TESTE *BOOTSTRAP* NÃO-PARAMÉTRICO PARA A IGUALDADE DE MATRIZES DE COVARIÂNCIAS DE DUAS POPULAÇÕES NORMAIS MULTIVARIADAS DEPENDENTES

Vanessa Siqueira Peres da SILVA[1]
Daniel Furtado FERREIRA[1]

- RESUMO: Este trabalho tem por objetivo avaliar as taxas de erro tipo I e poder do teste *bootstrap* não-paramétrico (t_{b_0}) para a igualdade de matrizes de covariâncias de duas populações normais multivariadas dependentes, com o intuito de comparar o seu desempenho com o dos testes apresentados por Jiang e Sarkar (1998) (W_2 e W_5) e Jiang et al. (1999) (LRT, LRT_1, LRT_2 e LRT_3). Para isso foram realizadas simulações Monte Carlo, considerando número de variáveis (p), tamanhos amostrais (n), matrizes de covariâncias ($\mathbf{\Sigma}$) e nível de significância (α) de 0,05. No primeiro caso, para $p = 2$, concluiu-se que dentre os testes que controlaram o erro tipo I, os testes t_{b_0}, LRT_3 e W_2 foram superiores aos seus competidores em todas as situações estudadas. Em relação ao poder, o teste t_{b_0} aproximou-se dos testes LRT_3 e W_2, sendo considerado intermediário. No segundo caso, em que considerou-se $p = 4$ e $p = 10$, concluiu-se que o teste t_{b_0} apresentou um desempenho elevado, na maioria das vezes igual a 100%, mesmo para pequenas amostras ($n = 20$). Portanto, recomenda-se a aplicação do testes proposto t_{b_0} em situações reais.

- PALAVRAS-CHAVE: Matrizes de covariâncias; teste *bootstrap*; simulação Monte Carlo; poder; erro tipo I.

1 Introdução

Em diversas situações das ciências biológicas, físicas e humanas é muito comum o pesquisador ter o interesse em comparar matrizes de variâncias e covariâncias de duas populações. Se as amostras de ambas as populações forem independentes nenhuma covariância entre elas é esperada, e alguns testes, bastantes conhecidos na

[1]Universidade Federal de Lavras – UFLA, Programa de Pós-graduação em Estatística e Experimentação Agropecuária, Caixa Postal 3037, CEP: 37200-000, Lavras, MG, Brasil. E-mail: *dexvanessa@gmail.com*; *danielff@dex.ufla.br*

literatura, podem ser aplicados, como o teste de Bartlett (1937). Por outro lado, os dados amostrais podem ser emparelhados, em situações em que um grupo de variáveis é mensurado antes e após a realização de um determinado tratamento, ou ainda, em situações na genética em que as medidas são realizadas em uma geração e, também, na geração filial derivada de cada indivíduo, sendo a genealogia preservada. Nesse caso, não é possível aplicar o teste de Bartlett (1937). A razão para isso é que, de alguma forma, esses dados são correlacionados e o teste possui baixo poder em detectar possíveis diferenças na matriz de covariâncias, uma vez que se viola a pressuposição, para sua aplicação, de que as duas populações são independentes.

Para o caso em que apenas uma variável é mensurada em cada situação, pré (X) e pós (Y) tratamento, Morgan (1939) e Pitman (1939) propuseram um teste t exato baseado na correlação entre as variáveis normais X e Y e na correlação de duas novas variáveis, U e V, que são combinações lineares de X e Y. O teste de Morgan (1939) e Pitman (1939) se torna bem mais poderoso do que as alternativas existentes para dados não correlacionados à medida que a correlação entre X e Y tende a 1 ou a -1. Esse teste considera, no entanto, apenas a situação de $q = 2$ populações e $p = 1$ variável. Para o caso de $q \geq 2$ populações e $p \geq 1$ variável, algumas soluções são propostas na literatura (BHATTACHARJEE e ALAM, 1982; BOGLE e HSU, 2002; CHU e PILLAI, 1980; HAN, 1968; HAYAKAWA, 1987; JIANG e SARKAR, 2000, 2002; JIANG et al., 1999; PIEPHO, 1997; SHAPIRO e COHEN, 1990). No entanto, todos são testes assintóticos.

Para os casos de dependência entre as populações, Jiang et al. (1999) propuseram alguns testes baseados na razão de verossimilhanças, que são limitados em relação ao número de variáveis e de populações, uma vez que suas estatísticas são baseadas em aproximações assintóticas da distribuição qui-quadrado (JIANG et al., 1999). De acordo com Cirillo et al. (2009), outro item que deve ser destacado é a complexidade de obter expressões analíticas para tais testes, além do mais, problemas numéricos para maximização das verossimilhanças considerando um maior número de variáveis e populações são comumente encontrados. Mediante esses problemas, como alternativa surge o uso de técnicas computacionais, das quais os métodos de computação intensiva, como técnicas de *bootstrap* são de grande importância, nas mais variadas situações reais (MANLY, 1997).

Segundo Ferreira (2013), a vantagem de utilizar os métodos de reamostragens como o *bootstrap* para realizar inferências ocorre quando não se conhece a distribuição de probabilidade da população e o modelo normal não é adequado para os dados ou resíduos. Esse método de reamostragem foi proposto por Efron (1979). O termo *bootstrap* refere-se a simulações Monte Carlo que tratam a amostra original como a pseudo-população, que é assumida para representar a verdadeira população de interesse. O pesquisador trata a sua amostra como se fosse a população de origem dos dados e utiliza a amostragem com reposição para gerar pseudo-amostras a partir da amostra original. Uma amostra original de tamanho n é considerada e cada elemento da amostra tem a mesma probabilidade de ser selecionado. Assim, pode-se estimar os parâmetros da população a partir das pseudo-amostras por meio da distribuição empírica amostral conhecida.

A ideia do método *bootstrap* não-paramétrico é realizar reamostragem da amostra original com reposição , formando novas amostras de mesmo tamanho. É usada quando não se conhece a distribuição de probabilidade da variável aleatória ou da estatística de interesse.

Suponha uma amostra aleatória de tamanho n, $X_1, X_2, \ldots, X_n$, de uma população com distribuição desconhecida e cujo interesse é um parâmetro θ. Essa amostra será reamostrada B vezes com reposição, de forma que cada reamostragem (subamostra) $(1, 2, \ldots, B)$ tenha tamanho n. A cada subamostra gerada calcula-se uma estatística de interesse, e após B reamostragens a distribuição empírica gerada é utilizada para a realização de inferências.

De acordo com Ferreira (2013), para a maior parte dos modelos probabilísticos teorias exatas não são conhecidas ou são intratáveis analiticamente. Nesse caso, usa-se a distribuição de probabilidade empírica, a qual atribui para cada observação amostral uma probabilidade $1/n$. Ainda segundo Ferreira (2013), a ideia de *bootstrap* é substituir a distribuição desconhecida da população pela distribuição empírica, que é conhecida. Por essa razão que se denomina esse método de *bootstrap* não-paramétrico.

Para lidar com o problema de comparar matrizes de covariâncias de distribuições normais dependentes este trabalho foi proposto procurando generalizar o teste de Morgan (1939) e Pitman (1939) para o caso multivariado, considerando a situação de $q = 2$ populações. Comparações utilizando simulação Monte Carlo foram realizadas para avaliar o desempenho do teste proposto em relação ao dos testes existentes para homogeneidade de covariâncias.

Na seção 2 buscou-se desenvolver teoricamente o teste proposto (*bootstrap* não-paramétrico (t_{b_0})). Posteriormente, foram construídas matrizes de covariâncias positivas definidas, provenientes de $p = 4$ e $p = 10$ populações, com o intuito de avaliar as taxas de erro tipo I e poder do teste t_{b_0}. Finalmente, na seção 3, comparações utilizando simulação Monte Carlo foram realizadas, considerando $p = 2$, para avaliar o desempenho dos testes propostos em relação aos testes existentes, propostos por Jiang e Sarkar (1998) (W_2 e W_5) e Jiang et al. (1999) (LRT, LRT_1, LRT_2 e LRT_3). Os testes propostos por Jiang e Sarkar (1998) e Jiang et al. (1999) se limitaram a apenas duas variáveis. Portanto, não foi possível fazer comparações desses testes com o teste proposto (t_{b_0}) quando $p = 4$ e $p = 10$. O teste proposto pode ser útil nos casos de $p = 4$ e $p = 10$, pois é muito comum termos mais de quatro variáveis em problemas reais.

2 Metodologia

Neste trabalho foi proposto um método, baseado na técnica de reamostragem *bootstrap*, para testar a hipótese a seguir:

$$H_0 : \mathbf{\Sigma}_{11} = \mathbf{\Sigma}_{22} \quad \text{vs} \quad H_1 : \mathbf{\Sigma}_{11} \neq \mathbf{\Sigma}_{22}, \qquad (1)$$

quando $\mathbf{\Sigma}_{12} \neq \mathbf{0}$, em que $\mathbf{\Sigma}_{11}$ e $\mathbf{\Sigma}_{22}$ são as matrizes de covariâncias das populações 1 e 2, respectivamente. Os testes para o caso em que $\mathbf{\Sigma}_{12} \neq \mathbf{0}$ são bastantes conhecidos na literatura (BARTLETT, 1937).

Para desenvolver o teste para $q = 2$ populações, sejam definidas as variáveis U e V como combinações lineares das variáveis de $\boldsymbol{X}$. Sejam os vetores $\boldsymbol{b}^\top = (\boldsymbol{a}^\top, \boldsymbol{a}^\top)$ e $\boldsymbol{c}^\top = (\boldsymbol{a}^\top, -\boldsymbol{a}^\top)$ de dimensão $(1 \times 2p)$. Portanto, sejam $U = \boldsymbol{b}^\top \boldsymbol{X} = \boldsymbol{a}^\top \boldsymbol{X}_1 + \boldsymbol{a}^\top \boldsymbol{X}_2 = U_1 + U_2$ e $V = \boldsymbol{c}^\top \boldsymbol{X} = \boldsymbol{a}^\top \boldsymbol{X}_1 - \boldsymbol{a}^\top \boldsymbol{X}_2 = U_1 - U_2$. Como U e V são combinações lineares de variáveis normais $\boldsymbol{X}$, então, $U \sim N_1(\mu_U, \sigma_U^2)$ e $V \sim N_1(\mu_V, \sigma_V^2)$ (ANDERSON, 1978), sendo:

$$\mu_U = \boldsymbol{a}^\top \boldsymbol{\mu}_1 + \boldsymbol{a}^\top \boldsymbol{\mu}_2 \quad \text{e} \quad \sigma_U^2 = \boldsymbol{a}^\top (\boldsymbol{\Sigma}_{11} + \boldsymbol{\Sigma}_{22} + 2\boldsymbol{\Sigma}_{12})\, \boldsymbol{a};$$
$$\mu_V = \boldsymbol{a}^\top \boldsymbol{\mu}_1 - \boldsymbol{a}^\top \boldsymbol{\mu}_2 \quad \text{e} \quad \sigma_V^2 = \boldsymbol{a}^\top (\boldsymbol{\Sigma}_{11} + \boldsymbol{\Sigma}_{22} - 2\boldsymbol{\Sigma}_{12})\, \boldsymbol{a}.$$

A correlação entre U e V é definida por

$$\rho_{UV} = \frac{\sigma_{UV}}{\sqrt{\sigma_U^2}\sqrt{\sigma_V^2}},$$

em que σ_{UV} é a covariância entre U e V definida por $\sigma_{UV} = E(UV) - E(U)E(V)$, σ_U^2 e σ_V^2 foram definidos anteriormente. Então,

$$\sigma_{UV} = Cov(\boldsymbol{b}^\top \boldsymbol{X}, \boldsymbol{c}^\top \boldsymbol{X}) = \boldsymbol{b}^\top \boldsymbol{\Sigma} \boldsymbol{c} = \boldsymbol{a}^\top (\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{22})\, \boldsymbol{a}.$$

