 | 40 |
| 50 | 200 |
| 75 | 50 |
| 100 | 80 |
| 150 | 120 |
| 200 | 150 |

Tabela 4.1: Combinações de p variáveis e amostras de tamanho n consideradas nas simulações

Complementarmente, foram ainda realizadas simulações partindo de amostras geradas aleatoriamente para populações t-Student Multivariadas para testar a robustez do teste quando as populações se afastam da distribuição Normal Multivariada, tendo aqui sido geradas amostras da distribuição t-Student Multivariadas com diferentes graus de liberdade.

Por último, para testar a versatilidade do teste, foram também realizadas algumas simulações para a utilização do mesmo nos casos em que se pretende testar a independência das variáveis assumindo estruturas de covariâncias específicas, ou seja, nos casos apresentados nas subsecções 3.2.1 e 3.2.2.

As simulações foram realizadas usando linguagem *Python* através da plataforma *Colab* do *Google Research*. O código utilizado nas simulações pode ser consultado no anexo I.

4.1 Estimativas das potências dos testes

Nesta secção serão analisados os resultados obtidos para as estimativas das potências dos testes nas simulações realizadas. Para cada cenário considerado, isto é, para cada estrutura foram considerados casos onde o número de variáveis é inferior ao tamanho da amostra, ou seja, $n > p$, mas também casos onde $n < p$, sendo que para este segundo cenário, apenas foram realizados os testes de redução univariada, tanto para as variâncias conhecidas como para as variâncias desconhecidas, e o teste de Schoot (Schoot, 2005 [19]) pois o teste de razão de verosimilhanças (Anderson, 2003 [1]) não pode ser aplicado nestes casos. Tal como já foi referido anteriormente, foram feitas simulações considerando diferentes matrizes de covariâncias construídas à custa de ρ , sendo que, o eixo dos xx em qualquer um dos gráficos representa esse valor. Nas figuras seguintes, as linhas a '–' e '· · ·' são os resultados das estimativas das potências do teste de redução univariada de variâncias desconhecidas e conhecidas, respetivamente, a '—' representa os resultados das estimativas das potências dos testes de razão de verosimilhanças (Anderson (2003 [1])) e a '—' temos as estimativas das potências do teste de Schoot (Schoot, 2005 [19]).

Na Figura 4.1 podemos consultar os resultados das simulações nos casos em que o número de variáveis é menor que o tamanho da amostra, para amostras geradas a partir de uma distribuição Normal Multivariada de matriz de covariâncias amostrais com estrutura simétrica composta, isto é, matriz com a estrutura definida em (4.1).

Na Figura 4.2 podemos consultar os resultados para amostras obtidas com a mesma estrutura para as covariâncias que na Figura (4.1), sendo que neste caso se apresentam as estimativas das potências dos testes no casos em que o número de variáveis é superior ao tamanho da amostra, isto é, $p > n$.

Analisando a Figura 4.1 é bastante visível que para, qualquer combinação de p e n , os resultados obtidos com qualquer um dos testes de redução univariada são significativamente melhores que os resultados obtidos com qualquer um dos outros dois testes, pois quando ρ se afasta de 0 a potência dos testes aproxima-se mais rapidamente de 1 do que em qualquer um dos outros casos. É também de notar que na grande maioria dos casos, e para quase todos os valores de ρ , a diferença entre a potência obtida nos testes de redução univariada para as variâncias conhecidas e desconhecidas é quase inexistente.

Na Figura 4.2 é ainda mais notório que o teste de redução univariada se comporta melhor que o teste de Schoot (Schoot, 2005 [19]), destacando-se esta diferença para os casos representados em 4.2(c) e 4.2(f) onde no caso do teste de redução univariada, independentemente das covariâncias serem conhecidas ou não, por mais mínima que sejam as covariâncias entre as variáveis, a regra de rejeição do teste é bastante precisa.

Nas Figuras 4.3 e 4.4, podemos visualizar os resultados obtidos para as potências dos testes nos casos em que as simulações foram feitas usando para a matriz de covariâncias da população a estrutura autorregressiva, ou seja, a estrutura dada em (4.2), e considerando o número de variáveis menor que o tamanho da amostra e também o número de variáveis maior que o tamanho da amostra, respetivamente. Notemos, que para esta estrutura, não

Figura 4.1: Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verossimilhanças quando $p < n$ para a matriz de covariâncias de estrutura simétrica composta

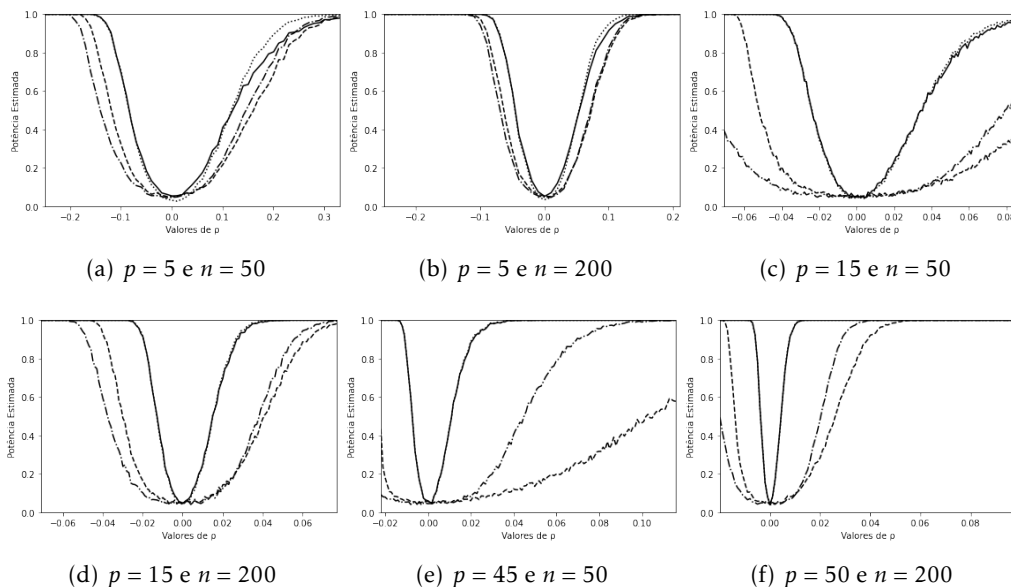
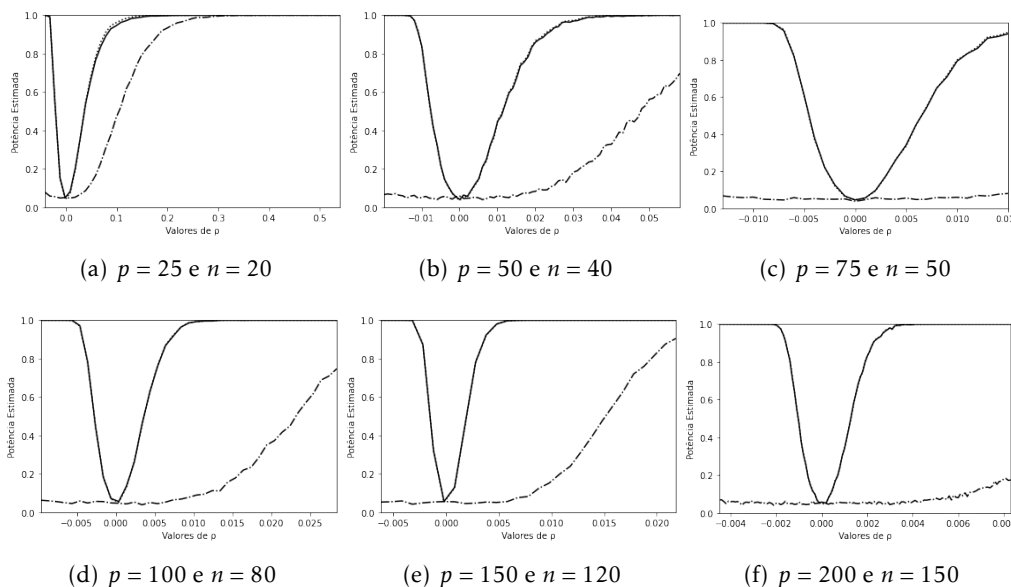
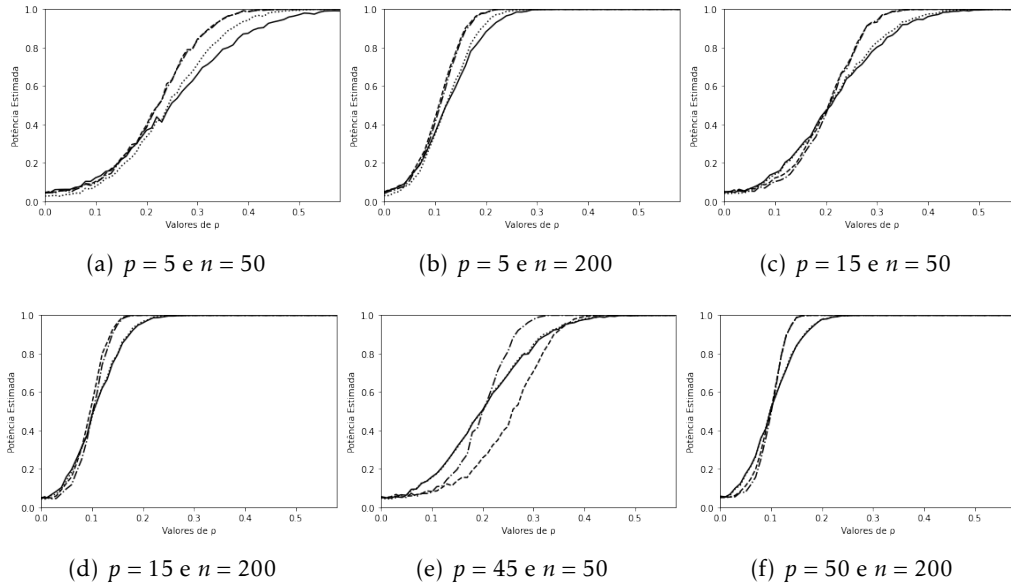


Figura 4.2: Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verossimilhanças quando $p > n$ para a matriz de covariâncias de estrutura simétrica composta



são apresentados resultados para diferentes valores de σ^2 pois esses casos foram testados não se tendo obtido resultados significativamente diferentes de quando se considera $\sigma^2 = 1$.

Figura 4.3: Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verossimilhanças com $p < n$ para a matriz de covariâncias de estrutura autorregressiva

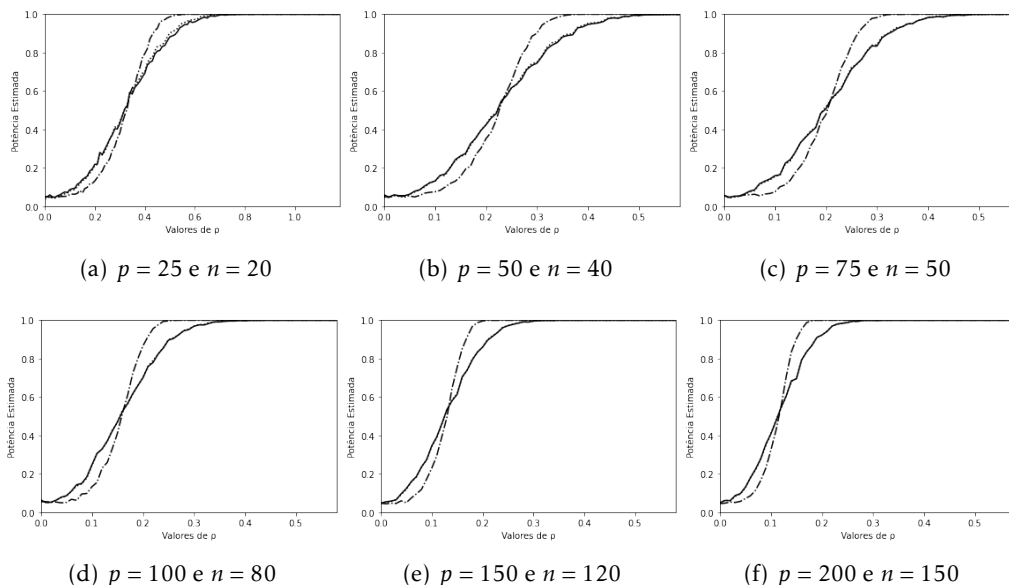


Quando para as simulações se usou a matriz de covariâncias de estrutura autoregressiva, (4.2), por a mesma ser contruída à custa de potências de ρ , não foi possível considerar para ρ valores negativos, pois isso originaria covariâncias positivas e negativas o que não cumpre com os pressuposto sob os quais é possível aplicar o teste de redução univariada. Ao analisarmos os resultados obtidos na Figura 4.3, é imediatamente notório que o teste de redução univariada, tanto no caso das variâncias serem conhecidas como no caso de serem desconhecidas, apenas apresenta melhores resultados que os testes de razão de verossimilhança (Anderson, 2003 [1]) e de Schoot (Schoot, 2005 [19]) quando ρ é pequeno. No entanto, à medida que ρ aumenta, apesar dos resultados obtidos para o teste de razão de verossimilhanças (Anderson, 2003 [1]) e para o teste de Schoot (Schoot, 2005 [19]) serem melhores, a diferença existente para os resultados obtidos com o teste de redução univariada não é muito significativa. Assim, podemos continuar a considerar que os resultados obtidos nas simulações foram bastante bons.

