

TESTE COMPUTACIONALMENTE INTENSIVO BASEADO NA DISTÂNCIA DE MAHALANOBIS PARA NORMALIDADE MULTIVARIADA

Adriele Giaretta BIASE¹
Daniel Furtado FERREIRA¹

- RESUMO: Os testes de normalidade multivariada influenciam diretamente na qualidade e confiabilidade da maioria das pesquisas científicas que realizam análises multivariadas, uma vez que a construção de procedimentos para a realização da inferência considera que a normal multivariada seja a distribuição apropriada para os dados ou resíduos dos modelos adotados. Alguns dos testes de normalidades multivariadas existentes possuem limitações e em grandes amostras os melhores deles não são aplicáveis. Portanto, esse trabalho objetivou propor um novo teste de normalidade multivariada ilimitado quanto ao tamanho da amostra denominado de teste Monte Carlo de normalidade multivariada baseado em distâncias e avaliar o seu desempenho, comparando-o com o do teste de normalidade de Royston. Para avaliação do desempenho foi usada simulação Monte Carlo, mensurando-se as taxas de erro tipo I e o poder. Todos os procedimentos foram implementados no *software* R. O teste Monte Carlo para normalidade multivariada baseado em distância teve grande sucesso no controle das taxas de erro tipo I e poder praticamente equivalente ao do teste de normalidade multivariada de Royston para grandes amostras, além disso, possui a vantagem de ser ilimitado quanto ao tamanho da amostra.
- PALAVRAS-CHAVE: Erro tipo I; normalidade multivariada; poder.

1 Introdução

As análises de dados podem ser realizadas tanto sobre a ótica da estatística univariada quanto da multivariada. Nos casos multivariados, a fundamentação da estatística inferencial é calcada na suposição de normalidade multivariada e/ou na

¹Universidade Federal de Lavras – UFLA, Departamento de Ciências Exatas, Caixa Postal 3037, CEP: 37200-000, Lavras, MG, Brasil. E-mail: adrielegbiase@yahoo.com.br / danielff@dex.ufla.br

extensão do teorema central do limite univariado em sua maioria. Dessa forma, a maior parte das análises de dados é feita por meio de métodos inferenciais da estatística multivariada que são baseados na suposição de normalidade multivariada. Destaca-se então a importância dos testes de normalidade multivariada que influenciam diretamente na qualidade e confiabilidade das pesquisas científicas, uma vez que a não checagem dessa pressuposição pode conduzir a resultados e conclusões incorretas.

Uma forma de verificação da normalidade multivariada é feita mediante o uso de testes baseados nos coeficientes de assimetria e curtose, como os testes de normalidade multivariado de Mardia (1970, 1974) e Doornik e Hansen (2008). A não rejeição da hipótese de distribuição simétrica e mesocúrtica não garante normalidade multivariada, pois uma condição necessária e suficiente para que duas distribuições sejam as mesmas é que todos seus momentos sejam idênticos e não apenas os quatro primeiros momentos. Além disso, esses testes caracterizam-se pelas propriedades assintóticas das distribuições de suas estatísticas, isto é, requerem amostras muito grandes para possuírem tamanho exato e apresentarem grandes valores de poder.

Existem outros testes de normalidade multivariada na literatura especializada, no entanto, apresentam limitações nas aplicações, isto é, são restritos quanto às dimensões das amostras multivariadas. Para esses testes, o número de variáveis não deve ser maior que o tamanho da amostra. Além disso, a amostra não deve ultrapassar 5.000 dados. Isso tem restringido a validade das inferências realizadas, uma vez que os testes de normalidade multivariada e até univariada não podem ser aplicados. Existem exceções, como, por exemplo, o teste de normalidade de Kolmogorov-Smirnov para o caso univariado e o teste de McNeil *et al.* (2005) baseado na distância de Mahalanobis para o caso multivariado.