Portanto,

$$\rho_{UV} = \frac{\boldsymbol{a}^\top (\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{22})\, \boldsymbol{a}}{\sqrt{\boldsymbol{a}^\top (\boldsymbol{\Sigma}_{11} + \boldsymbol{\Sigma}_{22} + 2\boldsymbol{\Sigma}_{12})\, \boldsymbol{a}}\sqrt{\boldsymbol{a}^\top (\boldsymbol{\Sigma}_{11} + \boldsymbol{\Sigma}_{22} - 2\boldsymbol{\Sigma}_{12})\, \boldsymbol{a}}}. \tag{2}$$

O estimador de máxima verossimilhança de sigma corrigido para viés ($\boldsymbol{S}$) em uma amostra de tamanho n, devidamente particionado é

$$\boldsymbol{S} = \left[\begin{array}{c|c} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \hline \boldsymbol{S}_{21} & \boldsymbol{S}_{22} \end{array}\right] = \frac{\displaystyle\sum_{j=1}^{n} (\boldsymbol{X}_j - \bar{\boldsymbol{X}})(\boldsymbol{X}_j - \bar{\boldsymbol{X}})^\top}{n-1}. \tag{3}$$

O estimador de ρ_{UV} é dado por

$$r_{UV} = \frac{\boldsymbol{a}^\top (\boldsymbol{S}_{11} - \boldsymbol{S}_{22})\, \boldsymbol{a}}{\sqrt{\boldsymbol{a}^\top (\boldsymbol{S}_{11} + \boldsymbol{S}_{22} + 2\boldsymbol{S}_{12})\, \boldsymbol{a}}\sqrt{\boldsymbol{a}^\top (\boldsymbol{S}_{11} + \boldsymbol{S}_{22} - 2\boldsymbol{S}_{12})\, \boldsymbol{a}}}. \tag{4}$$

A expressão (2) permite que se conclua que $\rho_{UV} = 0$ se e somente se $\boldsymbol{\Sigma}_{11} = \boldsymbol{\Sigma}_{22}$. Portanto, testar a hipótese $H_0 : \rho_{UV} = 0$ é equivalente a testar $H_0 : \boldsymbol{\Sigma}_{11} = \boldsymbol{\Sigma}_{22}$.

O teste *bootstrap* não-paramétrico será proposto a seguir, para fins de comparação com os testes propostos por Jiang e Sarkar (1998) e Jiang et al. (1999), quando $p = 2$. Foram construídas matrizes de covariâncias positivas definidas, considerando $p = 4$ e $p = 10$. Nesses casos não foi possível comparar o teste *bootstrap* não-paramétrico com os propostos por Jiang e Sarkar (1998) e Jiang et al. (1999), pois esses autores não apresentaram resultados de simulação Monte Carlo quando $p > 2$.

Para a construção do teste *bootstrap* não-paramétrico, considere uma amostra aleatória *bootstrap*, $\boldsymbol{X}_1,\ \boldsymbol{X}_2,\ \cdots,\ \boldsymbol{X}_j,\ \cdots,\ \boldsymbol{X}_n$, independente e identicamente distribuída, de uma população qualquer, em que $\boldsymbol{X}_j$ tem dimensão $(2p \times 1)$, $j = 1, \ldots, n$, com vetor de média $\boldsymbol{\mu}_{(2p\times 1)}$ e matriz de covariâncias $\boldsymbol{\Sigma}_{(2p\times 2p)}$ positiva definida. Em seguida, fixou-se o número de reamostragens *bootstrap* em $B = 1000$ e calculou-se o estimador $\boldsymbol{S}$ $(2p \times 2p)$, da matriz de covariâncias $\boldsymbol{\Sigma}$, dado por:

$$\boldsymbol{S} = \left[\begin{array}{c|c} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \hline \boldsymbol{S}_{21} & \boldsymbol{S}_{22} \end{array}\right].$$

Posteriormente, determinaram-se e isolaram-se as partições $\boldsymbol{S}_{11}$ e $\boldsymbol{S}_{22}$ de $\boldsymbol{S}$ para o cálculo da matriz $\hat{\boldsymbol{H}} = \hat{\boldsymbol{\Omega}}^{-1}(\boldsymbol{S}_{11} - \boldsymbol{S}_{22})(\hat{\boldsymbol{\Omega}}^{\top})^{-1}$, sendo a matriz $\hat{\boldsymbol{\Omega}}$ o fator de Cholesky de $\boldsymbol{S}_{11} + \boldsymbol{S}_{22}$. Assim, o máximo autovalor de $\hat{\boldsymbol{H}}$ é dado por $\hat{\rho}_{UV} = \hat{\lambda}_1$. Dessa forma, calculou-se

$$t_{b_0} = \frac{\hat{\rho}_{UV}\sqrt{n-2}}{\sqrt{1-\hat{\rho}^2_{UV}}} \tag{5}$$

para gerar a estatística original.

A distribuição nula de *bootstrap* (distribuição de reamostragem *bootstrap* sob H_0) é proposta conforme os passos descritos na sequência. Inicialmente recebeu-se a amostra original $\boldsymbol{X}_1,\ \boldsymbol{X}_2,\ \cdots,\ \boldsymbol{X}_j,\ \cdots,\ \boldsymbol{X}_n$, considerando $\boldsymbol{X}_j = [\boldsymbol{X}^{\top}_{j(1)} \,|\, \boldsymbol{X}^{\top}_{j(2)}]^{\top}$. Em seguida, obtiveram-se as médias amostrais

$$\bar{\boldsymbol{X}}_{(1)} = \frac{\sum_{j=1}^{n} \boldsymbol{X}_{j(1)}}{n} \qquad \text{e} \qquad \bar{\boldsymbol{X}}_{(2)} = \frac{\sum_{j=1}^{n} \boldsymbol{X}_{j(2)}}{n}$$

para o cálculo dos desvios

$$\boldsymbol{e}_j = [\boldsymbol{X}^{\top}_{j(1)} \,|\, \boldsymbol{X}^{\top}_{j(2)}]^{\top} - \left[\begin{array}{c} \bar{\boldsymbol{X}}_{(1)} \\ \hline \bar{\boldsymbol{X}}_{(2)} \end{array}\right], \quad j = 1, 2, \ldots, n.$$

Posteriormente sorteou-se uma amostra *bootstrap* com reposição da amostra original dos $\boldsymbol{e}_j$'s, ou seja, $\boldsymbol{X}^*_1,\ \boldsymbol{X}^*_2,\ \cdots,\ \boldsymbol{X}^*_j,\ \cdots,\ \boldsymbol{X}^*_n$, considerando $\boldsymbol{X}^*_j$ de dimensão $(2p \times 1)$, $j = 1, 2, \cdots, n$. Para a imposição da hipótese nula $H_0 : \boldsymbol{\Sigma}_{11} = \boldsymbol{\Sigma}_{22}$, quando $\boldsymbol{\Sigma}_{12} \neq 0$ considerou-se a partição $\boldsymbol{X}^*_j = \left[\boldsymbol{X}^{*\top}_{j(1)} \,|\, \boldsymbol{X}^{*\top}_{j(2)}\right]^{\top}_{(2p\times 1)}$. A hipótese nula é imposta reamostrando ao acaso cada partição aleatoriamente para compor a amostra de qualquer um dos grupos. A probabilidade de que uma das partições de um dos grupos seja sorteada para o outro foi de 50%. Logo, a amostra final de *bootstrap* será dada por $\boldsymbol{X}^+_j$ $(2p \times 1)$, $j = 1, 2, \cdots, n$, definido por

$$\begin{aligned} \boldsymbol{X}^+_j &= \left[\begin{array}{c} \boldsymbol{X}^*_{j(2)} \\ \hline \boldsymbol{X}^*_{j(1)} \end{array}\right], \quad \text{com probabilidade 50\% ou} \\ \boldsymbol{X}^+_j &= \boldsymbol{X}^*_j, \quad \text{com probabilidade 50\%.} \end{aligned}$$

De posse da amostra $\boldsymbol{X}_1^+, \boldsymbol{X}_2^+, \cdots, \boldsymbol{X}_j^+, \cdots, \boldsymbol{X}_n^+$ calculou-se $\boldsymbol{S}^+$, que é a matriz de covariâncias amostrais *bootstrap*. Posteriormente, tomaram-se as partições $\boldsymbol{S}_{11}^+$ e $\boldsymbol{S}_{22}^+$ e calculou-se $\hat{\boldsymbol{H}}^+ = \hat{\boldsymbol{\Omega}}^{-1}(\boldsymbol{S}_{11}^+ - \boldsymbol{S}_{22}^+)(\hat{\boldsymbol{\Omega}}^\top)^{-1}$, sendo a matriz $\hat{\boldsymbol{\Omega}}$ o fator de Cholesky de $(\boldsymbol{S}_{11}^+ + \boldsymbol{S}_{22}^+)$. Assim, o máximo autovalor de $\hat{\boldsymbol{H}}^+$ é dado por $\hat{\rho}_{UV}^+ = \hat{\lambda}_1^+$. Dessa forma, obteve-se a estatística

$$t_{b_k} = \frac{\hat{\rho}_{UV}^+ \sqrt{n-2}}{\sqrt{1 - (\hat{\rho}_{UV}^+)^2}}, \tag{6}$$

para a k-ésima amostra de *bootstrap*, de acordo com o proposto por Hall (1990) e Hall e Wilson (1991).

Essa estatística tem distribuição t de *Student* com $\nu = n-2$ graus de liberdade, se $\rho_{UV} = 0$. Em seguida repetiram-se os passos de (i) a (iv), fazendo k variar de 1 a B. Os B valores da estatística t_{b_k} foram armazenados juntamente com o valor na amostra original t_{b_0}, na posição $B+1$. Dessa forma, o valor-p definido por Manly (1997) foi calculado por:

$$\text{valor-}p = \frac{\displaystyle\sum_{k=1}^{B+1} I(t_{b_k} \geq t_{b_0})}{B+1},$$

em que $I(t_{b_k} \geq t_{b_0})$ é a função indicadora, que retorna 1 se $t_{b_k} \geq t_{b_0}$, e 0 caso contrário. Retornou-se o valor da estatística t_{b_0} e o valor-p. A decisão de rejeição da hipótese foi obtida confrontando-se o valor-p e o nível nominal de significância α. Se o valor-$p \leq \alpha$ rejeita-se H_0.

2.1 Simulação Monte Carlo

Para avaliar o desempenho do teste proposto (t_{b_0}) foram realizadas simulações Monte Carlo. A rotina com os comandos do software R (R Development core team, 2014) para sua aplicação está apresentada no Apêndice A. Foram avaliadas as taxas de erro tipo I e poder do desempenho do teste t_{b_0}. Foram simuladas amostras aleatórias normais multivariadas, de tamanho n, dadas por $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_j, \cdots, \boldsymbol{X}_n$, em que $\boldsymbol{X}_j \in \mathbb{R}^{2p}$, com vetor média $\boldsymbol{\mu}_{(2p\times 1)}$ e covariância $\boldsymbol{\Sigma}_{(2p\times 2p)}$ positiva definida, $\boldsymbol{X} \sim N_{2p}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

O processo de simulação de Monte Carlo foi repetido 10000 vezes e o teste acima foi aplicado em cada amostra. Os níveis de significância nominais, α, foram fixados em 1% e 5%. As taxas de erro tipo I foram calculadas para cada teste, como a proporção de vezes que a hipótese de nulidade foi (falsamente) rejeitada, e esta foi comparada com o nível de significância nominal. Uma vez que estas taxas de erro tipo I foram estimadas usando simulações de Monte Carlo, elas não são livres de erro.