Quando temos um maior número de variáveis, p , que o tamanho da amostra n , e foram simuladas amostras geradas a partir de uma população Normal Multivariada cuja estrutura de covariâncias é autorregressiva, (4.2), podemos verificar na Figura 4.4 que o comportamento é bastante idêntico ao da Figura 4.3. Assim, das Figuras 4.4 e 4.3, também podemos concluir que os resultados obtidos com as simulações são bastante bons.

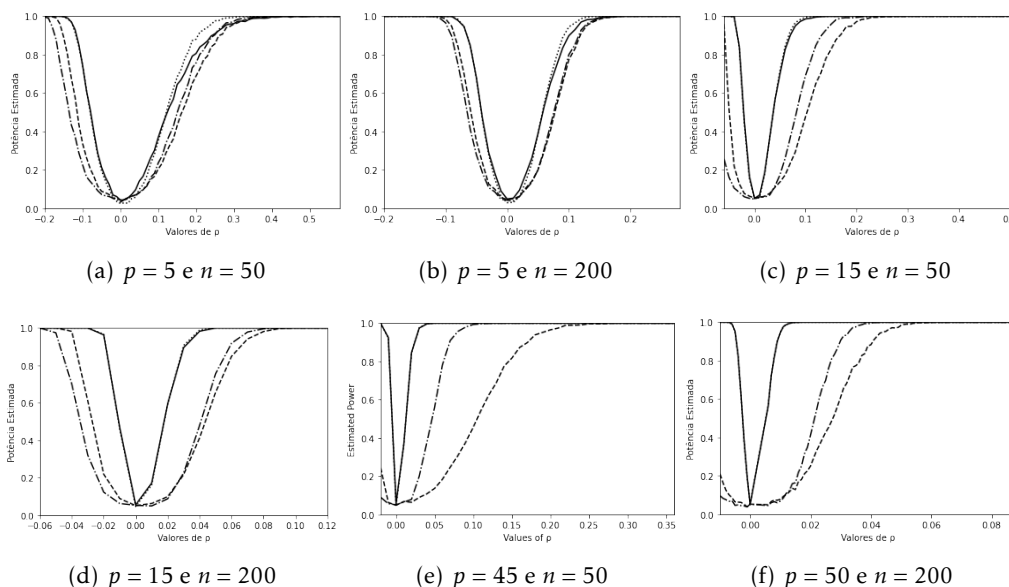
Por último, foram obtidos gráficos com os resultados das potências dos testes das simulações, para as amostras geradas considerando para a população a matriz de covariâncias cuja estrutura é circular, ou seja, a estrutura apresentada em (4.3). Estes resultados podem ser consultados nas Figuras 4.5 e 4.6 para os casos em que o número de variáveis, p , é menor que o tamanho da amostra, n , e em que o número de variáveis, p , é maior que

Figura 4.4: Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verossimilhanças com $p > n$ para a matriz de covariâncias de estrutura autorregressiva



o tamanho da amostra gerada, n , respetivamente.

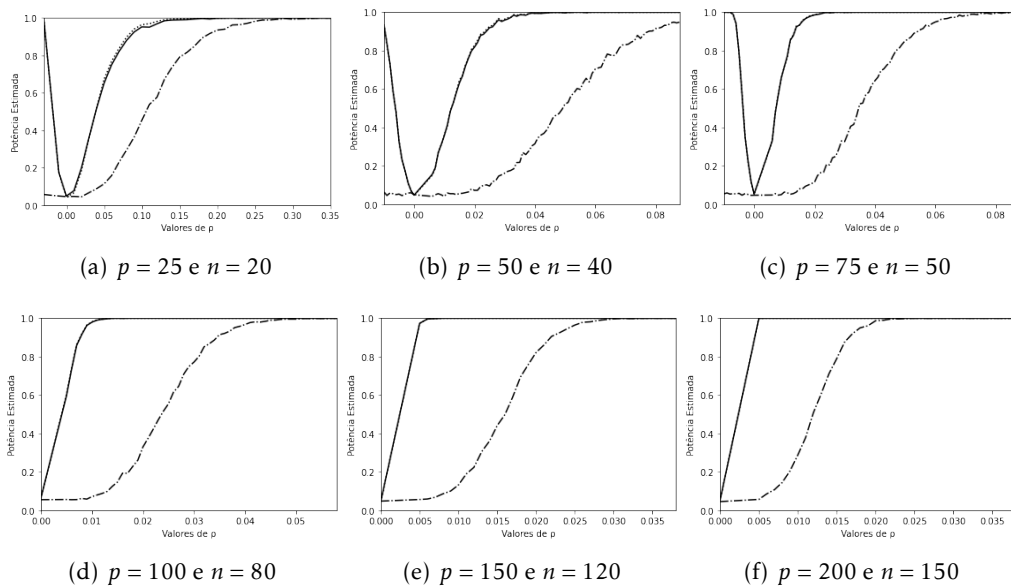
Figura 4.5: Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verossimilhanças com $p < n$ para a matriz de covariâncias de estrutura circular



Ao observarmos a Figura 4.5 é notório que quando a estrutura de covariâncias considerada para obter as amostras é a estrutura circular (4.3), os resultados obtidos para as potências dos testes são significativamente melhores quando usados os testes de redução

univariada, independentemente de se considerar as variâncias conhecidas ou desconhecidas, quando comparados com os resultados obtidos para as potências dos testes de razão de verosimilhanças (Anderson, 2003 [1]), e Schoot (Schoot, 2005 [19]). Tal como acontece quando se considera a estrutura simétrica composta, (4.1), cujos resultados são apresentados nas Figuras 4.1 e 4.2, conforme ρ se afasta de 0, as potências dos testes sobem muito mais rapidamente para 1 no caso dos testes de redução univariada do que nos restantes casos.

Figura 4.6: Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verosimilhanças com $p > n$ para a matriz de covariâncias de estrutura circular



Se consultarmos a Figura 4.6 facilmente percebemos que, os resultados obtidos são bastante favoráveis ao teste de redução univariada, independentemente de se conhecerem ou não as variâncias, pois as linhas que representam cada um dos casos acabam por estar maioritariamente sobrepostas. Os resultados aqui obtidos mostram, uma vez mais, que as potências de teste obtidas nos testes de redução univariada apresentam resultados bastante melhores que as obtidas com o teste de Schoot (Schoot, 2005 [19]).

4.2 Níveis de significância e potências dos testes

Nas Figuras apresentadas na Secção 4.1 correspondentes aos resultados das potências dos testes obtidas das simulações, é notório que quando a hipótese nula é verdadeira, a potência do teste está bastante aproximada do nível de significância do mesmo, para qualquer um dos casos considerados. Assim, nesta secção, vamos analisar com mais pormenor essa propriedade. Para isso foram simuladas 2000 amostras aleatórias provenientes de uma população Normal Multivariada de vetor de médias $\underline{\mu} = \underline{0}$ e matriz de covariâncias I_p , isto é, matriz de identidade de ordem p , para diferentes valores de p , número de variáveis,

e n , tamanho da amostra, de forma a obter estimativas dos níveis de significância. Estas simulações foram feitas para os níveis de significância $\alpha = 0.05$ e $\alpha = 0.01$, Tabelas 4.2 e 4.3 respectivamente. Destas simulações obtivemos as estimativas dos níveis de significância do teste de redução univariada para as variâncias desconhecidas. Estes resultados podem ser comparados com os das Tabelas 4.4 e 4.5 que foram retiradas de Schoot (2005) [19] e que apresentam, respectivamente, os mesmos resultados para o teste de Schoot (Schoot, 2005 [19]) com nível de significância $\alpha = 0.05$ e para o teste de razão de verossimilhanças (Anderson (2003 [1]) também para o nível de significância $\alpha = 0.05$.

| | $n = 4$ | $n = 8$ | $n = 16$ | $n = 32$ | $n = 64$ | $n = 128$ | $n = 256$ |
|-----------|---------|---------|----------|----------|----------|-----------|-----------|
| $p = 4$ | 0.029 | 0.022 | 0.023 | 0.022 | 0.023 | 0.022 | 0.025 |
| $p = 8$ | 0.032 | 0.037 | 0.035 | 0.040 | 0.038 | 0.037 | 0.039 |
| $p = 16$ | 0.039 | 0.040 | 0.041 | 0.040 | 0.045 | 0.040 | 0.041 |
| $p = 32$ | 0.044 | 0.044 | 0.043 | 0.048 | 0.047 | 0.047 | 0.050 |
| $p = 64$ | 0.045 | 0.047 | 0.043 | 0.049 | 0.051 | 0.046 | 0.051 |
| $p = 128$ | 0.046 | 0.048 | 0.050 | 0.048 | 0.046 | 0.049 | 0.046 |
| $p = 256$ | 0.045 | 0.043 | 0.043 | 0.045 | 0.054 | 0.052 | 0.045 |

Tabela 4.2: Resultados das estimativas do nível de significância do teste de redução univariada para $\alpha = 0.05$ e variâncias desconhecidas

Ao observarmos as Tabelas 4.2 e 4.3 é possível verificar que os valores são um pouco mais baixos do que os valores que pretendíamos obter, 0.05 e 0.01, respectivamente. No entanto, à medida que n e p aumentam os valores para as estimativas de significância dos testes também aumenta, aproximando-se dos valores usados nos testes. Quando comparamos os valores obtidos para $n > p$ com os obtidos para $n < p$, existe uma melhor aproximação das estimativas para o nível de significância ao valor usado para significância do teste quando $n < p$, apesar da diferença entre os dois casos não ser relevante.

| | $n = 4$ | $n = 8$ | $n = 16$ | $n = 32$ | $n = 64$ | $n = 128$ | $n = 256$ |
|-----------|---------|---------|----------|----------|----------|-----------|-----------|
| $p = 4$ | 0.004 | 0.007 | 0.005 | 0.003 | 0.003 | 0.002 | 0.003 |
| $p = 8$ | 0.007 | 0.008 | 0.005 | 0.007 | 0.006 | 0.006 | 0.005 |
| $p = 16$ | 0.007 | 0.009 | 0.012 | 0.008 | 0.007 | 0.008 | 0.009 |
| $p = 32$ | 0.007 | 0.011 | 0.010 | 0.007 | 0.009 | 0.010 | 0.010 |
| $p = 64$ | 0.008 | 0.009 | 0.009 | 0.009 | 0.009 | 0.010 | 0.010 |
| $p = 128$ | 0.009 | 0.010 | 0.008 | 0.008 | 0.010 | 0.008 | 0.010 |
| $p = 256$ | 0.009 | 0.009 | 0.012 | 0.009 | 0.013 | 0.010 | 0.008 |