A avaliação feita por Thode (2002) levou à conclusão que o teste univariado de Kolmogorov-Smirnov deveria ser evitado por possuir pequeno poder e, portanto, não deveria ser considerado para testar a normalidade dos dados ou dos resíduos de um modelo linear ou não-linear. Isso reflete diretamente no teste de McNeil *et al.* (2005) baseado na distância de Mahalanobis para o caso multivariado, uma vez que os autores fundamentaram a estatística do teste de normalidade na estatística de Kolmogorov-Smirnov, obtida a partir das estatísticas de ordem das distâncias de Mahalanobis de cada observação para a média amostral e as distribuições qui-quadrado ou beta.

O número de testes de normalidade multivariada disponíveis atualmente é muito maior do que em algumas décadas anteriores. Embora o desempenho desses recentes testes sejam excepcionais, muitos deles continuam possuindo restrições quanto aos tamanhos amostrais por usar como base a estatística do teste de normalidade univariada de Shapiro-Wilk, estendida para n até 5.000 por Royston (1993). Por exemplo, o teste Alva e Estrada (2009) é limitado à amostra de tamanho máximo de 5.000 e o teste Liang *et al.* (2009) é limitado para amostras de 2.000, por usar os procedimentos de Royston (1983).

Outro fato que pode ser destacado refere-se ao teste Shapiro-Francia de normalidade multivariada implementado no *software* R, do pacote *mvsf* pela função

de mesmo nome do pacote *mvst* implementado por Slawomir Jarek, em 2009, que é uma adaptação do teste Shapiro-Wilk de normalidade multivariada a partir da função *mshapiro.test*. Além desse teste ser limitado para amostras entre 5 e 5.000 não existem detalhes do teste na literatura e nem informações sobre o seu desempenho.

Existe também o falso conceito de que se a amostra for grande, não é necessário aplicar testes de normalidade. Isso não é verdade, embora seja possível se beneficiar dos resultados do teorema do limite central multivariado e justificar a inferência com base nos resultados desse teorema.

O presente trabalho foi realizado com o objetivo de propor um novo teste de normalidade multivariada, ilimitado quanto ao tamanho da amostra n , sendo denotado por teste Monte Carlo de normalidade multivariada baseado em distâncias (TMCNMD). Além disso, buscou-se comparar o desempenho do teste proposto com o teste de normalidade multivariada de Royston (1983, 1993), que é a extensão do teste de Shapiro-Wilk de normalidade univariado para o caso multivariado, o qual foi denotado por TSWNM. Essa avaliação comparativa consistiu na mensuração do desempenho quanto às taxas de erro tipo I e aos valores de poder dos testes usando simulação Monte Carlo.

2 Métodos

2.1 Teste de normalidade multivariada proposto

Foi considerada uma amostra aleatória dada por $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_j, \dots, \mathbf{X}_n$, sendo $\mathbf{X}_j \in \mathbb{R}^p$, de uma população com média $\boldsymbol{\mu}$ com dimensão $(p \times 1)$ e covariância $\boldsymbol{\Sigma}$, de dimensão $(p \times p)$. A hipótese nula de interesse é dada por

H_0 : Os dados amostrais seguem uma distribuição normal multivariada $f_{\mathbf{X}}(\mathbf{x})$,

em que

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

com média $\boldsymbol{\mu} \in \mathbb{R}^p$ e covariância $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$ positiva definida.

Para testar a hipótese nula anterior de normalidade multivariada foi construído um método baseado nas distâncias de cada observação p -variada da amostra de tamanho n para a média amostral $\bar{\mathbf{X}}$.

Assim, as observações multivariadas foram transformadas em distâncias, que são unidimensionais. A distância utilizada foi a de Mahalanobis. Calculadas cada uma das n distâncias entre os vetores de observações e a média amostral, os seus valores foram inicialmente transformados e, posteriormente, ordenados. Valores médios das estatísticas de ordem das distâncias transformadas foram estimados usando a sua distribuição de amostragem, que é a beta segundo Gnanadesikan e Kettenring (1972). Finalmente, a estatística do teste foi computada utilizando o coeficiente de determinação entre as estatísticas de ordem observadas relativas às

distâncias transformadas e às estatísticas de ordem esperadas. A distribuição nula da estatística foi obtida por simulação Monte Carlo e o valor observado na amostra original foi confrontado com os valores da distribuição nula Monte Carlo, para obtenção do valor- p e tomada de decisão. Todos os procedimentos descritos foram implementados no R e os detalhes podem ser vistos no algoritmo apresentado na sequência.