Os testes binomiais exatos foram aplicados usando o software Sisvar, criado por Ferreira (2011), considerando um nível de significância nominal de 1% para a hipótese $H_0 : \alpha = 5\%$ contra $H_1 : \alpha \neq 5\%$ e $H_0 : \alpha = 1\%$ versus $H_1 : \alpha \neq 1\%$

são aplicados para as taxas de erro de tipo I. Se a hipótese nula for rejeitada e as taxas de erro tipo I observadas forem consideradas de forma significativa ($p < 0,01$) inferior ao nível nominal, o teste deverá ser considerado conservativo; se as taxas de erro tipo I observados forem consideradas significativamente ($p < 0,01$) superior à do nível nominal, o teste deverá ser considerado liberal; e se as taxas de erro tipo I observadas não foram significativamente ($p > 0,01$) diferente do nível nominal, o teste deverá ser considerado exato (OLIVEIRA e FERREIRA, 2010).

Se y representa o número de hipóteses nulas rejeitadas em $N = 10000$ simulações de Monte Carlo para o nível de significância α nominal, então obtém-se a estatística do teste, utilizando-se a relação entre F e a distribuição binomial, com probabilidade de sucesso $p = \alpha$, dada por

$$F = \left(\frac{y+1}{N-y}\right)\left(\frac{1-\alpha}{\alpha}\right),$$

que, sob a hipótese nula, segue uma distribuição F com $\nu_1 = 2\,(N-y)$ e $\nu_2 = 2\,(y+1)$ graus de liberdade. Se $F \leq F_{0,005}$ ou $F \geq F_{0,995}$, a hipótese nula deve ser rejeitada ao nível de significância de 1%, onde $F_{0,005}$ e $F_{0,995}$ são quantis da distribuição F com ν_1 e ν_2 graus de liberdade (OLIVEIRA e FERREIRA, 2010).

Na simulação realizada neste trabalho e na apresentada por Jiang et al. (1999), para $p = 2$, foram considerados os mesmos valores para os parâmetros como no estudo de Jiang e Sarkar (1998), tais como: $n \in \{10, 15, 20, 25, 50, 75, 100\}$, para estimar o poder dos testes.

$$\boldsymbol{\Sigma}_1 = \boldsymbol{I}_4 = \begin{bmatrix} 1,0 & 0,0 & 0,0 & 0,0 \\ 0,0 & 1,0 & 0,0 & 0,0 \\ 0,0 & 0,0 & 1,0 & 0,0 \\ 0,0 & 0,0 & 0,0 & 1,0 \end{bmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1,0000 & -1,0000 & 0,7071 & -0,7071 \\ -1,0000 & 5,0000 & -0,7071 & 0,7071 \\ 0,7071 & -0,7071 & 1,0000 & -1,0000 \\ -0,7071 & 0,7071 & -1,0000 & 5,0000 \end{bmatrix},$$

$$\boldsymbol{\Sigma}_3 = \begin{bmatrix} 1,0 & 0,1 & 0,2 & 0,3 \\ 0,1 & 1,0 & 0,4 & 0,5 \\ 0,2 & 0,4 & 1,0 & 0,1 \\ 0,3 & 0,5 & 0,1 & 1,0 \end{bmatrix}, \quad \boldsymbol{\Sigma}_4 = \begin{bmatrix} 1,0000 & -1,0000 & 0,7071 & -1,4142 \\ -1,0000 & 5,0000 & -0,7071 & 1,4142 \\ 0,7071 & -0,7071 & 1,0000 & -1,0000 \\ -1,4142 & 1,4142 & -1,0000 & 5,0000 \end{bmatrix},$$

para estimar as taxas de erro tipo I, e

$$\boldsymbol{\Sigma}_5 = \begin{bmatrix} 1,0 & 0,1 & 0,2 & 0,3 \\ 0,1 & 1,0 & 0,3 & 0,5 \\ 0,2 & 0,3 & 1,0 & 0,6 \\ 0,3 & 0,5 & 0,6 & 1,0 \end{bmatrix}, \boldsymbol{\Sigma}_6 = \begin{bmatrix} 1,0 & 0,1 & 0,2 & 0,3 \\ 0,1 & 1,0 & 0,3 & 0,5 \\ 0,2 & 0,3 & 1,0 & 0,6 \\ 0,3 & 0,5 & 0,6 & 3,0 \end{bmatrix},$$

$$\boldsymbol{\Sigma}_7 = \begin{bmatrix} 1,0 & 0,1 & 0,2 & 0,3 \\ 0,1 & 2,0 & 0,3 & 0,5 \\ 0,2 & 0,3 & 3,0 & 0,6 \\ 0,3 & 0,5 & 0,6 & 4,0 \end{bmatrix}, \boldsymbol{\Sigma}_8 = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \boldsymbol{\Sigma}_9 = \begin{bmatrix} 1,0 & 0,1 & 0,2 & 0,3 \\ 0,1 & 1,0 & 0,4 & 0,5 \\ 0,2 & 0,4 & 1,0 & 0,6 \\ 0,3 & 0,5 & 0,6 & 1,0 \end{bmatrix},$$

$$\boldsymbol{\Sigma}_{10} = \begin{bmatrix} 1,0 & 0,1 & 0,2 & 0,3 \\ 0,1 & 1,0 & 0,4 & 0,5 \\ 0,2 & 0,4 & 1,0 & 0,6 \\ 0,3 & 0,5 & 0,6 & 3,0 \end{bmatrix}, \boldsymbol{\Sigma}_{111} = \begin{bmatrix} 1,0 & 0,1 & 0,2 & 0,3 \\ 0,1 & 2,0 & 0,4 & 0,5 \\ 0,2 & 0,4 & 3,0 & 0,6 \\ 0,3 & 0,5 & 0,6 & 4,0 \end{bmatrix},$$

Em todos os casos foram simuladas $N = 10000$ amostras e a proporção de rejeição da hipótese nula H_0 foi computada para cada teste ao longo de todas as simulações Monte Carlo. Os valores da taxa de erro tipo I empírica foram comparados com o valor nominal fixado em cada um dos casos. O teste foi aplicado a cada uma dessas amostras em cada uma das configurações formadas pela combinação de $\boldsymbol{\Sigma}$ e n, considerando o nível de significância (α) de 0,05. Posteriormente foram construídas matrizes de covariâncias positivas definidas, considerando $p = 4$ e $p = 10$, com o intuito de avaliar as taxas de erro tipo I e poder do teste t_{b_0}. Jiang e Sarkar (1998) e Jiang et al. (1999) não apresentaram resultados de simulação Monte Carlo para $p > 2$. Portanto, não foi possível comparar os resultados de simulação Monte Carlo do teste t_{b_0} com os dos testes propostos por Jiang e Sarkar (1998) e Jiang et al. (1999) quando $p = 4$ e $p = 10$. Os valores de n considerados na simulação para $p = 4$ e $p = 10$ foram: $n \in \{20, 25, 50, 75, 100, 300, 500, 1000\}$.

Matrizes de covariâncias positivas definidas, com $p = 4$ e $p = 10$, foram construídas para avaliar o desempenho das taxas de erro tipo I e poder do teste t_{b_0}.

Para avaliar as taxas de erro tipo I do teste t_{b_0} construiu-se matrizes de covariâncias $\boldsymbol{\Sigma}$ iguais a matrizes de correlações populacionais ($\boldsymbol{\rho}$). Nesse caso, serão denominadas de ($\boldsymbol{\rho}$). A simulação foi feita sob a hipótese nula $H_0 : \boldsymbol{\Sigma}_{11} = \boldsymbol{\Sigma}_{22}$ e, considerou-se as correlações de $\boldsymbol{\Sigma}_{12}$ como sendo valores menores ou iguais (de forma decrescente na linha e decrescente na coluna) às correlações fora da diagonal principal de $\boldsymbol{\Sigma}_{11}$. As correlações da diagonal principal de $\boldsymbol{\Sigma}_{11}$ são todas iguais a 1. Como $\boldsymbol{\Sigma}$ é simétrica tem-se que $\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21}^{\top}$. Em cada simulação foram consideradas as configurações formadas pela combinação dos valores de $\boldsymbol{\rho}$, n e nível de significância (α) de 0,05.

Em todos os casos $N = 10000$ simulações Monte Carlo foram gerados e, portanto, 10000 valores-p. Com esses valores foi calculada a proporção de vezes em que H_0 foi erroneamente rejeitada, ou seja, a proporção de vezes que o valor-p foi menor que α.

A simulação do poder do teste proposto foi feita sob a hipótese alternativa $H_1 : \boldsymbol{\Sigma}_{11} \neq \boldsymbol{\Sigma}_{22}$. Para avaliar o desempenho do poder do teste t_{b_0} foram construídas matrizes de covariâncias ($\boldsymbol{\Sigma}$), e serão descritas a seguir. Inicialmente considerou-se uma matriz diagonal de desvios padrões σ_k, com $k = 1, 2, \cdots, 2p$, definida por $\boldsymbol{V}^{1/2}$, de forma que o máximo desvio padrão de $\boldsymbol{\Sigma}_{11}$ é menor que o mínimo desvio padrão de $\boldsymbol{\Sigma}_{22}$. A partir da matriz de correlação ($\boldsymbol{\rho}$) e da matriz diagonal de desvios padrões, construídas anteriormente, construiu-se a matriz de covariâncias, dada por $\boldsymbol{\Sigma} = \boldsymbol{V}^{1/2}\boldsymbol{\rho}\boldsymbol{V}^{1/2}$, que por sua vez é uma matriz positiva definida.

Em cada simulação foram consideradas as configurações formadas pela combinação dos valores de $\boldsymbol{\Sigma}$ e n e níveis de significância (α) de 0,05. Em todos os casos o número de simulações Monte Carlo foi $N = 10000$, ou seja, foram gerados 10000 valores-p. Com esses valores foi calculada a proporção de vezes em que H_0 foi acertadamente rejeitada, ou seja, a proporção de vezes que o valor-p foi menor que α.

Para facilitar optou-se em denominar as matrizes de correlações ($\boldsymbol{\rho}$) por índice par e, as matrizes de covariâncias ($\boldsymbol{\Sigma}$) por índice ímpar. As matrizes com índice par foram construídas para avaliar o desempenho das taxas de erro tipo I do teste t_{b_0}, e as matrizes com índice ímpar o desempenho do poder desse mesmo teste. Algumas matrizes de correlações, considerando $p = 4$, foram construídas considerando a seguinte estrutura de correlação:

$$\boldsymbol{\rho} = \left[\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{B} \\ \hline \boldsymbol{B} & \boldsymbol{A} \end{array}\right],$$

em que

$$\boldsymbol{A} = \begin{bmatrix} 1,0 & \rho & \rho & \rho \\ \rho & 1,0 & \rho & \rho \\ \rho & \rho & 1,0 & \rho \\ \rho & \rho & \rho & 1,0 \end{bmatrix}, \qquad \boldsymbol{B} = \begin{bmatrix} \rho^* & \rho^* & \rho^* & \rho^* \\ \rho^* & \rho^* & \rho^* & \rho^* \\ \rho^* & \rho^* & \rho^* & \rho^* \\ \rho^* & \rho^* & \rho^* & \rho^* \end{bmatrix},$$

e em módulo $\rho^* \leq \rho$.