Tabela 4.3: Resultados das estimativas do nível de significância do teste de redução univariada para $\alpha = 0.01$ e variâncias desconhecidas

Quando comparamos os resultados simulados da Tabela 4.2 com os resultados que Schoot (2005) [19] obteve nas Tabelas 4.4 e 4.5, é possível de verificar que quando usado o teste de Schoot (Schoot, 2005 [19]) ocorre o oposto do que ocorre quando usado o teste de redução univariada para a variâncias desconhecidas, isto é, maioritariamente as estimativas do nível de significância obtidas são um pouco superiores ao valor desejado de

0.05 e além disso as aproximações ao nível de significância usado parecem ser melhores para $n < p$ do que para $n > p$.

| | $n = 4$ | $n = 8$ | $n = 16$ | $n = 32$ | $n = 64$ | $n = 128$ | $n = 256$ |
|-----------|---------|---------|----------|----------|----------|-----------|-----------|
| $p = 4$ | 0.062 | 0.063 | 0.066 | 0.069 | 0.072 | 0.071 | 0.071 |
| $p = 8$ | 0.062 | 0.061 | 0.060 | 0.060 | 0.059 | 0.065 | 0.062 |
| $p = 16$ | 0.065 | 0.060 | 0.055 | 0.060 | 0.057 | 0.056 | 0.055 |
| $p = 32$ | 0.066 | 0.060 | 0.060 | 0.056 | 0.050 | 0.054 | 0.055 |
| $p = 64$ | 0.072 | 0.054 | 0.050 | 0.056 | 0.051 | 0.046 | 0.057 |
| $p = 128$ | 0.065 | 0.068 | 0.058 | 0.052 | 0.050 | 0.050 | 0.050 |
| $p = 256$ | 0.065 | 0.055 | 0.054 | 0.057 | 0.047 | 0.047 | 0.056 |

Tabela 4.4: Resultados das estimativas do nível de significância do teste de Schoot (Schoot, 2005 [19]) para $\alpha = 0.05$, valores retirados da Tabela 1 de Schoot (2005) [19]

Por outro lado, quando comparamos os resultados do teste de redução univariada de variâncias desconhecidas, Tabela 4.2, com os resultados obtidos para o teste de razão de verossimilhanças (Anderson (2003 [1]), Tabela 4.5, é possível de verificar que para este último conforme n e p aumentam as estimativa para a significância dos testes vão se afastando do valor que se desejava ter, sendo que à medida que p e n aumentam e p se aproxima de n os resultados observados tendem para 1, o que não é uma boa propriedade para um teste de hipóteses pois mostra que este não é centrado.

| | $n = 4$ | $n = 8$ | $n = 16$ | $n = 32$ | $n = 64$ | $n = 128$ | $n = 256$ |
|-----------|---------|---------|----------|----------|----------|-----------|-----------|
| $p = 4$ | 0.132 | 0.060 | 0.053 | 0.052 | 0.050 | 0.054 | 0.052 |
| $p = 8$ | | 0.421 | 0.084 | 0.058 | 0.056 | 0.055 | 0.050 |
| $p = 16$ | | | 0.722 | 0.083 | 0.055 | 0.052 | 0.048 |
| $p = 32$ | | | | 0.989 | 0.135 | 0.060 | 0.056 |
| $p = 64$ | | | | | 0.995 | 0.192 | 0.068 |
| $p = 128$ | | | | | | 1.000 | 0.465 |
| $p = 256$ | | | | | | | 1.000 |

Tabela 4.5: Resultados das estimativas do nível de significância do teste de razão de verossimilhanças para $\alpha = 0.05$, valores retirados da Tabela 2 de Schoot (2005) [19]

De forma a dar uma melhor visualização sobre a robustez do teste de redução univariada, foram ainda realizadas simulações para obter a potência de teste do mesmo para as variâncias desconhecidas, partindo de uma população Normal Multivariada de vetor de médias $\underline{\mu} = \underline{0}$ e matriz de covariâncias com estrutura simétrica composta, (4.1), com $\rho = 0.1$ e $\sigma^2 = 1$. Os resultados destas simulações podem ser consultados na Tabela 4.6, e foram realizadas com o intuito de ser possível comparar com as simulações feitas por Schoot (2005) [19] e cujos resultados podem ser consultados nas Tabelas 4.7 e 4.8 para os testes de Schoot (Schoot, 2005 [19]) e de razão de verossimilhanças (Anderson (2003 [1]), respetivamente.

Quando comparamos os resultados na Tabela 4.6, teste de redução univariada para as variâncias desconhecidas, e na Tabela 4.7, teste de Schoot (Schoot, 2005 [19]), é notório que para o primeiro caso, os valores são maioritariamente melhores que os valores do 2ª

| | $n = 4$ | $n = 8$ | $n = 16$ | $n = 32$ | $n = 64$ | $n = 128$ | $n = 256$ |
|-----------|---------|---------|----------|----------|----------|-----------|-----------|
| $p = 4$ | 0.002 | 0.001 | 0.005 | 0.023 | 0.102 | 0.300 | 0.724 |
| $p = 8$ | 0.005 | 0.028 | 0.132 | 0.346 | 0.734 | 0.976 | 1.000 |
| $p = 16$ | 0.078 | 0.259 | 0.572 | 0.894 | 0.997 | 1.000 | 1.000 |
| $p = 32$ | 0.335 | 0.661 | 0.941 | 0.998 | 1.000 | 1.000 | 1.000 |
| $p = 64$ | 0.647 | 0.922 | 0.997 | 1.000 | 1.000 | 1.000 | 1.000 |
| $p = 128$ | 0.870 | 0.987 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| $p = 256$ | 0.961 | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Tabela 4.6: Potências estimadas do teste de redução univariada para as variâncias desconhecidas com base em amostras Normais Multivariadas de vetor de médias $\underline{\mu} = \underline{0}$ e matriz de covariâncias de estrutura simétrica composta (4.1) com $\sigma^2 = 1$ e $\rho = 0.1$ e para $\alpha = 0.05$

caso, sendo as exceções apenas para n e p muito pequenos. Outra característica que não passa despercebida é o facto de há medida que n e p aumentam as potências dos testes aumentam significativamente.

| | $n = 4$ | $n = 8$ | $n = 16$ | $n = 32$ | $n = 64$ | $n = 128$ | $n = 256$ |
|-----------|---------|---------|----------|----------|----------|-----------|-----------|
| $p = 4$ | 0.076 | 0.087 | 0.123 | 0.172 | 0.307 | 0.534 | 0.845 |
| $p = 8$ | 0.079 | 0.101 | 0.177 | 0.313 | 0.597 | 0.903 | 0.998 |
| $p = 16$ | 0.112 | 0.166 | 0.310 | 0.595 | 0.904 | 0.997 | 1.000 |
| $p = 32$ | 0.161 | 0.285 | 0.557 | 0.871 | 0.996 | 1.000 | 1.000 |
| $p = 64$ | 0.255 | 0.486 | 0.797 | 0.987 | 1.000 | 1.000 | 1.000 |
| $p = 128$ | 0.375 | 0.698 | 0.946 | 1.000 | 1.000 | 1.000 | 1.000 |
| $p = 256$ | 0.542 | 0.846 | 0.990 | 1.000 | 1.000 | 1.000 | 1.000 |

Tabela 4.7: Potências estimadas do teste de Schoot (Schoot, 2005 [19]) com base em amostras Normais Multivariadas de vetor de médias $\underline{\mu} = \underline{0}$ e matriz de covariâncias de estrutura simétrica composta (4.1) com $\sigma^2 = 1$ e $\rho = 0.1$ e para $\alpha = 0.05$, resultados retirados da Tabela 3 de Schoot (2005) [19]

No que diz respeito aos resultados obtidos com o teste de razão de verosimilhanças (Anderson (2003 [1]), Tabela 4.8, quando comparamos com os obtidos no teste de redução univariada, Tabela 4.6, os valores obtidos com este segundo são significativamente melhores que os obtidos com o primeiro, pois apesar de os resultados irem melhorando à medida que tanto p como n aumentam para o teste de razão de verosimilhanças (Anderson (2003 [1]), quando consideramos n muito maior que p os resultados são muito baixos, o que não acontece no teste de redução univariada.

4.3 Desvios da Normalidade

Para estudar o comportamento do teste de redução univariada nos casos em que a população não é proveniente de uma população Normal Multivariada foram realizadas simulações obtendo amostras de distribuições t-Student Multivariadas. A distribuição escolhida foi a distribuição t-Student Multivariada, pois, tal como a distribuição Normal Multivariada, é uma distribuição absolutamente contínua e simétrica, e na qual à medida

| | $n = 4$ | $n = 8$ | $n = 16$ | $n = 32$ | $n = 64$ | $n = 128$ | $n = 256$ |
|-----------|---------|---------|----------|----------|----------|-----------|-----------|
| $p = 4$ | 0.142 | 0.071 | 0.088 | 0.124 | 0.233 | 0.447 | 0.783 |
| $p = 8$ | | 0.445 | 0.154 | 0.224 | 0.446 | 0.817 | 0.993 |
| $p = 16$ | | | 0.793 | 0.370 | 0.705 | 0.982 | 1.000 |
| $p = 32$ | | | | 0.996 | 0.926 | 1.000 | 1.000 |
| $p = 64$ | | | | | 0.994 | 1.000 | 1.000 |
| $p = 128$ | | | | | | 1.000 | 1.000 |
| $p = 256$ | | | | | | | 1.000 |

Tabela 4.8: Potências estimadas do teste de razão de verosimilhanças com base em amostras Normais Multivariadas de vetor de médias $\underline{\mu} = \underline{0}$ e matriz de covariâncias de estrutura simétrica composta (4.1) com $\sigma^2 = 1$ e $\rho = 0.1$ e para $\alpha = 0.05$, resultados retirados da Tabela 4 de Schoot (2005) [19]

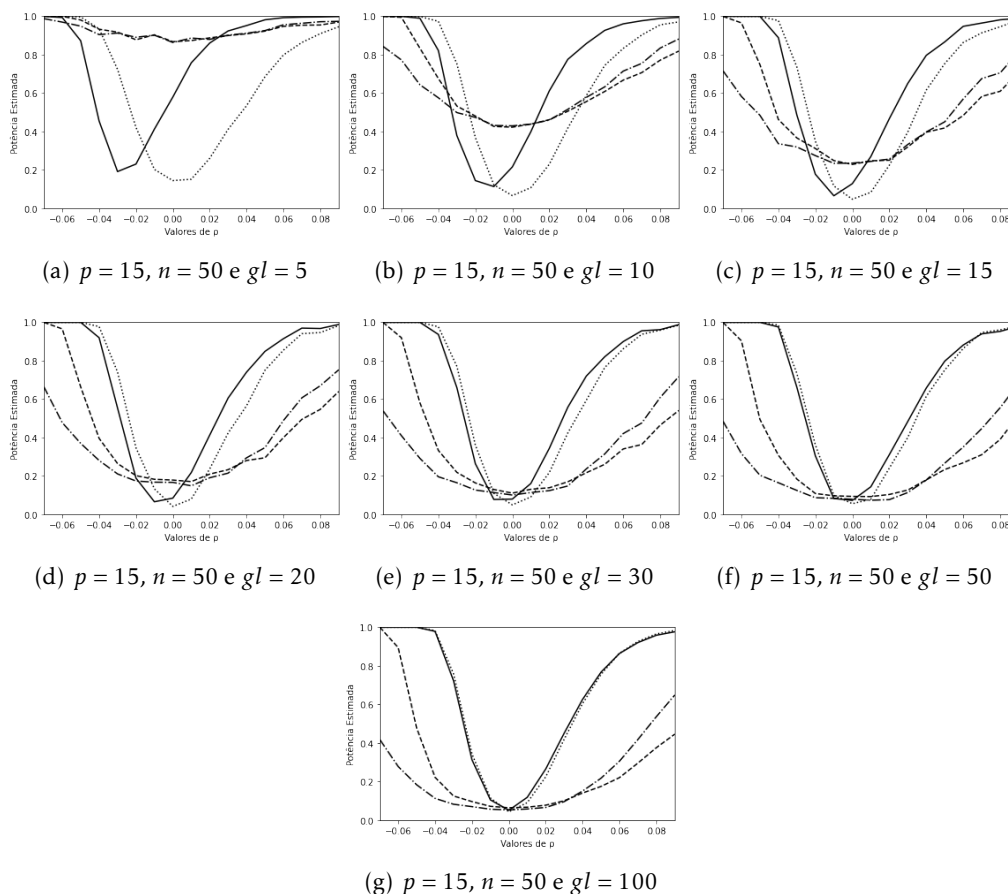
que os graus de liberdade aumentam a mesma se aproxima de uma distribuição Normal Multivariada.