A ideia principal do teste baseia-se no fato de que se a distribuição do conjunto de dados que se apresenta para o teste for de fato a normal multivariada, a distribuição das distâncias de Mahalanobis transformadas entre as observações multivariadas e a média amostral é a beta. Assim, o coeficiente de determinação entre o valor esperado e o valor observado das estatísticas de ordem das distâncias transformadas tenderá a 1. Se os dados submetidos ao teste não forem normais multivariados, então o coeficiente de determinação tenderá a zero. A distribuição dessa estatística, o coeficiente de determinação entre as estatísticas de ordem esperadas e observadas, não é conhecida. Assim, optou-se por construir um teste cuja distribuição da estatística fosse obtida por simulação Monte Carlo. Assim, o teste proposto é essencialmente um algoritmo, conforme descrição apresentada a seguir.

Algoritmo para o teste Monte Carlo de normalidade multivariada baseado em distâncias

Os passos para obtenção do teste Monte Carlo de normalidade multivariada baseado em distâncias são:

1. considerar $\mathbf{X}_{n \times p}$ a matriz de dados e definir N_{SM} como sendo o número de simulações Monte Carlo para realizar o teste;
2. calcular as distância de Mahalanobis

$$D_j^2 = (\mathbf{X}_j - \bar{\mathbf{X}})^\top \mathbf{S}^{-1} (\mathbf{X}_j - \bar{\mathbf{X}}), \text{ para } j = 1, 2, \dots, n,$$

em que \mathbf{X}_j é o vetor relativo a j -ésima observação p -variada posicionada na sua forma transposta na j -ésima linha da matriz de dados, $\bar{\mathbf{X}}$ é o vetor de médias das n observações e \mathbf{S} a matriz de variância e covariância amostral;

3. ordenar as distâncias da amostra original obtendo-se as estatísticas de ordem:

$$D_{(1)}^2, D_{(2)}^2, \dots, D_{(j)}^2, \dots, D_{(n)}^2;$$

4. realizar a seguinte transformação

$$b_{(j)} = \frac{nD_{(j)}^2}{(n-1)^2}, \quad (2)$$

que se distribui de forma exata, sob normalidade multivariada, como uma variável beta com parâmetros

$$\alpha = \frac{p}{2} \quad e \quad \beta = \frac{n-p-1}{2}, \quad (3)$$

conforme demonstrado por Gnanadesikan e Kettenring (1972);

5. estimar os valores esperados das estatística de ordem da distribuição beta por

$$b_j^* = F^{-1} \left[\frac{(j-a)}{(n-a-b+1)}; \alpha, \beta \right], \quad (4)$$

relativos a estatística de ordem $b_{(j)}$, que é j -ésimo menor valor da amostra aleatória, para $j = 1, 2, \dots, n$, sendo: F^{-1} a inversa da função de distribuição beta,

$$a = \frac{p-2}{2p} \quad e \quad b = \frac{n-p-3}{2(n-p-1)}; \quad (5)$$

6. calcular o coeficiente de determinação amostral, considerando a amostra original, r_c^2 , entre $b_{(j)}$ e b_j^* , em que r_c é dado pela expressão:

$$r_c^2 = \frac{\left(\sum_{j=1}^n b_{(j)} b_j^* - \frac{\sum_{j=1}^n b_{(j)} \sum_{j=1}^n b_j^*}{n} \right)^2}{\left[\sum_{j=1}^n b_{(j)}^2 - \frac{\left(\sum_{j=1}^n b_{(j)} \right)^2}{n} \right] \left[\sum_{j=1}^n (b_j^*)^2 - \frac{\left(\sum_{j=1}^n b_j^* \right)^2}{n} \right]}; \quad (6)$$