A partir da construção da matriz de correlação $\boldsymbol{\rho}_{(8\times 8)}$ e da matriz diagonal dos desvios padrões $(\boldsymbol{V}^{1/2})_{(8\times 8)}$, definida em (7), determinou-se a matriz de covariância $\boldsymbol{\Sigma}_{(8\times 8)}$.

$$\boldsymbol{V}^{1/2} = diag\,(\sigma_1,\, \sigma_2,\, \sigma_3,\, \sigma_4,\, \sigma_5,\, \sigma_6,\, \sigma_7,\, \sigma_8)\,, \tag{7}$$

em que σ_1, σ_2, σ_3 e σ_4 são desvios padrões menores que σ_5, σ_6, σ_7 e σ_8.

Para os desvios padrões foram escolhidos valores arbitrários que satisfazem a restrição imposta. Abaixo seguem as matrizes de correlações e de covariâncias que foram construídas considerando $p = 4$.

A matriz $\boldsymbol{\rho}_{12}$ construída para avaliar o desempenho da taxa de erro tipo I do teste foi construída considerando $\rho = \rho^* = 0,9$. A partir da matriz de correlação $\boldsymbol{\rho}_{12}$ e da matriz diagonal dos desvios padrões definida por $\boldsymbol{V}^{1/2} = diag\,(4, 5, 2, 1, 16, 20, 30, 10)$ construiu-se a matriz de covariâncias $\boldsymbol{\Sigma}_{13}$, para avaliar o desempenho do poder do teste. Assim, $\boldsymbol{\Sigma}_{13} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{12}\,\boldsymbol{V}^{1/2}$. Considerou-se $\rho = \rho^* = 0,8$ para construir a matriz de correlação $\boldsymbol{\rho}_{14}$ e a matriz diagonal dos desvios padrões foi dada por $\boldsymbol{V}^{1/2} = diag\,(4,\ 5,\ 2,\ 1,\ 16,\ 20,\ 30,\ 10)$. Assim, a matriz de covariâncias $\boldsymbol{\Sigma}_{15}$ será definida por $\boldsymbol{\Sigma}_{15} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{14}\,\boldsymbol{V}^{1/2}$. Na construção de $\boldsymbol{\rho}_{16}$ considerou-se $\rho = \rho^* = 0,5$, e para determinar $\boldsymbol{\Sigma}_{17}$ considerou-se a matriz $\boldsymbol{V}^{1/2} = diag\,(3,\ 5,\ 2,\ 4,\ 16,\ 20,\ 40,\ 10)$. Logo, $\boldsymbol{\Sigma}_{17} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{16}\,\boldsymbol{V}^{1/2}$. Para construir $\boldsymbol{\rho}_{18}$ considerou $\rho = \rho^* = 0,1$, e para determinar $\boldsymbol{\Sigma}_{19}$ considerou-se $\boldsymbol{V}^{1/2} = diag\,(3,\ 5,\ 2,\ 4,\ 16,\ 20,\ 40,\ 10)$. Portanto, $\boldsymbol{\Sigma}_{19} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{18}\,\boldsymbol{V}^{1/2}$. Na construção de $\boldsymbol{\rho}_{20}$ considerou-se $\rho = \rho^* = 0$ e, para determinar $\boldsymbol{\Sigma}_{211}$ definiu-se $\boldsymbol{V}^{1/2} = diag\,(2,\ 5,\ 4,\ 6,\ 15,\ 12,\ 20,\ 10)$. Dessa forma, $\boldsymbol{\Sigma}_{211} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{20}\,\boldsymbol{V}^{1/2}$. Para determinar $\boldsymbol{\rho}_{22}$ considerou-se $\rho = 0,9$ e $\rho^* = 0,8$ e, definiu-se a matriz $\boldsymbol{\Sigma}_{23}$ a partir de $\boldsymbol{V}^{1/2} = diag\,(2,\ 5,\ 4,\ 6,\ 15,\ 12,\ 20,\ 10)$. Logo, $\boldsymbol{\Sigma}_{23} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{22}\,\boldsymbol{V}^{1/2}$. Para construir $\boldsymbol{\rho}_{24}$ considerou-se $\rho = 0,9$ e $\rho^* = 0,5$ e, determinou-se a matriz $\boldsymbol{\Sigma}_{25}$ a partir de $\boldsymbol{V}^{1/2} = diag\,(1,\ 5,\ 7,\ 3,\ 15,\ 24,\ 10,\ 30)$. Assim, $\boldsymbol{\Sigma}_{25} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{24}\,\boldsymbol{V}^{1/2}$. Na construção de $\boldsymbol{\rho}_{26}$ considerou-se $\rho = 0,9$ e $\rho^* = 0,1$. Em seguida determinou-se a matriz $\boldsymbol{\Sigma}_{27}$ a partir de $\boldsymbol{V}^{1/2} = diag\,(1,\ 5,\ 7,\ 3,\ 15,\ 24,\ 10,\ 30)$. Logo, $\boldsymbol{\Sigma}_{27} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{26}\,\boldsymbol{V}^{1/2}$. A matriz de correlação $\boldsymbol{\rho}_{28}$ foi construída considerando $\rho = 0,9$ e $\rho^* = 0$. Em seguida determinou-se a matriz de covariâncias $\boldsymbol{\Sigma}_{29}$ a partir de $\boldsymbol{V}^{1/2} = diag\,(5, 6, 2, 4, 12, 15, 25, 20)$. Dessa forma, $\boldsymbol{\Sigma}_{29} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{28}\,\boldsymbol{V}^{1/2}$. Na construção da matriz de correlação $\boldsymbol{\rho}_{30}$ considerou-se outra estrutura de correlação definida por:

$$\boldsymbol{\rho} = \left[\begin{array}{cccc|cccc} 1,0 & \rho & \rho & \rho & \rho_1^* & \rho_2^* & \rho_3^* & \rho_4^* \\ \rho & 1,0 & \rho & \rho & \rho_2^* & \rho_3^* & \rho_4^* & \rho_5^* \\ \rho & \rho & 1,0 & \rho & \rho_3^* & \rho_4^* & \rho_5^* & \rho_6^* \\ \rho & \rho & \rho & 1,0 & \rho_4^* & \rho_5^* & \rho_6^* & \rho_7^* \\ \hline \rho_1^* & \rho_2^* & \rho_3^* & \rho_4^* & 1,0 & \rho & \rho & \rho \\ \rho_2^* & \rho_3^* & \rho_4^* & \rho_5^* & \rho & 1,0 & \rho & \rho \\ \rho_3^* & \rho_4^* & \rho_5^* & \rho_6^* & \rho & \rho & 1,0 & \rho \\ \rho_4^* & \rho_5^* & \rho_6^* & \rho_7^* & \rho & \rho & \rho & 1,0 \end{array}\right],$$

de forma que $\rho_1^* = 0,7$, $\rho_2^* = 0,6$, $\rho_3^* = 0,5$, $\rho_4^* = 0,4$, $\rho_5^* = 0,3$, $\rho_6^* = 0,2$ e $\rho_7^* = 0,1$. Definiu-se a matriz $\boldsymbol{\Sigma}_{31}$ a partir de $\boldsymbol{\rho}_{30}$ e $\boldsymbol{V}^{1/2} = diag\,(5,\ 6,\ 2,\ 4,\ 12,\ 15,\ 25,\ 20)$. Dessa forma, $\boldsymbol{\Sigma}_{31} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{30}\,\boldsymbol{V}^{1/2}$. Na construção de algumas matrizes de correlações ($\boldsymbol{\rho}$) de dimensão (20×20), para avaliar o desempenho das taxas de erro tipo I, considerou-se a seguinte estrutura de correlação:

$$\boldsymbol{\rho} = \left[\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{B} \\ \hline \boldsymbol{B} & \boldsymbol{A} \end{array}\right],$$

em que

$$\boldsymbol{A} = \begin{bmatrix} 1,0 & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho \\ \rho & 1,0 & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho \\ \rho & \rho & 1,0 & \rho & \rho & \rho & \rho & \rho & \rho & \rho \\ \rho & \rho & \rho & 1,0 & \rho & \rho & \rho & \rho & \rho & \rho \\ \rho & \rho & \rho & \rho & 1,0 & \rho & \rho & \rho & \rho & \rho \\ \rho & \rho & \rho & \rho & \rho & 1,0 & \rho & \rho & \rho & \rho \\ \rho & \rho & \rho & \rho & \rho & \rho & 1,0 & \rho & \rho & \rho \\ \rho & \rho & \rho & \rho & \rho & \rho & \rho & 1,0 & \rho & \rho \\ \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & 1,0 & \rho \\ \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & 1,0 \end{bmatrix},$$

$$\boldsymbol{B} = \begin{bmatrix} \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* \\ \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* \\ \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* \\ \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* \\ \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* \\ \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* \\ \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* \\ \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* \\ \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* \\ \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* & \rho^* \end{bmatrix},$$

e em módulo $\rho^* \leq \rho$.

A partir da matriz de correlação $\boldsymbol{\rho}_{(20\times 20)}$ e de uma matriz diagonal dos desvios padrões $(\boldsymbol{V}^{1/2})_{(20\times 20)}$, definida por (8), construiu-se a matriz de covariâncias $\boldsymbol{\Sigma}_{(20\times 20)}$.

$$\begin{aligned} \boldsymbol{V}^{1/2} &= diag\,(\sigma_1,\ \sigma_2,\ \sigma_3,\ \sigma_4,\ \sigma_5,\ \sigma_6,\ \sigma_7,\ \sigma_8,\ \sigma_9,\ \sigma_{10},\ \sigma_{11},\ \sigma_{12},\ \sigma_{13},\ \sigma_{14}, \sigma_{15}, \sigma_{16},\ \sigma_{17},\ \sigma_{18},\ \sigma_{19},\ \sigma_{20}\,) \\ &= diag\,(\,4,\ 5,\ 2,\ 1,\ 16,\ 20,\ 30,\ 10,\ 12,\ 16,\ 40,\ 45,\ 60,\ 55,\ 41,\ 39,\ 35,\ 65,\ 100,\ 50\,). \end{aligned} \tag{8}$$

Os valores dos desvios padrões da matriz $\boldsymbol{V}^{1/2}$, para $p = 10$, foram escolhidos arbitrariamente, de forma que satisfazem a restrição imposta: σ_1, σ_2, σ_3, σ_4, σ_5, σ_6, σ_7, σ_8, σ_9 e σ_{10} são desvios padrões menores que σ_{11}, σ_{12}, σ_{13}, σ_{14}, σ_{15}, σ_{16}, σ_{17}, σ_{18}, σ_{19} e σ_{20}.

Abaixo seguem as matrizes de correlações e de covariâncias que foram construídas, considerando $p = 10$. Na construção de todas as matrizes de covariâncias ($\boldsymbol{\Sigma}$) considerou-se a matriz diagonal dos desvios padrões definida em (8). Na construção de $\boldsymbol{\rho}_{32}$ considerou-se $\rho = \rho^* = 0,9$. Logo, $\boldsymbol{\Sigma}_{33} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{32}\,\boldsymbol{V}^{1/2}$. Na construção das matrizes de correlações $\boldsymbol{\rho}_{34}$, $\boldsymbol{\rho}_{36}$, $\boldsymbol{\rho}_{38}$, $\boldsymbol{\rho}_{40}$, $\boldsymbol{\rho}_{42}$, $\boldsymbol{\rho}_{44}$, $\boldsymbol{\rho}_{46}$ e $\boldsymbol{\rho}_{48}$ considerou-se $\rho = \rho^* = 0,8$; $\rho = \rho^* = 0,5$; $\rho = \rho^* = 0,1$; $\rho = \rho^* = 0$; $\rho = 0,9$ e $\rho^* = 0,8$; $\rho = 0,9$ e $\rho^* = 0,5$; $\rho = 0,9$ e $\rho^* = 0,1$; e, $\rho = 0,9$ e $\rho^* = 0$; respectivamente.