Como pretendemos compreender a eficácia do teste para diferentes possibilidades, aqui as simulações foram então realizadas para amostras cujo número de variáveis, p , é inferior ao tamanho da amostra, n , e também para um número de variáveis superior ao tamanho da amostra. Além disso, para cada, um destes dois casos, foram considerados diferentes graus de liberdade, gl , e ainda, tal como foi feito na Secção 4.1, foram consideradas matrizes de covariâncias para diferentes $\rho \in \underline{\rho} = (\rho_1, \dots, \rho_k)$, tendo se obtido $\underline{\rho}$ garantindo que ρ_1 é o menor valor para o qual a matriz é definida positiva e ρ_k é positivo, $\rho_1 < \dots < \rho_k$ e que $\exists! \rho_i = 0$ tal que $1 < i < k$. A metodologia aqui aplicada foi, portanto, a mesma que para a obtenção dos gráficos da Secção 4.1 e que vem explicada na Secção 4 sendo as únicas diferenças que aqui apenas se consideram duas combinações de n e p e para cada uma delas diferentes graus de liberdade gl .

Assim, na Figura 4.7 podemos consultar os resultados obtidos, para os diferentes graus de liberdade, gl , considerados para $n = 49$ e $p = 15$. Por outro lado, na Figura 4.7 temos os resultados obtidos quando $n = 40$ e $p = 50$ também para os vários graus de liberdade que foram considerados. Em ambos os casos, a linha a '-' são os resultados obtidos para as estimativas do teste de redução univariada de variâncias desconhecidas e a '...' são os resultados obtidos com o teste de redução univariada mas de variâncias desconhecidas, a linha '---' são as estimativas das potências do teste de razão de verosimilhanças (Anderson (2003 [1]) e a '-.' temos os resultados do teste de Schoot (Schoot, 2005 [19]).

Na Figura 4.7 é bastante notório que quanto menor os graus de liberdade, melhor são os resultados obtidos para o teste de redução univariada, quer as variâncias sejam conhecidas ou desconhecidas, quando comparados tanto com o teste de razão de verosimilhanças (Anderson (2003 [1]) como com o teste de Schoot (Schoot, 2005 [19]). Além disso, também se verifica que há medida que os graus de liberdade, gl , aumentam os resultados se aproximam dos obtidos quando as simulações foram realizadas considerando amostras provenientes de populações normais, sendo que este era o resultado esperado, pois como

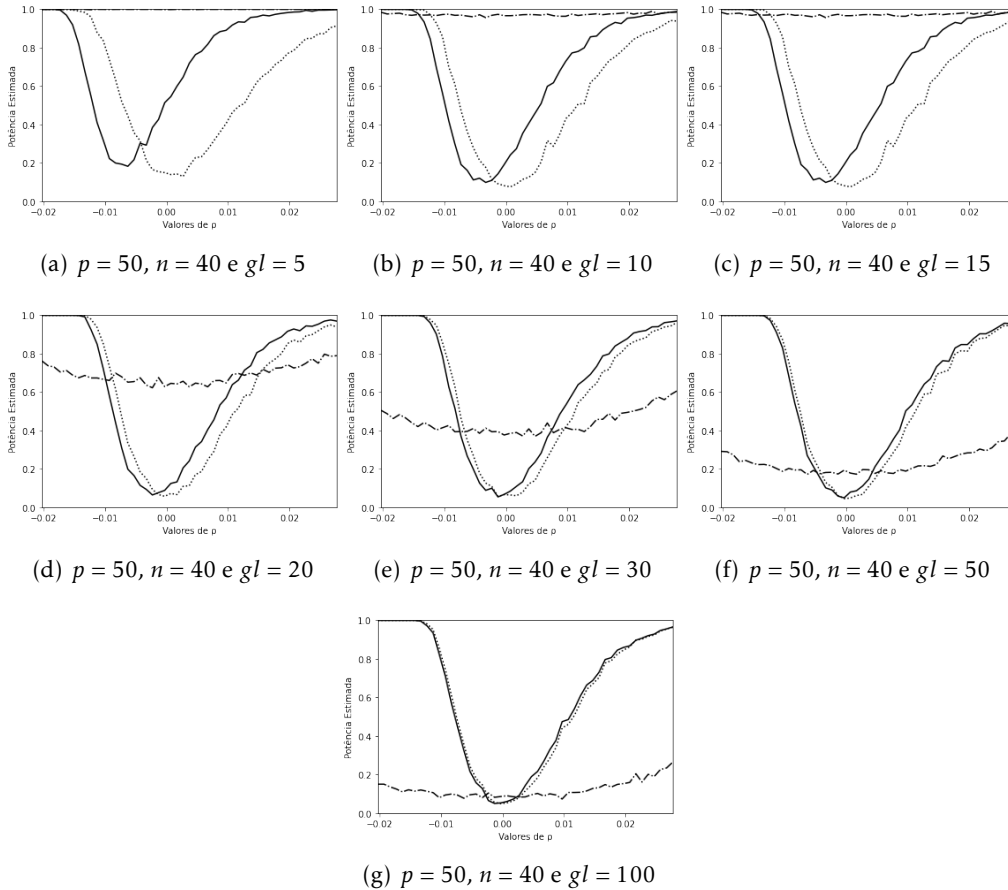
Figura 4.7: Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verossimilhanças para amostras de populações t-Student, com $p < n$ e para a matriz de covariâncias de estrutura simétrica composta



já foi referido à medida que os graus de liberdade aumentam a distribuição t-Student Multivariada aproxima-se da distribuição Normal Multivariada.

Relativamente ao teste de redução univariada, quer para variâncias conhecidas quer desconhecidas, quando a população a testar é proveniente de uma população t-Student Multivariada com um número de variáveis superior ao tamanho da amostra, com a análise da Figura 4.8, podemos concluir que o mesmo se comporta melhor que o teste de Schoot (Schoot, 2005 [19]) nestas situações, sendo a discrepância dos resultados mais significativa quanto menor forem os graus de liberdade. Tal como acontece para $n > p$, também neste caso, à medida que os graus de liberdade aumentam também os resultados mais se assemelham aos obtidos com amostras de populações Normais Multivariadas, o que novamente era de esperar.

Figura 4.8: Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas, teste de Schoot (Schoot, 2005 [19]) e teste de razão de verossimilhanças para amostras de populações t-Student, com $p < n$ e para a matriz de covariâncias de estrutura simétrica composta



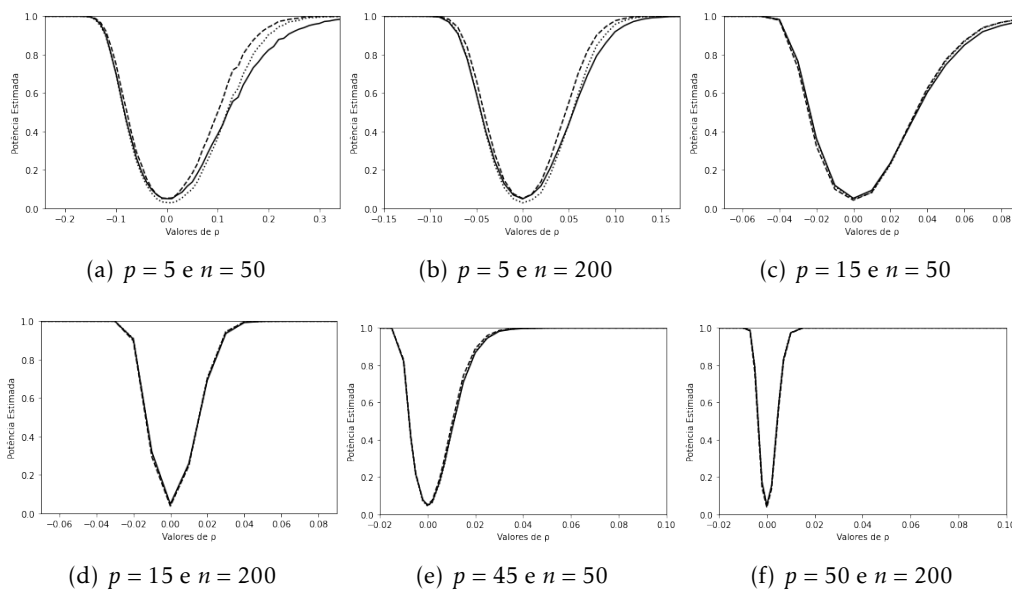
4.4 Testes de Esfericidade vs Equivariância-Equicorrelação

O teste de redução univariada pode ser usado para testar a hipótese nula no teste de esfericidade vs equivariância-equicorrelação, (3.4), sobre o qual falamos na Secção 3.2.1, pois o pressuposto das covariâncias serem todas não negativas ou não positivas está cumprido quer nos encontremos sob a hipótese nula, H_0 , ou sob a hipótese alternativa, H_1 .

Assim, para analisar de que forma o teste de redução univariada se comporta nesta situação, foram feitas simulações. A metodologia aqui aplicada foi bastante semelhante à aplicada nas secções 4.1 e 4.3. Ou seja, para algumas combinações de p variáveis e tamanho da amostra n foram geradas amostras de populações Normais Multivariadas de vetor de médias $\underline{\mu} = \underline{0}$ e matriz de covariâncias dada por uma estrutura simétrica composta (4.1). As simulações foram repetidas para diferentes valores de ρ , valor do qual depende a matriz de covariâncias, usando $\underline{\rho} = (\rho_1, \dots, \rho_k)$, com $k \in \mathbb{N}$, sendo que $\underline{\rho}$ foi calculado garantindo as mesmas condições que nas secções anteriores, isto é, garantindo

que $\rho_i \in \mathbb{R}$, $\rho_1 < \dots < \rho_k$, ρ_1 é o menor valor para o qual a matriz é definida positiva e que ρ_k é suficientemente grande para garantir que a potência estimada do teste de redução univariada quando se assume ρ_k é aproximadamente 1. Aqui, além de obtermos as potências estimadas para o teste de redução univariada, de variâncias conhecidas e desconhecidas, também se calcularam as potências estimadas para o teste proposto por Marques e Coelho (2015) [13] referido na Secção 3.2.1 e cuja estatística de teste é (3.5).