7. gerar uma amostra aleatória $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n$, de tamanho n da distribuição normal multivariada, com média $\boldsymbol{\mu} = \mathbf{0}$ e covariância $\boldsymbol{\Sigma} = \mathbf{I}$. A escolha desses parâmetros não interfere nos resultados do teste, uma vez que as distâncias são invariantes às transformações de escalas;
8. calcular as distâncias para a média amostral e ordenar os valores, obtendo-se as estatísticas de ordem:

$$b_{(1)}, b_{(2)}, \dots, b_{(n)},$$

considerando que $\bar{\mathbf{Z}}$ e \mathbf{S}_Z são o vetor de médias e a matriz de covariâncias da amostra gerada no passo 7, então as distâncias são calculadas por

$$D_j^2 = (\mathbf{Z}_j - \bar{\mathbf{Z}})^\top \mathbf{S}_{\bar{\mathbf{Z}}}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}), \text{ para } j = 1, 2, \dots, n,$$

e as estatísticas de ordem da beta, pela transformação (2), após as distâncias terem sido ordenadas;

9. calcular o coeficiente de determinação amostral r_i^2 entre $b_{(j)}$ e b_j^* , no i -ésimo passo da iteração Monte Carlo;
10. armazenar r_i^2 , juntamente com o valor original e os demais valores obtidos na iteração anterior, se houver ocorrido alguma iteração anterior;
11. se o número de valores armazenados, incluindo o valor original, for menor que $N_{SM} + 1$, então vá para o passo 7, senão vá para o passo 12;
12. calcular o valor- p por:

$$\text{valor} - p = \frac{\sum_{i=1}^{N_{SM}+1} I(r_i^2 \leq r_c^2)}{N_{SM} + 1}, \quad (7)$$

em que:

- i*) $I(\bullet)$ é uma função indicadora;
- ii*) r_i^2 é o coeficiente de determinação da i -ésima unidade amostral da simulação Monte Carlo;
- iii*) r_c^2 é o coeficiente de determinação calculado a partir da amostra original;

A ideia é estimar a proporção de estimativas do coeficiente de determinação, obtidas na distribuição nula, que são iguais ou inferiores a estimativa obtida na amostra original. Se essa proporção for menor que o valor nominal de significância α , a decisão deve ser de rejeição da hipótese nula, uma vez que o valor original é atípico na distribuição nula obtida sob H_0 ;

13. retornar o valor- p e o valor da estatística na amostra original r_c^2 .

Usou-se a distância de cada observação para a média, devido ao fato de não ser possível ordenar elementos de uma matriz $n \times p$, pois existem p colunas a serem ordenados e a ordenação de uma delas resultaria em pelo menos uma das demais fora de ordem.

Assim, as estatísticas de ordem observadas em uma amostra aleatória, se a hipótese de normalidade multivariada for verdadeira, devem se correlacionar fortemente com as estatísticas de ordem esperadas da distribuição beta. Os valores esperados foram obtidos de forma aproximada pela inversa da função de distribuição beta. A inversa depende das quantidades a e b apresentadas anteriormente, que servem para realizar correções de continuidade, sendo que para as suas definições foram considerados os aspectos de maior acurácia dos resultados dos valores esperados das estatísticas de ordem da distribuição beta.

2.2 Teste Shapiro-Wilk de normalidade multivariada

A descrição detalhada da extensão do teste Shapiro-Wilk de normalidade univariada para o caso multivariado proposto por Royston, pode ser encontrada em Ferreira (2008) e em Royston (1983, 1993). Foi usada a mesma implementação da rotina realizada por Oliveira e Ferreira (2009) para o teste de normalidade multivariada proposto por Royston (1983, 1993).

2.3 Validação do desempenho

Foi usada simulação Monte Carlo para avaliar o desempenho do novo teste de normalidade multivariada mensurando-se as taxas de erro tipo I e o poder. O teste de normalidade multivariada de Royston (1983, 1993) foi considerado para fins de comparação. Para execução desse procedimento duas etapas foram consideradas, uma referente ao erro tipo I e outra, relacionada ao poder do teste.