Na construção da matriz de correlação $\boldsymbol{\rho}_{50}$ considerou-se outra estrutura de correlação definida por:

$$\boldsymbol{\rho} = \left[\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{B} \\ \hline \boldsymbol{B} & \boldsymbol{A} \end{array}\right],$$

em que

$$\boldsymbol{A} = \begin{bmatrix} 1,0 & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho \\ \rho & 1,0 & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho \\ \rho & \rho & 1,0 & \rho & \rho & \rho & \rho & \rho & \rho & \rho \\ \rho & \rho & \rho & 1,0 & \rho & \rho & \rho & \rho & \rho & \rho \\ \rho & \rho & \rho & \rho & 1,0 & \rho & \rho & \rho & \rho & \rho \\ \rho & \rho & \rho & \rho & \rho & 1,0 & \rho & \rho & \rho & \rho \\ \rho & \rho & \rho & \rho & \rho & \rho & 1,0 & \rho & \rho & \rho \\ \rho & \rho & \rho & \rho & \rho & \rho & \rho & 1,0 & \rho & \rho \\ \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & 1,0 & \rho \\ \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & \rho & 1,0 \end{bmatrix}$$

e

$$\boldsymbol{B} = \begin{bmatrix} \rho_1^* & \rho_2^* & \rho_3^* & \rho_4^* & \rho_5^* & \rho_6^* & \rho_7^* & \rho_8^* & \rho_9^* & \rho_{10}^* \\ \rho_2^* & \rho_3^* & \rho_4^* & \rho_5^* & \rho_6^* & \rho_7^* & \rho_8^* & \rho_9^* & \rho_{10}^* & \rho_{11}^* \\ \rho_3^* & \rho_4^* & \rho_5^* & \rho_6^* & \rho_7^* & \rho_8^* & \rho_9^* & \rho_{10}^* & \rho_{11}^* & \rho_{12}^* \\ \rho_4^* & \rho_5^* & \rho_6^* & \rho_7^* & \rho_8^* & \rho_9^* & \rho_{10}^* & \rho_{11}^* & \rho_{12}^* & \rho_{13}^* \\ \rho_5^* & \rho_6^* & \rho_7^* & \rho_8^* & \rho_9^* & \rho_{10}^* & \rho_{11}^* & \rho_{12}^* & \rho_{13}^* & \rho_{14}^* \\ \rho_6^* & \rho_7^* & \rho_8^* & \rho_9^* & \rho_{10}^* & \rho_{11}^* & \rho_{12}^* & \rho_{13}^* & \rho_{14}^* & \rho_{15}^* \\ \rho_7^* & \rho_8^* & \rho_9^* & \rho_{10}^* & \rho_{11}^* & \rho_{12}^* & \rho_{13}^* & \rho_{14}^* & \rho_{15}^* & \rho_{16}^* \\ \rho_8^* & \rho_9^* & \rho_{10}^* & \rho_{11}^* & \rho_{12}^* & \rho_{13}^* & \rho_{14}^* & \rho_{15}^* & \rho_{16}^* & \rho_{17}^* \\ \rho_9^* & \rho_{10}^* & \rho_{11}^* & \rho_{12}^* & \rho_{13}^* & \rho_{14}^* & \rho_{15}^* & \rho_{16}^* & \rho_{17}^* & \rho_{18}^* \\ \rho_{10}^* & \rho_{11}^* & \rho_{12}^* & \rho_{13}^* & \rho_{14}^* & \rho_{15}^* & \rho_{16}^* & \rho_{17}^* & \rho_{18}^* & \rho_{19}^* \end{bmatrix},$$

de forma que $\rho = 0,8$, $\rho_1^* = 0,4$, $\rho_2^* = 0,38$, $\rho_3^* = 0,37$, $\rho_4^* = 0,36$, $\rho_5^* = 0,35$, $\rho_6^* = 0,34$, $\rho_7^* = 0,29$, $\rho_8^* = 0,28$, $\rho_9^* = 0,25$, $\rho_{10}^* = 0,23$, $\rho_{11}^* = 0,18$, $\rho_{12}^* = 0,15$, $\rho_{13}^* = 0,13$, $\rho_{14}^* = 0,12$, $\rho_{15}^* = 0,1$, $\rho_{16}^* = 0,09$, $\rho_{17}^* = 0,08$, $\rho_{18}^* = 0,07$ e $\rho_{19}^* = 0,05$.

A partir de $\boldsymbol{\rho}_{50}$ determinou-se $\boldsymbol{\Sigma}_{51}$, considerando $\boldsymbol{V}^{1/2}$, definida em (8). Dessa forma, $\boldsymbol{\Sigma}_{51} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{50}\,\boldsymbol{V}^{1/2}$. Considerou-se a mesma estrutura de correlação usada em $\boldsymbol{\rho}_{50}$ para construir a matriz de correlação $\boldsymbol{\rho}_{52}$, em que: $\rho = 0,5$, $\rho_1^* = 0,4$, $\rho_2^* = 0,38$, $\rho_3^* = 0,37$, $\rho_4^* = 0,36$, $\rho_5^* = 0,35$, $\rho_6^* = 0,34$, $\rho_7^* = 0,29$, $\rho_8^* = 0,28$, $\rho_9^* = 0,25$, $\rho_{10}^* = 0,23$, $\rho_{11}^* = 0,18$, $\rho_{12}^* = 0,15$, $\rho_{13}^* = 0,13$, $\rho_{14}^* = 0,12$, $\rho_{15}^* = 0,1$, $\rho_{16}^* = 0,09$, $\rho_{17}^* = 0,08$, $\rho_{18}^* = 0,07$ e $\rho_{19}^* = 0,05$. A partir de $\boldsymbol{\rho}_{52}$ determinou-se $\boldsymbol{\Sigma}_{53}$, considerando $\boldsymbol{V}^{1/2}$, definida em (8). Assim, $\boldsymbol{\Sigma}_{53} = \boldsymbol{V}^{1/2}\,\boldsymbol{\rho}_{52}\,\boldsymbol{V}^{1/2}$.

3 Resultados de simulação Monte Carlo

Para avaliar o desempenho dos testes, as taxas de erro tipo I e poder foram computadas e são apresentadas e discutidas separadamente nas subseções 3.1 e 3.2.

3.1 Erro tipo I

De acordo com Oliveira e Ferreira (2010), o tamanho do teste é uma característica fundamental para o seu desempenho. Se o teste tiver tamanho real equivalente ao tamanho nominal, α, então é considerado exato. Testes que têm um tamanho real menor do que o nominal são considerados conservativos, e aqueles que apresentam taxas de erro tipo I maiores do que os níveis nominais α são considerados liberais.

Pode-se observar que para $N = 10000$ simulações e $\alpha = 5\%$, valores inferiores a 445 e superiores a 559 rejeições levam a rejeição da hipótese $H_0 : \alpha = 5\%$, considerando o nível de significância de 1% para o teste. E valores inferiores a 76 e superiores a 129 levam à rejeição da hipótese nula $H_0 : \alpha = 1\%$, considerando o mesmo nível de significância de 1% para o teste realizado. Na Tabela 1 - Anexo estão apresentadas, em porcentagem, as taxas de erro tipo I de sete testes de igualdade de matrizes de covariâncias com dimensão 4×4: t_{b_0}, LRT, LRT_1, LRT_2, LRT_3, W_2 e W_5, obtidas em 10000 simulações de Monte Carlo, em função de (n) e $\boldsymbol{\Sigma}_{(4\times4)}$ e $\alpha = 5\%$.

Pode-se verificar que os testes t_{b_0}, LRT_3 e W_2 controlaram o erro tipo I em nível inferior ou no máximo igual ao valor nominal de 5%. Em todos os casos os testes t_{b_0}, LRT_3 e W_2 foram conservativos em algumas situações. Como pode ser visto na Tabela 1 - Anexo, os testes t_{b_0}, LRT_3 e W_2 se sobressaem em relação aos demais, por controlar o erro tipo I, mesmo quando $n = 10$. Os testes LRT, LRT_1, LRT_2 e W_5 mostraram-se liberais em alguns casos. Apesar disso, segundo Jiang et al. (1999), os testes LRT_2 e W_5 podem ser recomendados quando $n \geq 20$, com LRT_2 tendo uma ligeira vantagem sobre W_5. Para $n = 100$, observou-se que, em todas as configurações de $\boldsymbol{\Sigma}$ avaliadas, os tamanhos dos testes não foram significativamente diferentes dos valores nominais de 5%. No geral, os testes LRT, LRT_1 LRT_2 e W_5 foram liberais quando $n \leq 75$, $n \leq 50$, $n \leq 20$ e $n \leq 25$, respectivamente.

Por fim, o teste *bootstrap* não-paramétrico t_{b_0} foi conservativo para amostras pequenas $(n = 25)$ e tendeu a ser exato para amostras maiores $(n = 75)$.

Na Tabela 2 - Anexo estão apresentados, em porcentagem, as taxas de erro tipo I do teste t_{b_0}. Esse teste foi avaliado em função do tamanho da amostra $n \in \{20, 25, 50, 75, 100, 300, 500 \text{ e } 1000\}$, e das matrizes de correlações para $p = 4$: $\boldsymbol{\rho}_{12}$, $\boldsymbol{\rho}_{14}$, $\boldsymbol{\rho}_{16}$, $\boldsymbol{\rho}_{18}$, $\boldsymbol{\rho}_{20}$, $\boldsymbol{\rho}_{22}$, $\boldsymbol{\rho}_{24}$, $\boldsymbol{\rho}_{26}$, $\boldsymbol{\rho}_{28}$ e $\boldsymbol{\rho}_{30}$, resultantes de $N = 10000$ simulações Monte Carlo, considerando o nível nominal de significância de 5%. Pode-se observar que para amostras menores ou iguais a 100 o teste *bootstrap* não-paramétrico t_{b_0} apresentou taxas significativamente menores do que 5%, indicando um teste conservativo. Para $n \geq 300$, o teste t_{b_0} apresentou taxas de erro tipo I iguais ao valor nominal, o que indica um teste exato.

Na Tabela 3 - Anexo estão apresentados, em porcentagem, os resultados de erro tipo I do teste de igualdade de matrizes de covariâncias t_{b_0}, considerando diferentes tamanhos amostrais $n \in \{20{,}25{,}50{,}75{,}100{,}300{,}500 \text{ e } 1000\}$, diferentes matrizes de correlações com $p = 10$: $\boldsymbol{\rho}_{32}$, $\boldsymbol{\rho}_{34}$, $\boldsymbol{\rho}_{36}$, $\boldsymbol{\rho}_{38}$, $\boldsymbol{\rho}_{40}$, $\boldsymbol{\rho}_{42}$, $\boldsymbol{\rho}_{44}$, $\boldsymbol{\rho}_{46}$, $\boldsymbol{\rho}_{48}$, $\boldsymbol{\rho}_{50}$ e $\boldsymbol{\rho}_{52}$ e valor nominal de significância de 5%. É possível observar que o teste t_{b_0} foi conservativo quando $n \leq 300$ e tendeu a ser exato quando $n \geq 500$.

3.2 Poder

O poder de um teste é a probabilidade $(1 - \beta)$ de rejeitar H_0 quando ela é falsa. Costa (1992) *apud* Gebert (2014) define o poder como sendo a sensibilidade da região crítica para perceber e rejeitar uma hipótese falsa.