Figura 4.9: Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas para a matriz de covariâncias de estrutura de simétrica composta e do teste de razão de verossimilhanças dado por Marques e Coelho (2015) [13]

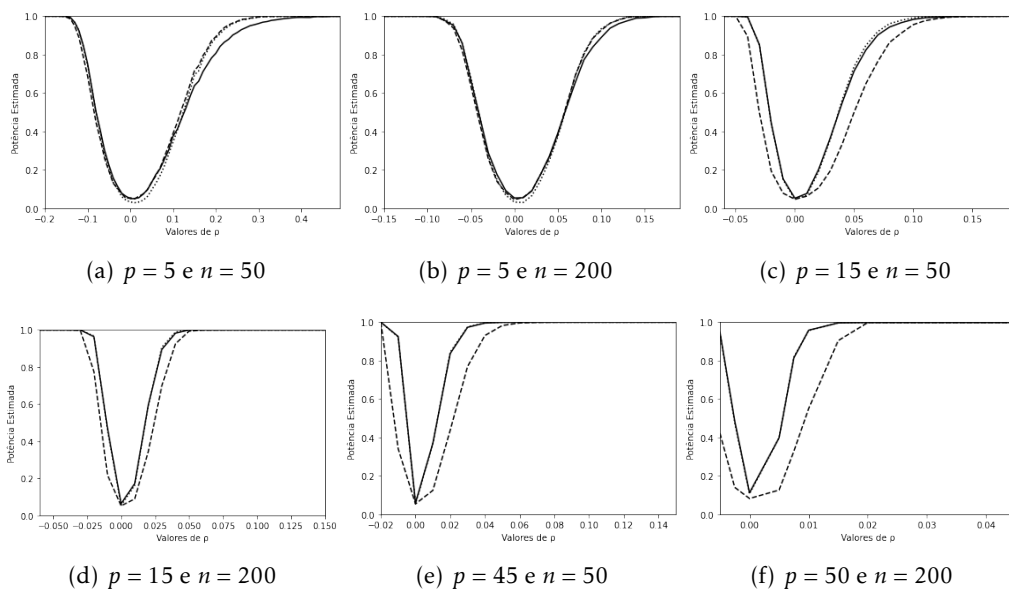


Nos Gráficos 4.9 podemos consultar os resultados obtidos com estas simulações, sendo que a linha a '—' são os resultados para o teste de redução univariada de variâncias conhecidas, a '⋯' são os resultados para o teste de redução univariada de variâncias desconhecidas e a '—' são os resultados para o teste proposto por Marques e Coelho (2015) [13]. Através destes resultados podemos concluir que independentemente do teste usado, os resultados são bastante semelhantes, pois apenas se notam algumas diferenças pouco significativas nos casos de $p = 5$ e $n = 50$ (4.9(a)) e $p = 5$ e $n = 200$ (4.10(b)), tendo os restantes casos as 3 linhas bastante sobrepostas. Neste caso, apesar de as diferenças obtidas com o teste de redução univariada, independentemente das variâncias serem ou não conhecidas, não serem significativamente diferentes das obtidas quando se usa a estatística de teste de razão de verossimilhanças (3.5), é de notar, que no caso dos testes de redução univariada, a estatística de teste é baseada numa distribuição χ^2 que é muito mais fácil de usar do que a distribuição de (3.5), além disso, o teste de redução univariada é possível de usar nos casos em que $p > n$ enquanto que o teste de razão de verossimilhanças não.

4.5 Testes de Esfericidade vs Circularidade

Outro teste para o qual é possível utilizar o teste dado na Secção 3.4, o teste de redução univariada, é no teste de esfericidade vs circularidade, sobre o qual falamos na Secção 3.2.2 e que foi apresentado por Olkin e Press (1969) [17]. Analogamente à Secção 4.4, foram simuladas amostras que permitiram estudar a utilização do teste de redução univariada, para as variâncias conhecidas e desconhecidas, quando comparada com o teste (3.8) cuja estatística de teste é (3.9) (Olkin e Press, 1969 [17]). A metodologia usada para estas simulações, foi a mesma utilizada na Secção 4.4.

Figura 4.10: Gráficos das potências estimadas dos testes de redução univariada de variâncias conhecidas e desconhecidas para a matriz de covariâncias de estrutura de circular e do teste de razão de verosimilhanças dado por Olkin e Press (1969) [17]



O Gráfico 4.10 tem os resultados das potências de teste estimadas nas simulações para ‘-’ e ‘...’ os resultados dos testes de redução univariada de variâncias conhecidas e desconhecidas, respetivamente, e ‘- - -’ os resultados do teste proposto por Olkin e Press (1969) [17]. Aqui é possível de constatar que, à excepção de quando o número de variáveis, p , é muito pequeno, qualquer um dos testes de redução univariada se comporta um pouco melhor que de razão de verosimilhanças dado em Olkin e Press (1969) [17]. Aliado a isso, uma vez mais, o teste de redução univariada é mais fácil de utilizar por se tratar de uma estatística de teste de distribuição χ^2 , além de ser possível de utilizar quando $p > n$.

CONCLUSÕES

Ao longo deste trabalho foram feitas diversas simulações para estudar de que forma os testes de redução univariada, para as variâncias conhecidas e desconhecidas, se comportam, tendo sido considerados diferentes cenários. Para provar a robustez dos resultados quando mudam as condições iniciais da população, as simulações foram feitas considerando amostras de diferentes tamanhos de populações Normais Multivariadas com diferentes dimensões, para diferentes estruturas de covariâncias, considerando casos em que as variáveis são independentes e casos em que não são. Além de tudo isto, ainda se testou de que forma o teste se comporta quando as amostras não são provenientes de populações Normais Multivariadas, tendo sido usada uma população t-Student pelas suas características e também se estudou a utilização da estatística de teste proposta para o teste de esfericidade vs equivariância-equicorrelação e para o teste de esfericidade vs circularidade.

Os resultados obtidos em todas estas simulações foram bastante coerentes, mostrando que na maioria dos casos o teste de redução univariada comporta-se melhor que o teste de razão de verossimilhanças (Anderson (2003 [1]), quando este pode ser aplicado, e que o teste de Schoot (Schoot, 2005 [19]). Nos vários gráficos que tivemos oportunidade de analisar ao longo do trabalho, é notório, que o teste de redução univariada é mais sensível a covariâncias baixas entre variáveis que os restantes testes usados nas simulações. A única exceção, ou seja, o único caso em que o teste de redução univariada não apresenta melhores resultados que os testes de razão de verossimilhanças (Anderson (2003 [1]) e de Schoot (Schoot, 2005 [19]), é quando se considera para matriz de covariâncias uma estrutura autorregressiva. No entanto, neste caso, apesar de o teste de redução univariada não se destacar dos restantes, apresenta resultados muito semelhantes, sendo que, por se tratar de um teste de aplicação muito simples e cuja distribuição da estatística é fácil de utilizar, continua a ser vantajoso usar o mesmo.

No que diz respeito aos desvios da normalidade da população, apesar dos resultados obtidos para o teste de redução univariada não serem os ideais, quando comparados com os outros testes, têm um comportamento um pouco melhor.

Quanto à utilização da estatística do teste de redução univariada para testar a esfericidade vs a equivariância-equicorrelação e para testar a esfericidade vs a circularidade, os resultados obtidos mantêm a coerência. Ou seja, quando usada a estatística de teste de redução univariada obtivemos resultados melhores do que quando usadas as restantes estatísticas. É ainda de salientar que o teste de redução univariada pode ser utilizada com outras estruturas, desde que se verifique o pressuposto de que as covariâncias são todas não positivas ou não negativas.

Assim, é possível de concluir que o teste proposto, o teste de redução univariada, apresenta várias vantagens relativamente a outros testes já conhecidos, sendo algumas delas, a facilidade de aplicação do mesmo, a facilidade de manuseamento da distribuição da estatística de teste e a robustez dos resultados obtidos quando se utiliza esta estatística de teste.

Como trabalhos futuros, pretendemos explorar e analisar a forma mais adequada de aplicar o teste desenvolvido a problemas da vida real, principalmente em contextos de alta dimensionalidade. Adicionalmente, temos como objetivo desenvolver um método semelhante para testes sobre matrizes de covariância bloco diagonais, onde inclui o teste de independência de grupos de variáveis e para outro tipo de estruturas de covariância usuais em contexto multivariado.

BIBLIOGRAFIA

- [1] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. Third. John Wiley Sons, 2003. ISBN: 0-471-36091-0 (ver pp. 1, 9, 15, 21, 24–26, 28, 30–34, 39).
- [2] G. E. BOX. “A general distribution theory for a class of likelihood criteria”. Em: *Biometrika* 36.3-4 (1949), pp. 317–346. DOI: 10.1093/biomet/36.3-4.317 (ver pp. 1, 12).
- [3] R. W. Butler, S. Huzurbazar e J. G. Booth. “Saddlepoint Approximations for Tests of Block Independence, Sphericity and Equal Variances and Covariances”. Em: *Journal of the Royal Statistical Society: Series B (Methodological)* 55.1 (1993), pp. 171–183. DOI: 10.1093/biomet/36.3-4.317 (ver pp. 1, 12).
- [4] C. A. Coelho. “Likelihood Ratio Tests for Elaborate Covariance Structures and for MANOVA Models with Elaborate Covariance Structures—A Review”. Em: *Journal of the Indian Institute of Science* 1-39 (2022) (ver p. 2).
- [5] C. A. Coelho. “Likelihood Ratio Tests for Elaborate Covariance Structures and for MANOVA Models with Elaborate Covariance Structures—A Review.” Em: *Gervasi O, Murgante B, Misra S, Borruso G, Torre CM, Rocha AMAC, Taniar D, Apduhan BO, Stankova E, Cuzzocrea A (eds) Computational Science and Its Applications. Lecture Notes in Computer Science* 10408 5 (2017), pp. 20–32 (ver p. 2).
- [6] C. A. Coelho. “The generalized near-integer Gamma distribution: A basis for ‘near-exact’ approximations to the distribution of statistics which are the product of an odd number of independent Beta random variables”. Em: *Journal of Multivariate Analysis* 89.2 (2004), pp. 191–218. DOI: 10.1016/j.jmva.2003.12.001 (ver pp. 1, 12).
- [7] C. A. Coelho e F. J. Marques. “Near-exact distributions for the independence and sphericity likelihood ratio test statistics”. Em: *Journal of Multivariate Analysis* 101.3 (2010), pp. 583–593. DOI: 10.1016/j.jmva.2009.09.012 (ver pp. 1, 12, 15).
- [8] R. M. Gray. “The sphericity versus equivariance-euicorrelation test”. Em: *Foundations and Trends® in Communications and Information Theory* 2.3 (2006), pp. 155–239 (ver p. 17).

- [9] O. Ledoit e M. Wolf. “Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size”. Em: *The Annals of Statistics* 30.4 (2002), pp. 1081–1102 (ver p. 2).
- [10] W. Li e Z. Liu. “A test for the complete independence of high-dimensional random vectors”. Em: *Journal of Statistical Computation and Simulation* 86.16 (2016), pp. 3135–3140 (ver p. 2).
- [11] F. J. Marques e C. A. Coelho. “Near-exact distributions for the sphericity likelihood ratio test statistic”. Em: *Journal of Statistical Planning and Inference* 138 (2008), pp. 726–741 (ver p. 15).
- [12] F. J. Marques e C. A. Coelho. “Testing simultaneously different covariance block diagonal structures – the multi-sample case”. Em: *Journal of Applied Statistics* 47.13-15 (2020), pp. 2765–2784 (ver p. 1).
- [13] F. J. Marques e C. A. Coelho. “The sphericity versus equivariance-euicorrelation test”. Em: *AIP Conference Proceedings* 1648.1 (2015) (ver pp. 15, 16, 37).
- [14] F. J. Marques, C. A. Coelho e B. C. Arnold. “A general near-exact distribution theory for the most common likelihood ratio test statistics used in Multivariate Analysis”. Em: *Test* 20 (2011), pp. 180–203 (ver p. 1).
- [15] J. W. Mauchly. “Significance Test for Sphericity of a Normal n-Variate Distribution”. Em: *The Annals of Mathematical Statistics* 11.2 (1940), pp. 204–209. DOI: 10.1214/aoms/1177731915 (ver p. 15).
- [16] G. S. Mudholkar, M. C. Trivedi e C. T. Lin. “An Approximation to the Distribution of the Likelihood Ratio Statistic for Testing Complete Independence”. Em: *Technometrics* 24.2 (1982), pp. 139–143 (ver p. 1).
- [17] I. Olkin e S. J. Press. “Testing and Estimation for a Circular Stationary Model”. Em: *The Annals of Mathematical Statistics* 40.4 (1969), pp. 1358–1373. DOI: 10.1214/aoms/1177697508 (ver pp. 15, 17, 38).
- [18] A. C. Rencher e W. F. Christensen. *Methods of Multivariate Analysis*. Third. John Wiley Sons, 2012. ISBN: 978-0-470-17896-6 (ver pp. 10, 12–14).
- [19] J. R. Schoolt. “Testing for Complete Independence in High Dimensions”. Em: *Biometrika* 92.4 (2005), pp. 951–956. DOI: 10.1093/biomet/92.4.9513 (ver pp. xi, xiii, 2, 17, 18, 24–36, 39).
- [20] M. S. Srivastava. “Some Tests Concerning the Covariance Matrix in High Dimensional Data”. Em: *Journal of the Japan Statistical Society* 35.2 (2005), pp. 251–272 (ver p. 2).
- [21] S. S. Wilks. “Sample Criteria for Testing Equality of Means, Equality of Variances, and Equality of Covariances in a Normal Multivariate Distribution”. Em: *The Annals of Mathematical Statistics* 17.3 (1946), pp. 257–281. DOI: 10.1214/aoms/1177730940 (ver p. 15).