2.3.1 Avaliação do erro tipo I

Para avaliar a taxa de erro tipo I , amostras normais p -dimensionais foram geradas com o vetor de médias $\boldsymbol{\mu} = \mathbf{0}$ e matriz de covariância $\boldsymbol{\Sigma}$ com estrutura de correlação dada por

$$\boldsymbol{\Sigma} = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix} = \sigma^2 \boldsymbol{\rho}. \quad (8)$$

A matriz definida em (8) é a matriz de equicorrelação multiplicada por uma constante positiva real, σ^2 . Para a realização das simulações o parâmetro σ^2 foi fixado em 1 e a correlação, ρ , em 0,5 sem perda de generalidade. Isso decorre do fato de que as distâncias são invariantes a transformações de escalas (Thode, 2002).

As configurações, quanto aos testes de normalidade multivariados avaliados, juntamente com as combinações de n e p para a dimensão da amostra dada pela matriz $\mathbf{X}_{n \times p}$, foram especificadas na Tabela 1. Foi necessário assumir a restrição, dada por $n > p$, na geração das amostras normais multivariadas, evitando-se que estimativas singulares da matriz de covariâncias fossem obtidas.

O processo de simulação Monte Carlo foi repetido 2.000 vezes e o TMCNMD e o TSWNM aplicados em cada amostra considerando o nível de significância nominal, α , fixado em 1% e 5%. As taxas de erro tipo I foram calculadas para cada teste como a proporção de vezes que a hipótese nula foi rejeitada. Os resultados foram comparados com o nível de significância nominal.

Uma vez que estas taxas de erro tipo I foram estimadas utilizando simulações de Monte Carlo, elas não estão livres de erro. Assim, testes binomiais exatos,

Tabela 1 - Testes de normalidade multivariados aplicados nas avaliações de desempenho, com os tamanhos amostrais n e números de variáveis p usados nas simulações Monte Carlo

Testes de normalidade multivariada	n	p
TMCNMD e TSWNM	10	2, 4, 6, 8
TMCNMD e TSWNM	20	2, 5, 10, 18
TMCNMD e TSWNM	50	2, 10, 25, 48
TMCNMD e TSWNM	100	2, 10, 50, 98
TMCNMD e TSWNM	200	2, 10, 50, 100
TMCNMD	10000	5

considerando um nível de significância nominal de 1%, para as hipóteses

$$\begin{cases} H_0 : \alpha = 1\% \text{ versus } H_1 : \alpha \neq 1\%, \\ H_0 : \alpha = 5\% \text{ versus } H_1 : \alpha \neq 5\%, \end{cases} \quad (9)$$

foram aplicados.

Se a hipótese nula for rejeitada e as taxas de erro tipo I observadas forem considerados significativamente (valor- $p < 0,01$) inferiores ao nível de significância nominal considerado, o teste será *conservativo*. Se a hipótese nula for rejeitada e as taxas de erro tipo I observadas forem consideradas significativamente (valor- $p < 0,01$) superiores ao nível nominal de significância, o teste será *liberal*. Porém, se as taxas de erro tipo I observadas não forem consideradas significativamente (valor- $p > 0,01$) diferentes do nível nominal de significância considerado, o teste é *exato*.

Se x representa o número de hipóteses nula rejeitadas em $N = 2.000$ simulações Monte Carlo, para o nível de significância nominal α , Oliveira e Ferreira (2009) consideraram a estatística do teste, usando a relação entre a F e a distribuição binomial, com probabilidade de sucesso $p = \alpha$, por:

$$F_c = \left(\frac{x+1}{N-x} \right) \left(\frac{1-\alpha}{\alpha} \right), \quad (10)$$

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

2.3.2 Avaliação do poder do teste

Foram simuladas amostras aleatórias de distribuições não-normais multivariadas para avaliar o poder dos testes em rejeitar a hipótese nula que por construção é falsa. Foi utilizado o mesmo procedimento para realizar a simulação descrita na seção (2.1). A diferença consiste na geração de realizações de variáveis aleatórias de distribuições não-normais como a *t-student* multivariada com diferentes graus de liberdade (com $\nu = 1$ e $\nu = 30$), a log-normal multivariada, a mistura de normais

multivariadas e a uniforme multivariada. Foram utilizadas as funções geradoras de variáveis aleatórias multidimensionais do programa R para a realização de tal tarefa.