Na Tabela 4 - Anexo estão apresentados, em porcentagem, os valores de poder de sete testes de igualdade de matrizes de covariâncias: t_{b_0}, LRT, LRT_1, LRT_2, LRT_3, W_2, W_5, em função de n, $p = 2$ e $\alpha = 5\%$. Os testes t_{b_0}, LRT_3 e W_2 apresentaram controle adequado do erro tipo I. Como consequência disso, espera-se que os testes que são conservativos (t_{b_0}, LRT_3 e W_2) sejam menos poderosos que os considerados liberais. Isso realmente aconteceu, como pode ser visto na Tabela 4 - Anexo. Como era de se esperar, o poder de todos os testes aumentou quando o tamanho amostral (n) e as taxas de erro tipo I foram aumentadas. Para grandes amostras ($n = 100$), o desempenho dos testes se igualou e seus valores de poder se aproximaram de 100%.

De acordo com Jiang et al. (1999), no caso de duas normais bivariadas dependentes, o teste da razão de verossimilhanças modificado LRT_2 e o teste assintótico W_5 são recomendados para serem usados em amostras de tamanhos moderados ou grandes $n \geq 20$, enquanto que os testes LRT_3 e W_2 podem ser aplicados para pequenas amostras. Pode-se verificar que em alguns casos de pequenas amostras ($n \leq 25$) o teste W_5 é mais poderoso do que W_2, e para amostras maiores ($n \geq 75$) os dois testes mantiveram o mesmo poder. Segundo Jiang e Sarkar (1998) isso se deve ao fato de W_5 tornar-se cada vez mais poderoso quando a diferença entre $\mathbf{\Sigma}_{11}$ e $\mathbf{\Sigma}_{22}$ torna-se mais proeminente. Dentre os testes apresentados por Jiang et al. (1999), o teste LRT_2 supera os demais por ter poder elevado e controlar a taxa de erro tipo I para amostras de tamanho $n \geq 20$ de duas populações normais bivariadas dependentes.

Os resultados de poder do teste t_{b_0}, considerando as matrizes de covariâncias ($\mathbf{\Sigma}_{13}, \mathbf{\Sigma}_{15}, \mathbf{\Sigma}_{17}, \mathbf{\Sigma}_{19}, \mathbf{\Sigma}_{211}, \mathbf{\Sigma}_{23}, \mathbf{\Sigma}_{25}, \mathbf{\Sigma}_{27}$, $\mathbf{\Sigma}_{29}$ e $\mathbf{\Sigma}_{31}$), $n \in \{$ 20, 25, 50, 75, 100, 300, 500 e 1000$\}$, com $p = 4$ e $\alpha = 5\%$ foram bastante elevados, sendo superiores a 98%, mesmo para amostras pequenas ($n = 15$).

Os valores do poder do teste t_{b_0}, considerando diferentes matrizes de covariâncias com $p = 10$, diferentes tamanhos amostrais (n) e $\alpha = 5\%$, se aproximaram de 100% em todas as configurações avaliadas, mesmo para pequenas amostras ($n = 20$).

4 Considerações gerais

Considerando as matrizes de covariâncias quando $p = 2$, pode-se afirmar que os testes t_{b_0}, LRT_3 e W_5 controlaram o erro tipo I em todos os casos analisados. O poder do teste t_{b_0} se aproximou com o dos testes LRT_3 e W_5 quando $n \geq 75$. Na maioria dos casos, o teste *bootstrap* não paramétrico t_{b_0} foi conservativo quando $n \leq 50$ e exato para $n > 50$.

Nas simulações para $p = 4$, o teste t_{b_0} foi conservativo quando $n \leq 100$, exceto para $\boldsymbol{\rho}_{16}, \boldsymbol{\rho}_{26}$ e $\boldsymbol{\rho}_{28}$, que foi conservativo quando $n \leq 300$; e exato para $n > 100$, com exceção dos casos das matrizes de correlações $\boldsymbol{\rho}_{16}, \boldsymbol{\rho}_{26}$ e $\boldsymbol{\rho}_{28}$, que foi exato quando $n > 300$. Analisando as matrizes de correlações quando $p = 10$, observa-se que o teste t_{b_0} foi conservativo quando $n \leq 1000$; exceto para $\boldsymbol{\rho}_{34}$, $\boldsymbol{\rho}_{38}$, $\boldsymbol{\rho}_{40}$, $\boldsymbol{\rho}_{44}$ e $\boldsymbol{\rho}_{52}$, que foi conservativo quando $n \leq 500$, e exato nesses casos de matrizes de correlações citadas acima, quando $n = 1000$. Comparando o teste t_{b_0} com todos os testes apresentados por Jiang e Sarkar (1998) e Jiang et al. (1999), pode-se afirmar que em relação ao erro tipo I eles foram tão bons quanto ou melhores. E nas situações não avaliadas por Jiang e Sarkar (1998) e Jiang et al. (1999), ($p = 4$ e $p = 10$), o teste t_{b_0} controlou o erro tipo I em todos os casos analisados. Nas situações comparáveis, o poder do teste t_{b_0} se aproximou, em alguns casos, dos testes apresentados por Jiang e Sarkar (1998) e Jiang et al. (1999), sendo superior a 95%, quando $n = 100$.

De uma maneira geral, o que se pode dizer sobre as taxas de erro tipo I é que o aumento do número de variáveis não acarretou variação nos valores das taxas de erro tipo I. Em relação ao poder, o aumento do número de variáveis provocou um aumento expressivo no poder, quase sempre igual a 100%, mesmo para amostras pequenas $n = 20$.

Os resultados de simulação Monte Carlo, obtidos por Jiang e Sarkar (1998) e Jiang et al. (1999), se limitaram a apenas duas variáveis. Portanto, para um número maior de variáveis não se viram relatos na literatura que possam ser comparados com os resultados do teste t_{b_0}.

Conclusões

Neste trabalho os objetivos propostos foram alcançados e obtiveram-se as seguintes conclusões:

Os testes t_{b_0}, LRT_3 e W_2 controlaram em todas as situações o erro tipo I, em níveis iguais ou inferiores aos valores nominais de significância, e apresentaram desempenho superior aos testes LRT, LRT_1, LRT_2 e W_5, que são liberais nas situações de pequenas amostras. Nas situações em que foi possível comparar o teste proposto neste trabalho com os apresentados por Jiang e Sarkar (1998) (W_2 e W_5) e Jiang et al. (1999) (LRT, LRT_1, LRT_2 e LRT_3), pode-se dizer que em alguns casos os resultados de poder do teste t_{b_0} foi superior a 95% e dos testes LRT_3 e W_2 superiores a 97%. Portanto, pode-se afirmar que dentre os testes que controlaram o erro tipo I, para $p = 2$, o teste t_{b_0} se aproximou dos testes LRT_3 e W_2, que foram superiores aos seus competidores em todas as situações estudadas.

Para os casos em que não foi possível fazer comparação ($p = 4$ e $p = 10$), o teste t_{b_0} apresentou um ótimo desempenho, com marcas de quase sempre 100% para $n \geq 20$. Portanto, recomenda-se a aplicação do teste t_{b_0} em situações reais.

SILVA, V. S. P.; FERREIRA, D. F. Nonparametric bootstrap test for equal covariance matrices of two dependent multivariate normal populations. *Rev. Bras. Biom.,* Lavras, v.34, n.2, p.210-232, 2016.

- *ABSTRACT: This study aims to evaluate the type I error rates and power of the nonparametric bootstrap test (t_{b_0}) for equality of covariance matrices of two dependent multivariate normal populations in order to compare its performance with the test presented by Jiang and Sarkar (1998) (W_2 and W_5) and Jiang et al. (1999) (LRT, LRT_1, LRT_2 and LRT_3). For this simulations Monte Carlo were performed, considering the number of variables (p), sample sizes (n), covariance matrices ($\boldsymbol{\Sigma}$) and significance level (α) of 0.05. In the first case, for $p = 2$, it was concluded that among the tests that controlled the type I error, the tests t_{b_0}, LRT_3 and W_2 were greater than its competitors* *in all cases studied. In relation to power, the test t_{b_0} approached the testing LRT_3 and W_2 and is considered intermediate. In the second case, it is considered that $p = 4$ and $p = 10$, it was concluded that the test t_{b_0} showed high performance, in most cases equal to 100% even for small samples ($n = 20$). Therefore, we recommend the application of the proposed test t_{b_0} in real situations.*

- *KEYWORDS: Covariance matrices; bootstrap test; Monte Carlo simulation; power; type I error.*

References

ANDERSON,T.W. *Maximum likelihood estimation for vector autorregressive moving average models*. California: Stanford University, 1978. 16p.

BARTLETT, M. S. Properties of sufficiency and statistical tests. *Proceedings of the Royal Society. Series A*, London, v.160, n.901, p.268–282, May 1937.

BHATTACHARJEE, G. P.; ALAM, S. S. Testing equality of variances after a preliminary test of significance for correlation-coefficient. *Journal of Statistical Planning and Inference*, North-Holland, v.6, n.1, p.17–28, 1982.

BOGLE, W.; HSU, Y. S. Sample size determination in comparing two population variances with paired-data: application to bilirubin tests. *Biometrical Journal*, Berlin, v.44, n.5, p.594–602, July 2002.

CHU, S. S.; PILLAI, K. C. S. Some complex variable transformations and exact power comparisons of two-sided tests of equality of two Hermitian covariance matrices. *Journal of Statistical Planning and Inference*, North-Holland, v.4, n.3, p.267–290, 1980.

CIRILLO, M. A. ; FERREIRA, D. F. ; SAFADI, T. Comparação de matrizes de covariâncias de populações normais dependentes: um estudo de caso. *Ciência e Agrotecnologia*, Lavras, v.33, p.1741-1791, 2009.

COSTA, S. F. *Introdução ilustrada à estatística (com muito humor)*. São Paulo: Harbra, 1992. 303p.

EFRON, B. Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, Philadelphia, v.7, n.1, p.1–26, 1979.

FERREIRA, D. F. Sisvar: a computer statistical analysis system. *Ciência e Agrotecnologia*, Lavras, v.35, n.6, p.1039–1042, 2011.

FERREIRA, D. F. *Estatística computacional em java*. Lavras: Editora da UFLA, 2013. 695p.

GEBERT, D. M. P. *Uma solução via bootstrap paramétrico para o problema de Behrens-Fisher multivariado*. 2014. 127 p. Tese (Doutorado em Estatística e Experimentação Agropecuária) — Universidade Federal de Lavras, Lavras, 2014.

HALL, P. Two Guidelines for Bootstrap Hypothesis Testing. *Biometrics*, v. 47, p. 757–762, 1990.

HALL, P.; WILSON, S. R. On Bootstrap Hypothesis Testing. *Australian Journal Statistics*, v.32, n.2, p.177–190, 1991.

HAN, C. P. Testing homogeneity of a set of correlated variances. *Biometrika*, Oxford, v.55, n.2, p.317–326, July 1968.

HAYAKAWA, T. Normalizing and variance stabilizing transformations of multivariate statistics under an elliptical population. *Annals of the Institute of Statistical Mathematics. Part A*, Tokyo, v.39, n.1, p.299–306, 1987.

JIANG, G.; SARKAR, S. K. Some asymptotic tests for the equality of the covariance matrices of two dependent bivariate normals. *Biometrical Journal*, Berlin, v.40, n.2, p.205–225, 1998.

JIANG, G.; SARKAR, S. K. The likelihood ratio test for the homogeneity of the variances in a covariance matrix with block compound symmetry. *Communications in Statistics. Part A, Theory and Methods*, New York, v.29, n.5–6, p.1155–1178, 2000.