- [22] S. S. Wilks. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. Em: *The Annals of Mathematical Statistics* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360 (ver pp. 1, 12).

ANEXO 1: CÓDIGO PYTHON UTILIZADO NAS SIMULAÇÕES

Nesta secção apresentamos o código Python utilizado nas simulações que foram realizadas utilizando a plataforma..

I.1 Código usado para a estrutura de covariância simétrica composta

De seguida é possível consultar o código utilizado nas simulações realizadas usando a estrutura de covariâncias simétrica composta cujos resultados foram apresentados na secção 4.1. Notemos, que apenas está representado o código utilizado para quando o número de variáveis considerado foi $p = 5$ e se geraram amostras de tamanho $n = 50$, no entanto para os restantes casos utilizou-se o mesmo, apenas substituindo os valores de n e p e calculando o vetor vec de forma a cumprir com os requisitos mencionados na secção 4.1, considerando que este vetor é o vetor $\underline{\rho}$ que contém os diferentes valores de ρ utilizados nas simulações.

```
import numpy as np
from scipy.stats import chi2
from scipy.stats import norm
from scipy.linalg import helmert
from math import *
import matplotlib.pyplot as plt

def is_pos_def(x):
    return np.all(np.linalg.eigvals(x) > 0)

np.random.seed(10)
nt = 10000
NN = 50
```

```
n = NN-1
p = 5

mu = [0 for j in range(1,p+1)]
veca = [0 for j in range(1,nt+1)]
for i in range(1,nt):
    x = np.random.multivariate_normal(mu, np.identity(p), NN).T
    vecS=(NN-1)*np.cov(x)
    veca[i] =np.linalg.det(vecS)/np.prod(np.diagonal(vecS))
quantil=np.quantile(veca,0.05)

vec=[j/10000 for j in range(-5000,1,1)]
for gamma in vec:
    Cov = np.dot(gamma,([[1] * p] * p))\
    + np.dot((1-gamma),np.identity(p))
    if is_pos_def(Cov):
        print(gamma)
        break
value= round(gamma,4)+0.0005
limit= 30
vec = [value + (j-1)/100 for j in range(1,2*limit)]

Config1 = [0 for j in range(1,len(vec)+1)]
Config2 = [0 for j in range(1,len(vec)+1)]
LRsimfig = [0 for j in range(1,len(vec)+1)]
sco = [0 for j in range(1,len(vec)+1)]
count = 0
ns = 2000
for gamma in vec:
    count = count + 1
    Cov = np.dot(gamma,([[1] * p] * p))\
    + np.dot((1-gamma),np.identity(p))
    Con1 = [0 for j in range(1,ns+1)]
    Con2 = [0 for j in range(1,ns+1)]
    Con3 = [0 for j in range(1,ns+1)]
    LRsim = [0 for j in range(1,ns+1)]
    for j in range(1,ns):
        Data = np.random.multivariate_normal(mu, Cov, NN).T
        vecS = (NN-1)*np.cov(Data)
        LR = np.linalg.det(vecS)/np.prod(np.diagonal(vecS))
```

```

if LR < quantil: LRsim[j-1] = 1
sig = np.diagonal(np.cov(Data))
datan1 = Data.sum(axis=0)
if n*np.var(datan1)/sum(sig) > chi2.ppf(0.975,n) \
or n*np.var(datan1)/sum(sig) < chi2.ppf(0.025,n):
    Con1[j-1] = 1
if n*np.var(datan1)/sum(np.diagonal(Cov))\
    > chi2.ppf(0.975,n)\
    or n*np.var(datan1)/sum(np.diagonal(Cov))\
    < chi2.ppf(0.025,n):
    Con2[j-1] =1
vecSS = np.corrcoef(Data)
stat = (sum(sum([[vecSS[i][jj]**2\
                    for jj in range(0,i)]\
                    for i in range(1,p)],[]))\
        -(p*(p-1)/(2*n)))\
        /((p*(p-1)*(n-1)/((n**2)*(n+2)))*(1/2))
if stat > norm.ppf(0.975,0,1)\
    or stat < norm.ppf(0.025,0,1): Con3[j-1] = 1
sco[count-1] = sum(Con3)/ns
Config1[count-1] = sum(Con1)/ns
Config2[count-1] = sum(Con2)/ns
LRsimfig[count-1] = sum(LRsim)/ns

plt.plot(vec, Config2, label="\u03C3's known", \
         linestyle='-', color='black')
plt.plot(vec, Config1, label="\u03C3's unknown", \
         linestyle=':', color='black')
plt.plot(vec, LRsimfig, label="LRT", \
         linestyle='—', color='black')
plt.plot(vec, sco, label="SCH", \
         linestyle='-.', color='black')
plt.xlabel("Valores de \u03C1")
plt.ylabel("Potencia Estimada")
plt.ylim(0,1)
plt.xlim(vec[0], vec[len(vec)-1])
plt.show()

```

Notemos que no código apresentado acima, quando foram considerados valores para o número de variáveis p superiores ao tamanho da amostra n , não foi considerado para as simulações o teste de razão de verosimilhança, sendo que, por esse motivo, toda a parte

do código correspondente a esse teste foi omitida.

I.2 Código usado para a estrutura de covariância autorregressiva

Quando para as simulações se considerou a estrutura autorregressiva para a matriz de covariâncias o código usado foi o seguinte:

```
import numpy as np
from scipy.stats import chi2
from scipy.stats import norm
from scipy.linalg import helmert
from math import *
import matplotlib.pyplot as plt

def is_pos_def(x):
    return np.all(np.linalg.eigvals(x) > 0)

def matriz_cov(gamma, p):
    matriz = np.zeros((p,p))
    for i in range(0,p):
        matriz[i,i] = gamma ** 0
        for j in range(1,p-i):
            matriz[i,i+j] = gamma ** j
            matriz[i+j,i] = gamma ** j
    return(matriz)

np.random.seed(10)

nt = 10000

NN = 50
n = NN-1
p = 5

mu = [0 for j in range(1,p+1)]
veca = [0 for j in range(1,nt+1)]
for i in range(1,nt):
    x = np.random.multivariate_normal(mu, np.identity(p), NN).T
    vecS=(NN-1)*np.cov(x)
    veca[i] =np.linalg.det(vecS)/np.prod(np.diagonal(vecS))
```

```

quantil=np.quantile(veca,0.05)

value= 0
print(value)
limit= 30
vec = [value + (j-1)/100 for j in range(1,2*limit)]
print(vec)

Config1 = [0 for j in range(1,len(vec)+1)]
Config2 = [0 for j in range(1,len(vec)+1)]
LRsimfig = [0 for j in range(1,len(vec)+1)]
sco = [0 for j in range(1,len(vec)+1)]
count = 0
ns = 2000
for gamma in vec:
    count = count + 1
    Cov = matriz_cov(gamma, p)
    mu = [0 for j in range(1,p+1)]
    Con1 = [0 for j in range(1,ns+1)]
    Con2 = [0 for j in range(1,ns+1)]
    Con3 = [0 for j in range(1,ns+1)]
    LRsim = [0 for j in range(1,ns+1)]
    for j in range(1,ns):
        Data = np.random.multivariate_normal(mu, Cov, NN).T
        vecS = (NN-1)*np.cov(Data)
        LR = np.linalg.det(vecS)/np.prod(np.diagonal(vecS))
        if LR < quantil: LRsim[j-1] = 1
        sig = np.diagonal(np.cov(Data))
        datan1 = Data.sum(axis=0)
        if n*np.var(datan1)/sum(sig) > chi2.ppf(0.975,n) \
        or n*np.var(datan1)/sum(sig) < chi2.ppf(0.025,n): \
        Con1[j-1] = 1
        if n*np.var(datan1)/sum(np.diagonal(Cov)) > \
        chi2.ppf(0.975,n) or \
        n*np.var(datan1)/sum(np.diagonal(Cov)) < \
        chi2.ppf(0.025,n): Con2[j-1] =1
        vecSS = np.corrcoef(Data)
        stat = (sum(sum([[vecSS[i][jj]]**2 \
                        for jj in range(0,i)] \
                        for i in range(1,p)],[]))\

```

```

-(p*(p-1)/(2*n)))/((p*(p-1)*(n-1)/\
((n**2)*(n+2)))*(1/2))
    if stat > norm.ppf(0.975,0,1) \
    or stat < norm.ppf(0.025,0,1): \
    Con3[j-1] = 1
    sco[count-1] = sum(Con3)/ns
    Config1[count-1] = sum(Con1)/ns
    Config2[count-1] = sum(Con2)/ns
    LRsimfig[count-1] = sum(LRsim)/ns

plt.plot(vec, Config2, label="\u03C3's known", \
         linestyle='-', color='black')
plt.plot(vec, Config1, label="\u03C3's unknown", \
         linestyle=':', color='black')
plt.plot(vec, LRsimfig, label="LRT", \
         linestyle='—', color='black')
plt.plot(vec, sco, label="SCH", \
         linestyle='-.', color='black')
plt.xlabel("Valores de \u03C1")
plt.ylabel("Potencia Estimada")
plt.ylim(0,1)
plt.xlim(vec[0], vec[len(vec)-1])
plt.show()

```

O resultado do código acima foi apresentado na secção 4.1. É de notar que novamente, este código foi o usado para simular para o caso em que o número de variáveis foi $p = 5$ e o tamanho da amostra foi $n = 50$, sendo que para os restantes casos se usou o mesmo código mudando n e p e calculando o vetor vec de forma a fazer sentido para cada caso em particular considerando sempre os requisitos dados na secção 4.1.