Para gerar os vetores aleatórios da mistura de normais multivariadas, a seguinte função de densidade foi considerada

$$f_{\mathbf{X}}(\mathbf{x}) = \delta(2\pi)^{-p/2}|\boldsymbol{\Sigma}_1|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}_1^{-1}(\mathbf{x} - \boldsymbol{\mu}_1)\right\} \\ + (1 - \delta)(2\pi)^{-p/2}|\boldsymbol{\Sigma}_2|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}_2^{-1}(\mathbf{x} - \boldsymbol{\mu}_2)\right\}, \quad (11)$$

em que $0 \leq \delta \leq 1$, $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ são as matrizes de covariância positiva definida e $\boldsymbol{\mu}_1$ e $\boldsymbol{\mu}_2$ são os vetores médios p -dimensionais. Para realizar as simulações os seguintes valores foram considerados: $\delta = 0,30$, $\boldsymbol{\mu}_1 = \mathbf{0}$, $\boldsymbol{\mu}_2 = 10 \mathbf{J}$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\rho}$, $\boldsymbol{\Sigma}_2 = 10 \boldsymbol{\rho}$, em que \mathbf{J} é um vetor ($p \times 1$) de 1's e a matrix $\boldsymbol{\rho}$ foi definida na Equação (8) com $\rho = 0,5$.

Realizações de vetores aleatórios da uniforme multivariada \mathbf{y} foram gerados a partir de vetores aleatórios com distribuição normal multivariada, ou seja, $\mathbf{X} \sim N_p(\mathbf{0}, \boldsymbol{\rho})$, em que $\boldsymbol{\rho}$ é a matriz de correlação obtida da equação (8) considerando $\sigma^2 = 1$. O vetor aleatório \mathbf{Y} , com distribuição uniforme multivariada é obtido calculando-se o seu i -ésimo componente pela seguinte transformação $\mathbf{Y}_i = \Phi(X_i)$, $i = 1, 2, \dots, p$, em que $\Phi(z)$ é a função de distribuição acumulada da normal padrão avaliada em z . Assim, \mathbf{Y} possui distribuição uniforme $U(0, 1)^p$, ou seja, possui distribuição uniforme $(0, 1)$ p -dimensional, conforme descrito em Oliveira e Ferreira (2009).

Geradas as amostras de cada uma dessas distribuições, em cada configuração, foram aplicados os testes de normalidade multivariada descritos anteriormente em 2.000 repetições. Foram computadas as proporções de vezes que a hipótese nula foi rejeitada para os dois testes (TMCNMD e o TSWNM), considerando os níveis nominais fixos para $\alpha = 1\%$ e $\alpha = 5\%$. Essas proporções estimam o poder dos dois testes, que foram comparados entre si.

2.4 Exemplo

Um subconjunto do conjunto de dados apresentado por Royston (1983) foi utilizado com o objetivo de ilustrar a aplicação de ambos os testes de normalidade multivariado. Nesse exemplo as variáveis X_1 e X_2 referem-se à concentração de hemoglobina e contagem de linfócitos, respectivamente, mensuradas em $n = 20$ pintores de carro. Os dados observados estão apresentados a seguir:

Pintores	X_1	X_2	Pintores	X_1	X_2
1	13,4	14	11	15,2	26
2	14,6	15	12	16,9	28
3	13,5	19	13	14,8	24
4	15,0	23	14	16,2	26
5	14,6	17	15	14,7	23
6	14,0	20	16	14,7	9
7	16,4	21	17	16,5	18
8	14,8	16	18	15,4	28
9	15,2	27	19	15,1	17
10	15,5	34	20	14,2	14

3 Resultados e discussão

Dividiu-se esta seção em duas etapas: uma referente aos resultados para o erro tipo I e outra, para os resultados do poder dos testes.