JIANG, G.; SARKAR, S. K. Combination tests for the equality of the covariance matrices of two dependent bivariate normal. *Journal of Statistical Computation and Simulation*, New York, v.72, n.6, p.495–505, 2002.

JIANG, G.; SARKAR, S. K.; HSUAN, F. A likelihood ratio test and its modifications for the homogeneity of the covariance matrices of dependent multivariate normals. *Journal of Statistical Planning and Inference*, North-Holland, v.81, n.1, p.95–111, Oct. 1999.

MANLY, B. F. J. *Randomization, bootstrap and Monte Carlo methods in biology.* 2nd ed. New Zealand: University of Otago, 1997. 356p.

MORGAN, W. A. A test for the significance of the difference between the two variances in a sample from a normal bivariate population. *Biometrika*, Oxford, v.31, n.1-2, p.13–19, July 1939.

OLIVEIRA, I. R. C.; FERREIRA, D. F. Multivariate extension of chi-squared univariate normality test. *Journal of Statistical Computation and Simulation*, New York, v.80, n.5, p.513–526, 2010.

PIEPHO, H. P. Tests for equality of dispersion in bivariate samples - Review and empirical comparison. *Journal of Statistical Computation and Simulation*, New York, v.56, n.4, p.353–372, 1997.

PITMAN, E. J. G. A note on normal correlation. *Biometrika*, Oxford, v.31, n.1-2, p.9–12, July 1939.

R CORE TEAM. *R*: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2014. Disponível em: <http://www.R-project.org>. Acesso: 14 mar. 2014.

SHAPIRO, A.; COHEN, A. Testing and estimation of equal variances for correlated variables. *Statistics and Probability Letters*, Amsterdam, v.10, p.231–234, Aug. 1990.

Recebido em 08.05.2015.

Aprovado após revisão em 20.02.2016.

APÊNDICE

APÊNDICE A: Rotina no R para aplicação do teste *bootstrap* não-paramétrico

```
# Algoritmo implementado em R para testar duas matrizes de covariâncias na
# presença de correlação
# A Função  recebe S_{11} e S_{22} e retorna
# o autovalor "lambda _ 1" e o vetor a do bootstrap não paramétrico

maxRhob <- function(S11, S22)
{
  Gt      <- chol(S11+S22)
  GtI     <- solve(Gt)
  H       <- t(GtI) %*% (S11-S22) %*% GtI
  aVaVe   <- eigen(H)
  rho_max <- aVaVe$values[1]
  return(rho_max)
}

# Função para obtenção da estatística t do teste: recebe a, S11, s12, S22,
# n e  retorna o tb do bootstrap

tTestb.rhomax <- function(s11, s22, s12, n)
{
  Rho_max <- maxRhob(s11, s22)
  tb      <- Rho_max*(n-2)^0.5/(1-Rho_max^2)^0.5
  return(tb)
}

# Função para trocar X's com Y's: recebe  o  vetor x (1 x 2p) (linha)
# a dimensão p (troca os p primeiros com os p's últimos)

trocaXY <- function(x, p)
{
  if (runif(1) <= 0.5) {
    y <- x[(p+1):(2*p)]
    y <- c(y, x[1:p])
  } else y <- x
  return(y)
}

# Função para reamostrar a matriz X (n x 2p)
# Retorna: Xb impondo H0, por meio de uma amostragem de uma amostra
# combinada preservando a covariância - lembrar de retirar a média

bootSample <- function(X)
{
  n <- nrow(X)
  p <- ncol(X) %/% 2
  sorteio <- sample(n, replace=TRUE)
  Xb <- t(apply(X[sorteio,], 1, trocaXY, p))
  return(Xb)
}

# Função para executar o teste bootstrap
```

```
bootTest <- function(it, X)
{
  Xb <- bootSample(X)
  S <- var(Xb)
  p <- ncol(Xb) %/% 2
  n <- nrow(Xb)
  S11b <- S[1:p,1:p]
  S22b <- S[(p+1):(2*p),(p+1):(2*p)]
  S12b <- S[1:p,(p+1):(2*p)]
  tcb <- tTestb.rhomax(S11b, S22b, S12b, n)
  tcb
  return(tcb)
}

# Função para aplicar o teste bootstrap
# Recebe: X (n x 2p), B
# Retorna: o valor original tc e o valor-p

TestBootMaxRho <- function(X, B=1000)
{
  S <- var(X)
  p <- ncol(X) %/% 2
  n <- nrow(X)
  S11 <- S[1:p,1:p]
  S22 <- S[(p+1):(2*p),(p+1):(2*p)]
  S12 <- S[1:p,(p+1):(2*p)]
  # cálculo do tc original
  tc <- tTestb.rhomax(S11, S22, S12, n)
  it <- matrix(1:B, B, 1)  # número de bootstrap, usa ao invés do loop,
                           # é o contador do bootSample.
  tcb <- apply(it, 1, bootTest, X)

  valor.p <- (length(tcb[abs(tcb) >= abs(tc)])+1)/(B+1)
  return(list(tc=tc, valor.p=valor.p))
}

# Função para simular dados e aplicar o teste bootstrap
# Recebe: a Matriz de Covariâncias pop., n e N
# Dependência: library MASS

library(MASS)
simMCB <- function(Sig, n = 10, N = 1000)
{
  p <- nrow(Sig) / 2
  ct <- matrix(c(0,0),2,1)
  value <- c("0,05", "0,01")
  rownames(ct) <- value
  for (i in 1:N)
  {
    X <- mvrnorm(2, rep(0, 2 * p), Sig)
    tc <- TestBootMaxRho(X, B)
    if (tc$valor.p <= 0.05) ct[1,1] <- ct[1,1] + 1/N
    if (tc$valor.p <= 0.01) ct[2,1] <- ct[2,1] + 1/N
  }
  return(ct)
}
```

```
# Exemplo de Simulação

N <- 10000
n <- 10
##sob H0 # Erro tipo I - matriz de correlação 8x8
Sig <- matrix(c(1,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,1,0.9,0.9,0.9,0.9,0.9,0.9,0.9,
  0.9,1,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,1,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,1,0.9,
  0.9,0.9,0.9,0.9,0.9,0.9,0.9,1,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,1,0.9,0.9,0.9,
  0.9,0.9,0.9,0.9,0.9,1),8,8)
eigen(Sig)
simMCB(Sig, n, N)

##sob H1 # Poder do teste - matriz de covariâncias 8x8
Vmeio <- diag(c(4,5,2,1,16,20,30,10))
Sig1 <- Vmeio %*% Sig %*% Vmeio
simMCB(Sig1, n, N)
```

Tabela 1 - Taxas de erro tipo I de sete testes de igualdade de matrizes de covariâncias: LRT, LRT_1, LRT_2, LRT_3, W_2, W_5 e t_{b_0}(%), considerando diferentes tamanhos amostrais (n), diferentes matrizes de covariâncias $\mathbf{\Sigma}_{(4\times4)}$ e valor nominal de significância de 5%.

$\mathbf{\Sigma}$	n	LRT [a]	LRT_1 [a]	LRT_2 [a]	LRT_3 [a]	W_2 [b]	W_5 [b]	t_{b_0}
$\mathbf{\Sigma}_1$	10	$14,00^*$	$10,80^*$	$7,50^*$	$3,80^+$	$5,30^{ns}$	$7,20^*$	$1,68^+$
	15	$9,20^*$	$7,40^*$	$5,70^*$	$3,80^+$	$4,20^+$	$6,10^*$	$2,55^+$
	20	$8,10^*$	$6,90^*$	$5,70^*$	$4,20^+$	$4,00^+$	$5,90^*$	$2,71^+$
	25	$7,30^*$	$6,30^*$	$5,50^{ns}$	$4,30^+$	$4,00^+$	$5,70^*$	$3,90^+$
	50	$6,10^*$	$5,70^*$	$5,30^{ns}$	$4,80^{ns}$	$4,20^+$	$5,20^{ns}$	$4,50^{ns}$
	75	$5,60^*$	$5,30^{ns}$	$5,10^{ns}$	$4,70^{ns}$	$4,50^{ns}$	$5,10^{ns}$	$4,46^{ns}$
	100	$5,40^{ns}$	$5,30^{ns}$	$5,10^{ns}$	$4,80^{ns}$	$4,70^{ns}$	$5,20^{ns}$	$4,70^{ns}$
$\mathbf{\Sigma}_2$	10	$13,50^*$	$10,20^*$	$7,10^*$	$3,70^+$	$3,70^+$	$7,40^*$	$1,34^+$
	15	$9,70^*$	$7,90^*$	$6,10^*$	$4,00^+$	$3,30^+$	$6,40^*$	$2,58^+$
	20	$8,30^*$	$7,20^*$	$6,00^*$	$4,50^{ns}$	$3,30^+$	$6,00^*$	$2,97^+$
	25	$7,40^*$	$6,50^*$	$5,50^{ns}$	$4,40^+$	$3,30^+$	$5,70^*$	$3,90^+$
	50	$6,20^*$	$5,80^*$	$5,30^{ns}$	$4,80^{ns}$	$4,00^+$	$5,40^{ns}$	$3,92^+$
	75	$5,70^*$	$5,40^{ns}$	$5,20^{ns}$	$4,80^{ns}$	$4,30^+$	$5,30^{ns}$	$4,61^{ns}$
	100	$5,50^{ns}$	$5,30^{ns}$	$5,10^{ns}$	$4,80^{ns}$	$4,50^{ns}$	$5,20^{ns}$	$4,54^{ns}$
$\mathbf{\Sigma}_3$	10	$13,80^*$	$10,50^*$	$7,20^*$	$3,90^+$	$5,00^{ns}$	$7,50^*$	$1,65^+$
	15	$9,50^*$	$7,80^*$	$6,00^*$	$4,00^+$	$4,20^+$	$6,50^*$	$2,76^+$
	20	$8,30^*$	$7,10^*$	$5,80^*$	$4,40^+$	$3,80^+$	$5,90^*$	$3,16^+$
	25	$7,30^*$	$6,40^*$	$5,50^{ns}$	$4,40^+$	$4,00^+$	$5,90^*$	$3,47^+$
	50	$6,10^*$	$5,70^*$	$5,30^{ns}$	$4,80^{ns}$	$4,30^+$	$5,40^{ns}$	$4,02^+$
	75	$5,90^*$	$5,60^*$	$5,30^{ns}$	$5,00^{ns}$	$4,50^{ns}$	$5,30^{ns}$	$4,53^{ns}$
	100	$5,50^{ns}$	$5,40^{ns}$	$5,20^{ns}$	$4,90^{ns}$	$4,60^{ns}$	$5,10^{ns}$	$4,56^{ns}$
$\mathbf{\Sigma}_4$	10	$13,20^*$	$10,30^*$	$7,00^*$	$3,30^+$	$4,00^+$	$7,50^*$	$1,60^+$
	15	$9,60^*$	$7,80^*$	$6,10^*$	$4,10^+$	$3,40^+$	$6,50^*$	$2,60^+$
	20	$8,30^*$	$7,00^*$	$5,90^*$	$4,30^+$	$3,40^+$	$6,00^*$	$3,11^+$
	25	$7,50^*$	$6,60^*$	$5,60^*$	$4,50^{ns}$	$3,50^+$	$5,70^*$	$3,75^+$
	50	$6,10^*$	$5,70^*$	$5,30^{ns}$	$4,80^{ns}$	$4,10^+$	$5,40^{ns}$	$4,16^+$
	75	$5,60^*$	$5,40^{ns}$	$5,10^{ns}$	$4,80^{ns}$	$4,50^{ns}$	$5,40^{ns}$	$4,75^{ns}$
	100	$5,50^{ns}$	$5,30^{ns}$	$5,10^{ns}$	$4,90^{ns}$	$4,60^{ns}$	$5,20^{ns}$	$4,53^{ns}$

[a] Taxas de erro tipo I dos testes LRT, LRT_1, LRT_2 e LRT_3, apresentados por Jiang et al. (1999).