I.3 Código usado para a estrutura de covariância circular

Para as simulações foi ainda considerada a estrutura circular para a matriz de covariâncias, sendo que de seguida apresentamos o código utilizado nesses casos. Tal com anteriormente, também aqui apenas se apresenta o caso em que foram consideradas $p = 5$ variáveis e amostras do tamanho $n = 50$, sendo que para os outros casos considerados se alterou os valores de n e p e se adaptou o cálculo do vetor dos valores de ρ , dado por vec no código, considerando sempre os critérios definidos anteriormente para este vetor. Neste caso em particular, a construção da matriz de covariâncias depende de p ser par ou ímpar, assim, apesar de o caso apresentado ser para p par, incluiu-se no código também a forma de cálculo da matriz para p ímpar, sendo fácil de identificar cada um dos casos.

```
import numpy as np
from scipy.stats import chi2
from scipy.stats import norm
from scipy.linalg import helmert
from math import *
import matplotlib.pyplot as plt
import math
import random
from scipy.linalg import circulant

def is_pos_def(x):
    return np.all(np.linalg.eigvals(x) > 0)

def vec_circ(r, j):
    a = [-1/(n*100) + j for n in range(2, r+2)]
    b = [-1/(n*100) + j for n in range(2, r+2)]
    circ=np.concatenate(([1], a, b[::-1]), axis=None)
    return(circ)

def vec_circ(r, j):
    a = [-1/(n*100) + j for n in range(2, r+2)]
    b = [-1/(n*100) + j for n in range(2, r+1)]
    circ=np.concatenate(([1], a, b[::-1]), axis=None)
    return(circ)

def matriz_cov_circ(r, j):
    if j == 0:
        Cov = np.identity(p)
    else:
        Cov = circulant(vec_circ(r, j))
    return(Cov)

def min_rho_def_pos(p, r):
    vec = [i/100 for i in range(-20, 50, 1)]
    for j in vec:
        Cov = matriz_cov_circ(r, j)
        if is_pos_def(Cov):
            break
    return(j)
```

```
def verify_gammas(r, j):
    vec = vec_circ(r, j)
    vec_len = [n for n in range(0, r+1)]
    contador_mais = 0
    contador_menos = 0
    for i in vec_len:
        if vec[i] < 0:
            contador_menos = contador_menos + 1
        else:
            contador_mais = contador_mais + 1
    if contador_menos > 0 and contador_mais != 1:
        result = j
    return(result)

np.random.seed(10)

nt = 10000

NN = 50
n = NN-1
p = 5
r=math.floor((p)/2)

mu = [0 for j in range(1, p+1)]
veca = [0 for j in range(1, nt+1)]
for i in range(1, nt):
    x = np.random.multivariate_normal(mu, np.identity(p), NN).T
    vecS=(NN-1)*np.cov(x)
    veca[i] =np.linalg.det(vecS)/np.prod(np.diagonal(vecS))
quantil=np.quantile(veca, 0.05)

value= min_rho_def_pos(p, r)
limit= 40
vec = [value + (j-1)/100 for j in range(1, 2*limit)]
print(vec)

gammas = []
for i in vec:
    if verify_gammas(r, i) == i:
        gammas.append(i)
```

```

vec = [i for i in vec if i not in gammas]
print(vec)

vec_len = [n for n in range(0, len(vec))]
for i in vec_len:
    if vec[i] == 0:
        if i == 0:
            before_zero = 0
            after_zero = vec[i+1]
        else:
            before_zero = vec[i-1]
            after_zero = vec[i+1]

def verify_signal(r, j):
    vec = vec_circ(r, j)
    vec_len = [n for n in range(0, r+1)]
    contador_mais = 0
    contador_menos = 0
    for i in vec_len:
        if vec[i] < 0:
            contador_menos = contador_menos + 1
        else:
            contador_mais = contador_mais + 1
    print("Contador_Mais:", contador_mais, \
          "Contador_Menos:", contador_menos)

Config1 = [0 for j in range(1, len(vec)+1)]
Config2 = [0 for j in range(1, len(vec)+1)]
LRsimfig = [0 for j in range(1, len(vec)+1)]
sco = [0 for j in range(1, len(vec)+1)]
count = 0
ns = 2000
for gamma in vec:
    count = count + 1
    Cov = matriz_cov_circ(r, gamma)
    mu = [0 for j in range(1, p+1)]
    Con1 = [0 for j in range(1, ns+1)]
    Con2 = [0 for j in range(1, ns+1)]
    Con3 = [0 for j in range(1, ns+1)]
    LRsim = [0 for j in range(1, ns+1)]

```

```

for j in range(1, ns):
    Data = np.random.multivariate_normal(mu, Cov, NN).T
    vecS = (NN-1)*np.cov(Data)
    LR = np.linalg.det(vecS)/np.prod(np.diagonal(vecS))
    if LR < quantil: LRsim[j-1] = 1
    sig = np.diagonal(np.cov(Data))
    datan1 = Data.sum(axis=0)
    if n*np.var(datan1)/sum(sig) > chi2.ppf(0.975, n) \
    or n*np.var(datan1)/sum(sig) < chi2.ppf(0.025, n): \
    Con1[j-1] = 1
    if n*np.var(datan1)/sum(np.diagonal(Cov)) > \
    chi2.ppf(0.975, n) or \
    n*np.var(datan1)/sum(np.diagonal(Cov)) \
    < chi2.ppf(0.025, n): Con2[j-1] = 1
    vecSS = np.corrcoef(Data)
    stat = (sum(sum([[vecSS[i][jj]]**2 \
                    for jj in range(0, i)] \
                    for i in range(1, p)], [])) - \
            (p*(p-1)/(2*n)))/((p*(p-1)*(n-1)/\
                               ((n**2)*(n+2)))* (1/2))

    if stat > norm.ppf(0.975, 0, 1) \
    or stat < norm.ppf(0.025, 0, 1): Con3[j-1] = 1
    sco[count-1] = sum(Con3)/ns
    Config1[count-1] = sum(Con1)/ns
    Config2[count-1] = sum(Con2)/ns
    LRsimfig[count-1] = sum(LRsim)/ns

plt.plot(vec, Config2, label="\u03C3's known", \
         linestyle='-', color='black')
plt.plot(vec, Config1, label="\u03C3's unknown", \
         linestyle=':', color='black')
plt.plot(vec, LRsimfig, label="LRT", \
         linestyle='—', color='black')
plt.plot(vec, sco, label="SCH", \
         linestyle='-.', color='black')
plt.xlabel("Valores de \u03C1")
plt.ylabel("Potencia Estimada")
plt.ylim(0, 1)
plt.xlim(vec[0], vec[len(vec)-1])
plt.show()

```

I.4 Código usado para as simulações usando populações t-Student

Para simular a partir de populações t-Student, cujos resultados são apresentados na secção 4.3, foi utilizado o código seguinte:

```
!pip install scipy==1.6
import numpy as np
import random
from scipy.stats import chi2
from scipy.stats import norm
from scipy.stats import beta
from scipy.stats import multivariate_t
from scipy.linalg import helmert
from math import *
import math
import matplotlib.pyplot as plt
from scipy.linalg import toeplitz
from scipy.linalg import circulant
import numpy as np

def is_pos_def(x):
    return np.all(np.linalg.eigvals(x) > 0)

def multivariate_t_rvs(m, S, df=np.inf, n=1):
    '''generate random variables of multivariate
    t distribution
    Parameters
    -----
    m : array_like
        mean of random variable, length determines
        dimension of random variable
    S : array_like
        square array of covariance matrix
    df : int or float
        degrees of freedom
    n : int
        number of observations, return random
        array will be (n, len(m))
    Returns
    -----
```

```
    rvs : ndarray, (n, len(m))
        each row is an independent draw of a
        multivariate t distributed
        random variable
    '''
    m = np.asarray(m)
    d = len(m)
    if df == np.inf:
        x = np.ones(n)
    else:
        x = np.random.chisquare(df, n)/df
    z = np.random.multivariate_normal(np.zeros(d), S, (n,))
    return m + z/np.sqrt(x)[:,None]

np.random.seed(10)
nt=10000
n = 49
NN = n + 1
p = 15
mu = [0 for j in range(1,p+1)]
veca = [0 for j in range(1,nt+1)]
vecS = 0
l = random.sample(range(1,100), p)
for i in range(1,nt):
    x = np.random.multivariate_normal(mu, np.diag(l), NN).T
    vecS=(NN-1)*np.cov(x)
    veca[i] =np.linalg.det(vecS)/np.prod(np.diagonal(vecS))
quantil=np.quantile(veca,0.05)
print(quantil)

def is_pos_def(x):
    return np.all(np.linalg.eigvals(x) > 0)
vec=[j/10000 for j in range(-5000,1,1)]
for gamma in vec:
    Cov = np.dot(gamma,([[1] * p] * p)) + \
    np.dot((1-gamma),np.identity(p))
    if is_pos_def(Cov):
        print(gamma)
        break
```

```

value=-0.07
limit= 7+2
vec = [value + (j-1)/100 for j in range(1,2*limit)]
Config = [0 for j in range(1,len(vec)+1)]
Config1 = [0 for j in range(1,len(vec)+1)]
Config2 = [0 for j in range(1,len(vec)+1)]
LRsimfig = [0 for j in range(1,len(vec)+1)]
sco = [0 for j in range(1,len(vec)+1)]
count = 0
ns = 1000
for gamma in vec:
    count = count + 1
    Cov = np.dot(gamma, ([[1] * p] * p)) \
    + np.dot((1-gamma), np.identity(p))
    mu = [0 for j in range(1,p+1)]
    Con = [0 for j in range(1,ns+1)]
    Con1 = [0 for j in range(1,ns+1)]
    Con2 = [0 for j in range(1,ns+1)]
    Con3 = [0 for j in range(1,ns+1)]
    LRsim = [0 for j in range(1,ns+1)]
    for j in range(1,ns):
        #Data = np.random.multivariate_normal
        #(mu, Cov, NN).T
        Data = multivariate_t_rvs(np.zeros(p),Cov, 10, NN).T
        vecS = (NN-1)*np.cov(Data)
        LR = np.linalg.det(vecS)/np.prod(np.diagonal(vecS))
        sig = np.diagonal(np.cov(Data))
        datan1 = Data.sum(axis=0)
        if n*np.var(datan1)/sum(sig) > chi2.ppf(0.975,n) \
        or n*np.var(datan1)/sum(sig) < chi2.ppf(0.025,n): \
        Con1[j-1] = 1
        if n*np.var(datan1)/sum(np.diagonal(Cov)) > \
        chi2.ppf(0.975,n) or \
        n*np.var(datan1)/sum(np.diagonal(Cov)) < \
        chi2.ppf(0.025,n): Con2[j-1] =1
        if LR < quantil: LRsim[j-1] = 1
        vecSS = np.corrcoef(Data)
        stat = (sum(sum([[vecSS[i][jj]]**2 \
                        for jj in range(0,i)] \
                        for i in range(1,p)],[]))-\

```

```

        (p*(p-1)/(2*n)))/((p*(p-1)*(n-1)/\
                               ((n**2)*(n+2)))*(1/2))
    if stat > norm.ppf(0.975,0,1) \
    or stat < norm.ppf(0.025,0,1): Con3[j-1] = 1
    sco[count-1] = sum(Con3)/ns
    Config1[count-1] = sum(Con1)/ns
    Config2[count-1] = sum(Con2)/ns
    LRsimfig[count-1] = sum(LRsim)/ns

plt.plot(vec, Config2, label="\u03C3's known", \
         linestyle='-', color='black')
plt.plot(vec, Config1, label="\u03C3's unknown", \
         linestyle=':', color='black')
plt.plot(vec, LRsimfig, label="LRT", \
         linestyle='—', color='black')
plt.plot(vec, sco, label="SCH", \
         linestyle='-.', color='black')
plt.xlabel("Valores de \u03C1")
plt.ylabel("Potencia Estimada")
plt.ylim(0,1)
plt.xlim(vec[0], vec[len(vec)-1])
plt.show()

```

Tal como nos restantes casos, apenas aqui apresentamos o código utilizado para $p = 15$ variáveis e amostra de tamanho $n = 50$, e também para uma distribuição t-student com $gl = 10$ graus de liberdade, no entanto para os restantes casos estudados utilizou-se o mesmo código substituindo apenas n , p e os graus de liberdade da função que gera amostras de uma distribuição t-student.