[b] Taxas de erro tipo I dos testes W_2 e W_5, apresentados por Jiang e Sarkar (1998).

* O valor pode ser considerado estatisticamente diferente e maior que o α correspondente, considerando uma confiança de 99%.

$^+$ O valor pode ser considerado estatisticamente diferente e menor que o α correspondente, considerando uma confiança de 99%.

^{ns} O valor não pode ser considerado estatisticamente diferente do α correspondente, considerando uma confiança de 99%.

Tabela 2 - Taxas de erro tipo I do teste de igualdade de matrizes de covariâncias $t_{b_0}(\%)$, considerando diferentes tamanhos amostrais (n), diferentes matrizes de correlações $\boldsymbol{\rho}_{(8\times8)}$ e valor nominal de significância de 5%.

n	ρ_{12}	ρ_{14}	ρ_{16}	ρ_{18}	ρ_{20}	ρ_{22}	ρ_{24}	ρ_{26}	ρ_{28}	ρ_{30}
20	$1,80^{+}$	$1,76^{+}$	$1,69^{+}$	$1,72^{+}$	$1,89^{+}$	$1,67^{+}$	$1,65^{+}$	$1,56^{+}$	$1,61^{+}$	$1,68^{+}$
25	$1,97^{+}$	$2,32^{+}$	$2,22^{+}$	$2,32^{+}$	$2,31^{+}$	$2,25^{+}$	$2,17^{+}$	$2,29^{+}$	$1,92^{+}$	$2,10^{+}$
50	$3,46^{+}$	$2,83^{+}$	$3,34^{+}$	$3,38^{+}$	$3,41^{+}$	$3,30^{+}$	$3,00^{+}$	$3,57^{+}$	$3,11^{+}$	$2,95^{+}$
75	$3,65^{+}$	$3,95^{+}$	$3,87^{+}$	$3,62^{+}$	$3,85^{+}$	$3,85^{+}$	$3,76^{+}$	$3,87^{+}$	$3,91^{+}$	$3,47^{+}$
100	$3,86^{+}$	$3,97^{+}$	$4,15^{+}$	$4,13^{+}$	$4,13^{+}$	$3,91^{+}$	$4,25^{+}$	$4,03^{+}$	$4,37^{+}$	$3,90^{+}$
300	$4,59^{ns}$	$4,56^{ns}$	$4,50^{+}$	$4,67^{ns}$	$4,89^{ns}$	$4,77^{ns}$	$4,63^{ns}$	$4,45^{+}$	$4,30^{+}$	$4,61^{ns}$
500	$4,67^{ns}$	$4,57^{ns}$	$4,88^{ns}$	$4,70^{ns}$	$4,99^{ns}$	$4,88^{ns}$	$4,66^{ns}$	$5,11^{ns}$	$4,83^{ns}$	$5,12^{ns}$
1000	$4,92^{ns}$	$4,79^{ns}$	$5,05^{ns}$	$4,84^{ns}$	$4,65^{ns}$	$4,74^{ns}$	$4,70^{ns}$	$5,10^{ns}$	$4,91^{ns}$	$4,84^{ns}$

$^{+}$ O valor pode ser considerado estatisticamente diferente e menor que o α correspondente, considerando uma confiança de 99%.

^{ns} O valor não pode ser considerado estatisticamente diferente do α correspondente, considerando uma confiança de 99%.

Tabela 3 - Taxas de erro tipo I do teste de igualdade de matrizes de covariâncias $t_{b_0}(\%)$, considerando diferentes tamanhos amostrais (n), diferentes matrizes de correlações $\boldsymbol{\rho}_{(20\times20)}$ e valor nominal de significância de 5%.

n	ρ_{32}	ρ_{34}	ρ_{36}	ρ_{38}	ρ_{40}	ρ_{42}	ρ_{44}	ρ_{46}	ρ_{48}	ρ_{50}	ρ_{52}
20	$0,01^{+}$	$0,01^{+}$	$0,05^{+}$	$0,04^{+}$	$0,00^{+}$	$0,02^{+}$	$0,01^{+}$	$0,02^{+}$	$0,01^{+}$	$0,05^{+}$	$0,03^{+}$
25	$0,18^{+}$	$0,19^{+}$	$0,13^{+}$	$0,31^{+}$	$0,20^{+}$	$0,21^{+}$	$0,15^{+}$	$0,23^{+}$	$0,26^{+}$	$0,23^{+}$	$0,18^{+}$
50	$1,00^{+}$	$0,87^{+}$	$1,39^{+}$	$1,00^{+}$	$1,08^{+}$	$1,13^{+}$	$1,23^{+}$	$1,17^{+}$	$1,13^{+}$	$1,03^{+}$	$1,14^{+}$
75	$1,89^{+}$	$1,94^{+}$	$1,70^{+}$	$1,87^{+}$	$2,11^{+}$	$1,84^{+}$	$1,85^{+}$	$1,90^{+}$	$1,78^{+}$	$2,13^{+}$	$1,96^{+}$
100	$2,28^{+}$	$2,25^{+}$	$2,43^{+}$	$2,42^{+}$	$2,30^{+}$	$2,48^{+}$	$2,40^{+}$	$2,27^{+}$	$2,49^{+}$	$2,44^{+}$	$2,26^{+}$
300	$3,73^{+}$	$3,76^{+}$	$3,59^{+}$	$3,89^{+}$	$3,68^{+}$	$3,74^{+}$	$3,79^{+}$	$3,88^{+}$	$3,87^{+}$	$3,74^{+}$	$3,75^{+}$
500	$4,54^{+}$	$4,19^{+}$	$4,14^{+}$	$4,24^{+}$	$3,95^{+}$	$4,51^{+}$	$4,20^{+}$	$4,17^{+}$	$4,13^{+}$	$4,05^{+}$	$4,20^{+}$
1000	$4,37^{+}$	$4,80^{ns}$	$4,46^{+}$	$4,70^{ns}$	$4,76^{ns}$	$4,53^{+}$	$4,61^{ns}$	$4,49^{+}$	$4,43^{+}$	$4,50^{+}$	$4,66^{ns}$

$^{+}$ O valor pode ser considerado estatisticamente diferente e menor que o α correspondente, considerando uma confiança de 99%.

^{ns} O valor não pode ser considerado estatisticamente diferente do α correspondente, considerando uma confiança de 99%.

Tabela 4 - Poder de sete testes de igualdade de matrizes de covariâncias: LRT, LRT_1, LRT_2, LRT_3, W_2, W_5 e t_{b_0}(%), considerando diferentes tamanhos amostrais (n), diferentes matrizes de covariâncias $\mathbf{\Sigma}_{(4\times4)}$ e $\alpha = 5\%$.

$\mathbf{\Sigma}$	n	LRT [a]	LRT_1 [a]	LRT_2 [a]	LRT_3 [a]	W_2 [b]	W_5 [b]	t_{b_0}
$\mathbf{\Sigma}_5$	10	28,40	22,50	16,70	9,10	18,30	18,50	4,81
	15	32,20	28,30	23,90	18,20	25,40	25,50	11,49
	20	39,3	36,30	32,80	27,00	34,00	33,90	20,75
	25	46,30	43,70	40,90	37,00	42,30	42,00	29,63
	50	76,60	75,60	74,50	73,00	75,50	75,10	65,33
	75	92,10	91,80	91,50	91,00	91,80	91,70	86,68
	100	97,50	97,40	97,30	97,20	97,60	97,60	95,26
$\mathbf{\Sigma}_6$	10	39,20	33,30	26,40	16,40	10,20	27,80	7,55
	15	49,30	45,00	39,80	32,50	15,90	41,20	22,00
	20	60,30	57,10	53,60	48,00	27,50	54,40	36,61
	25	69,60	67,50	64,90	61,20	43,40	65,70	51,96
	50	94,90	94,50	94,10	93,60	90,70	94,30	92,00
	75	99,40	99,40	99,40	99,30	99,10	99,40	98,99
	100	99,90	99,90	99,90	99,90	99,90	99,90	99,89
$\mathbf{\Sigma}_7$	10	48,50	42,10	34,90	24,50	8,90	34,20	8,85
	15	60,00	56,10	51,30	43,80	19,90	50,50	23,71
	20	71,00	68,40	65,20	60,30	39,10	64,80	41,06
	25	80,60	78,80	76,70	73,50	58,40	76,40	56,92
	50	98,30	98,10	98,00	97,70	96,60	97,90	94,32
	75	99,90	99,90	99,90	99,90	99,80	99,90	99,61
	100	100,00	100,00	100,00	100,00	100,00	100,00	99,99
$\mathbf{\Sigma}_8$	10	92,40	89,00	84,70	72,50	88,20	88,10	46,03
	15	98,90	98,40	97,80	96,90	98,30	98,50	84,31
	20	99,90	99,80	99,70	99,70	99,80	99,90	96,76
	25	100,00	100,00	100,00	100,00	100,00	100,00	99,52
	50	100,00	100,00	100,00	100,00	100,00	100,00	100,00
	75	100,00	100,00	100,00	100,00	100,00	100,00	100,00
	100	100,00	100,00	100,00	100,00	100,00	100,00	100,00
$\mathbf{\Sigma}_9$	10	28,10	22,50	16,50	8,60	18,30	18,20	5,16
	15	32,60	28,90	24,40	18,60	25,50	25,50	11,82
	20	39,40	36,30	32,80	28,10	34,20	34,00	20,59
	25	47,00	44,50	41,70	37,40	42,90	42,50	29,54
	50	78,00	77,10	76,00	74,60	76,20	75,70	66,10
	75	92,60	92,40	92,00	91,60	92,40	92,20	86,79
	100	97,90	97,90	97,80	97,70	97,90	97,80	95,21

Tabela 4 -

"continua"

$\boldsymbol{\Sigma}$	n	LRT [a]	LRT_1 [a]	LRT_2 [a]	LRT_3 [a]	W_2 [b]	W_5 [b]	t_{b_0}
$\boldsymbol{\Sigma}_{10}$	10	41,00	35,00	27,40	16,90	10,40	28,20	7,66
	15	50,00	46,00	40,80	33,50	16,20	41,90	21,39
	20	60,70	57,70	54,00	48,70	28,00	54,60	37,21
	25	69,90	67,70	65,10	61,10	43,80	66,20	52,18
	50	94,90	94,60	94,20	93,70	91,00	94,5	91,59
	75	99,50	99,40	99,40	99,30	99,10	99,40	99,17
	100	99,90	99,90	99,90	99,90	99,90	100,00	99,93
$\boldsymbol{\Sigma}_{111}$	10	48,60	42,50	35,20	25,20	9,10	34,20	8,79
	15	60,30	56,30	51,20	43,50	20,00	50,70	23,77
	20	71,20	68,70	65,40	60,50	39,40	65,30	40,99
	25	80,80	79,10	77,00	73,90	58,90	76,80	57,07
	50	98,30	98,10	97,90	97,70	96,70	98,00	94,40
	75	99,90	99,90	99,90	99,90	99,90	99,90	99,63
	100	100,00	100,00	100,00	100,00	100,00	100,00	99,99

[a] Poder dos testes LRT, LRT_1, LRT_2 e LRT_3, apresentados por Jiang et al. (1999).

[b] Poder dos testes W_2 e W_5, apresentados por Jiang e Sarkar (1998).