1.5 Código usado para as simulações do teste de esfericidade vs equivariância-equicorrelação

De seguida apresentamos o código utilizado nas simulações cujos resultados são apresentados na secção 4.4, onde se apresenta o código utilizado para simular para $p = 5$ variáveis e amostras de tamanho $n = 50$, sendo que para os restantes casos apenas se alteraram estes valores, e o vetor *vec* que contém os valores para ρ foi calculado de forma a cumprir os requisitos definidos para ρ .

```

import numpy as np
from scipy.stats import chi2
from scipy.stats import norm
from scipy.linalg import helmert

```

I.5. CÓDIGO USADO PARA AS SIMULAÇÕES DO TESTE DE ESFERICIDADE VS EQUIVARIÂNCIA-EQUICORRELAÇÃO

```
from math import *
import matplotlib.pyplot as plt
import math
import random
from scipy.linalg import circulant

np.random.seed(10)
nt=1000
n = 49
NN = n + 1
p = 5
mu = [0 for j in range(1,p+1)]
vecaa = [0 for j in range(1,nt+1)]
vecaa1 = [0 for j in range(1,nt+1)]
vecaa2 = [0 for j in range(1,nt+1)]
vecS = 0
vecSS = 0
Mat = 0
l = random.sample(range(1,100), p)
for i in range(1,nt):
    x = np.random.\
    multivariate_normal(mu, np.identity(p), NN).T
    vecS = ((NN-1)/NN)*np.cov(x)
    veca[i] = np.linalg.det(vecS)/\
    ((1/p*np.sum(np.diagonal(vecS)))**p)
    Mat = np.dot(helmert(p,True),np.cov(x))
    vecSS = (NN-1)/NN*\
    (np.dot(Mat,helmert(p,True).T))
    vecaa[i] = np.linalg.det(vecSS)/\
    ((1/p*np.sum(np.diagonal(vecSS)))**p)
    vecaa1[i] = np.linalg.det(vecSS)/\
    (vecSS[0,0]*\
    (1/\(p-1)*\
    np.sum(np.diagonal(vecSS[1:p,1:p]))))** (p-1))
    vecaa2[i] = vecaa[i]/vecaa1[i]
quantil=np.quantile(vecaa2,0.05)
print(quantil)

value=-0.24
```

```
limit= 25+5
vec = [value + (j-1)/100 \
       for j in range(1,2*limit)]
Config = [0 for j in range(1,len(vec)+1)]
Config1 = [0 for j in range(1,len(vec)+1)]
Config2 = [0 for j in range(1,len(vec)+1)]
LRsimfig = [0 for j in range(1,len(vec)+1)]
sco = [0 for j in range(1,len(vec)+1)]
count = 0
ns = 2000
for gamma in vec:
    count = count + 1
    Cov = np.dot(gamma,([1] * p) * p) + \
          np.dot((1-gamma),np.identity(p))
    mu = [0 for j in range(1,p+1)]
    Con = [0 for j in range(1,ns+1)]
    Con1 = [0 for j in range(1,ns+1)]
    Con2 = [0 for j in range(1,ns+1)]
    Con3 = [0 for j in range(1,ns+1)]
    LRsim = [0 for j in range(1,ns+1)]
    for j in range(1,ns):
        Data = np.random.multivariate_normal(mu, Cov, NN).T
        Mat = np.dot(helmert(p, True), np.cov(Data))
        vecSS = (NN-1)/NN*(np.dot(Mat, helmert(p, True).T))
        LR = (np.linalg.det(vecSS)/\
              ((1/p*np.sum(np.diagonal(vecSS)))**p))/\
              (np.linalg.det(vecSS)/\
               (vecSS[0,0]*(1/(p-1))*\
                np.sum(np.diagonal(vecSS[1:p,1:p])))**(p-1)))
        sig = np.diagonal(np.cov(Data))
        datan1 = Data.sum(axis=0)
        if n*np.var(datan1)/sum(sig) > chi2.ppf(0.975,n) or \
            n*np.var(datan1)/sum(sig) < chi2.ppf(0.025,n): \
            Con1[j-1] = 1
        if n*np.var(datan1)/ sum(np.diagonal(Cov)) > \
            chi2.ppf(0.975,n) or \
            n*np.var(datan1)/ sum(np.diagonal(Cov)) \
            < chi2.ppf(0.025,n): Con2[j-1] =1
        if LR < quantil: LRsim[j-1] = 1
    Config1[count-1] = sum(Con1)/ns
```

```
Config2[count-1] = sum(Con2)/ns
LRsimfig[count-1] = sum(LRsim)/ns

plt.plot(vec, Config2, label="\u03C3's\u2099known", \
         linestyle='-', color='black')
plt.plot(vec, Config1, label="\u03C3's\u2099unknown", \
         linestyle=':', color='black')
plt.plot(vec, LRsimfig, label="LRT", \
         linestyle='—', color='black')
plt.xlabel("Valores de \u03C1")
plt.ylabel("Potencia Estimada")
plt.ylim(0,1)
plt.xlim(vec[0],vec[len(vec)-1])
plt.show()
```

I.6 Código usado para as simulações do teste de esfericidade vs circularidade

As simulações na secção 4.5 foram realizadas com o seguinte código:

```
import numpy as np
from scipy.stats import chi2
from scipy.stats import norm
from scipy.linalg import helmert
from math import *
import matplotlib.pyplot as plt
import math
import random
from scipy.linalg import circulant

np.random.seed(10)
nt=10000
n = 49
NN = n + 1
p = 5
mu = [0 for j in range(1,p+1)]
vecaa = [0 for j in range(1,nt+1)]
vecaa1 = [0 for j in range(1,nt+1)]
vecaa2 = [0 for j in range(1,nt+1)]
vecS = 0
vecSS = 0
```

```

Mat = 0
hel= np.zeros( (p,p) )
for k in range(1,p+1):
    for j in range(1,p+1):
        hel[j-1,k-1]=p**(-1/2)*\
            (math.cos(2*math.pi*p**(-1)*(j-1)*(k-1))+\
             math.sin(2*math.pi*p**(-1)*(j-1)*(k-1)))
for i in range(1,nt):
    x = np.random.multivariate_normal(mu,np.identity(p), NN).T
    Mat = np.dot(hel.T,np.cov(x))
    V = np.dot(Mat,hel)
    m = math.floor((p)/2)
    vv = [0 for j in range(1,m+2)]
    vj = np.diagonal(V)
    vv[0]=vj[0]
    if p % 2 == 0:
        for j in range(2,m+1):
            vv[j-1]=vj[j-1]+vj[len(vj)-j+1]
        vv[m]=vj[m]
    else:
        for j in range(2,m+2):
            vv[j-1]=vj[j-1]+vj[len(vj)-j+1]
    if p % 2 == 0:
        vvv=np.concatenate((vj[0],vv[1:m+1],\
                             vv[1:m+1][::-1][1:m+1]),axis=None)
    else:
        vvv=np.concatenate((vj[0],vv[1:m+1],\
                             vv[1:m+1][::-1]),axis=None)
    vecaa1[i] = 2**(2*(p-m-1))*\
np.linalg.det(V)/(np.prod(vvv))
    vecS = (NN-1)*np.cov(x)
    vecaa[i] = np.linalg.det(vecS)/\
((1/p*np.sum(np.diagonal(vecS)))*p)
    vecaa2[i] = vecaa[i]/vecaa1[i]
quantil=np.quantile(vecaa2,0.05)
print(quantil)

def is_pos_def(x):
    return np.all(np.linalg.eigvals(x) > 0)
r=math.floor((p)/2)

```

I.6. CÓDIGO USADO PARA AS SIMULAÇÕES DO TESTE DE ESFERICIDADE VS CIRCULARIDADE

```
vec = [i/100 for i in range(-20,50,1)]
for j in vec:
    a = [-1/(n*100) + j for n in range(2,r+2)]
    b = [-1/(n*100) + j for n in range(2,r+2)]
    circ=np.concatenate(([1],a, b[::-1]),axis=None)
    Cov=circulant(circ)
    print(Cov)
    if not is_pos_def(Cov):
        print(j)
print("OK")
print(Cov)
len(vec)

Config = [0 for j in range(1,len(vec)+1)]
Config1 = [0 for j in range(1,len(vec)+1)]
Config2 = [0 for j in range(1,len(vec)+1)]
LRsimfig = [0 for j in range(1,len(vec)+1)]
sco = [0 for j in range(1,len(vec)+1)]
count = 0
ns = 5000
for gamma in vec:
    count = count + 1
    r=math.floor((p)/2)
    a = [-1/(n*100) + gamma for n in range(2,r+2)]
    b = [-1/(n*100) + gamma for n in range(2,r+2)]
    circ=np.concatenate(([1],a, b[::-1]), axis=None)
    Cov=circulant(circ)
    vecaa = [0 for j in range(1,nt+1)]
    vecaa1 = [0 for j in range(1,nt+1)]
    vecaa2 = [0 for j in range(1,nt+1)]
    mu = [0 for j in range(1,p+1)]
    Con = [0 for j in range(1,ns+1)]
    Con1 = [0 for j in range(1,ns+1)]
    Con2 = [0 for j in range(1,ns+1)]
    Con3 = [0 for j in range(1,ns+1)]
    LRsim = [0 for j in range(1,ns+1)]
    for jj in range(1,ns):
        Data = np.random.multivariate_normal(mu, Cov, NN).T
        Mat = np.dot(hel.T,np.cov(Data))
        V = np.dot(Mat,hel)
```

```

m = math.floor((p)/2)
vv = [0 for j in range(1,m+2)]
vj = np.diagonal(V)
vv[0]=vj[0]
if p % 2 == 0:
    for j in range(2,m+1):
        vv[j-1]=vj[j-1]+vj[len(vj)-j+1]
        vv[m]=vj[m]
else:
    for j in range(2,m+2):
        vv[j-1]=vj[j-1]+vj[len(vj)-j+1]
if p % 2 == 0:
    vvv=np.concatenate\
    ((vj[0],vv[1:m+1],\
    vv[1:m+1][::-1][1:m+1]),axis=None)
else:
    vvv=np.\
    concatenate((vj[0],\
    vv[1:m+1],vv[1:m+1][::-1]),axis=None)
vecaa1[jj] = 2*(2*(p-m-1))*\
np.linalg.det(V)/(np.prod(vvv))
vecS = (N-1)*np.cov(Data)
vecaa[jj] = np.linalg.det(vecS)/\
((1/p*np.sum(np.diagonal(vecS))))*p)
LR = vecaa[jj]/vecaa1[jj]
sig = np.diagonal(np.cov(Data))
datan1 = Data.sum(axis=0)
if n*np.var(datan1)/sum(sig) > \
chi2.ppf(0.975,n) or \
n*np.var(datan1)/sum(sig) < \
chi2.ppf(0.025,n): Con1[jj-1] = 1
if n*np.var(datan1)/sum(np.diagonal(Cov)) >\
chi2.ppf(0.975,n) or \
n*np.var(datan1)/sum(np.diagonal(Cov)) < \
chi2.ppf(0.025,n): Con2[jj-1] =1
if LR < quantil: LRsim[jj-1] = 1
Config1[count-1] = sum(Con1)/ns
Config2[count-1] = sum(Con2)/ns
LRsimfig[count-1] = sum(LRsim)/ns

```

I.6. CÓDIGO USADO PARA AS SIMULAÇÕES DO TESTE DE ESFERICIDADE VS CIRCULARIDADE

```
plt.plot(vec, Config2, label="\u03C3's known", \
         linestyle='-', color='black')
plt.plot(vec, Config1, label="\u03C3's unknown", \
         linestyle=':', color='black')
plt.plot(vec, LRsimfig, label="LRT", \
         linestyle='—', color='black')
plt.xlabel("Valores de \u03C1")
plt.ylabel("Pot\u00eancia Estimada")
plt.ylim(0,1)
plt.xlim(vec[0], vec[len(vec)-1])
plt.show()
```

o c\u00f3digo anterior \u00e9 relativo ao caso onde se simulou para $p = 5$ vari\u00e1veis e amostras de tamanho $n = 50$, no entanto para os restantes casos estudados usou-se o mesmo caso, alterando apenas os valores de p e n e ainda consoante p seja par ou \u00edmpar usou-se a matriz respetiva, sendo que o c\u00f3digo para ambos os casos v\u00eam aqui representado, sendo necess\u00e1rio usar apenas um para as imula\u00e7\u00f5es funcionarem corretamente.