3.1 Erro tipo I

Na Tabela 2 estão apresentadas as taxas de erro tipo I do TMCNMD e do TSWNM considerando os níveis de significância nominais de 1% e 5%, de acordo com as combinações de n e p apresentados na Tabela 1. Os resultados foram confrontados com cada nível de significância nominal α associado, por intermédio do teste binomial exato com nível de significância de 1%. O TMCNMD apresentou controle da taxa de erro tipo I em todas as situações, independente do tamanho da amostra e do número de variáveis, como também para ambos os níveis nominais de significância. Sendo assim as taxas de erro tipo I observadas não foram consideradas significativamente (valor- $p > 0,01$) diferentes do nível nominal de significância e este foi então considerado exato em todas as situações. Isso já era esperado por construção (definição) do novo teste, o que corrobora a adequação do procedimento proposto.

O TSWNM teve um bom desempenho em relação ao controle das taxas de erro tipo I , mostrando-se, na maioria das vezes, exato. Para valores de p grandes com $n = 20$, o TSWNM mostrou-se conservativo para $\alpha = 5\%$. O mesmo ocorreu com $n = 100$ e $p = 50$. Para situação em que o número de variáveis p foi próximo de n , com $n = 100$, o teste foi liberal para $\alpha = 1\%$.

Buscou comparar algumas das taxas de erro tipo I do TMCNMD e do TSWNM com outros testes já avaliados na literatura. Para a amostra $n = 100$ com $p = 2$, o teste de normalidade de curtose multivariado avaliado por Cirillo e Ferreira (2003) foi significativamente (valor- $p < 0,01$) inferior ao nível nominal de 5% considerado, ou seja, conservativo. Para o mesmo tamanho amostral, os testes apresentados no presente trabalho, o teste de normalidade baseado no coeficiente de correlação Quantil-Quantil multivariado e o teste de normalidade multivariada baseado no coeficiente de assimetria avaliados em Cirillo e Ferreira (2003) foram exatos.

Tabela 2 - Erro tipo I dos testes TMCNMD, TNMBP e TSWNM, obtido de 2.000 amostras simuladas sob a hipótese de nulidade de distribuição normal multivariada, em função de n e p para $\alpha = 1\%$ e 5%

n	p	1%		5%	
		TMCNMD	TSWNM	TMCNMD	TSWNM
10	2	0,0070	0,0125	0,0450	0,0535
	4	0,0085	0,0110	0,0525	0,0495
	6	0,0100	0,0050	0,0515	0,0440
	8	0,0110	0,0055	0,0510	0,0450
20	2	0,0115	0,0110	0,0555	0,0460
	5	0,0085	0,0080	0,0510	0,0405
	10	0,0085	0,0070	0,0550	0,0355 ⁻
	18	0,0085	0,0080	0,0505	0,0340 ⁻
50	2	0,0105	0,0110	0,0550	0,0465
	10	0,0080	0,0095	0,0505	0,0405
	25	0,0085	0,0115	0,0455	0,0475
	48	0,0090	0,0165	0,0430	0,0420
100	2	0,0080	0,0100	0,0435	0,0460
	10	0,0095	0,0090	0,0525	0,0470
	50	0,0135	0,0095	0,0530	0,0350 ⁻
	98	0,0095	0,0175 ⁺	0,0475	0,0435
200	2	0,0090	0,0075	0,0455	0,0440
	10	0,0115	0,0110	0,0495	0,0455
	50	0,0110	0,0085	0,0455	0,0410
	100	0,0135	0,0135	0,0455	0,0385
10.000	5	0,0082	-	0,0418	-

⁺ Taxas de erro tipo I superiores ao valor nominal de 1% de significância ($P < 0,01$).

⁻ Taxas de erro tipo I inferiores ao valor nominal de 1% de significância ($P < 0,01$).

Os testes avaliados por Farrell *et al.* (2007) para 10.000 procedimento Monte Carlo, como o teste de Doornik e Hansen (1994) e o teste de Henze e Zirkler (1990), para amostras de tamanho $n = 50$ com $p = 2$, $n = 50$ com $p = 10$ e $n = 100$ com $p = 2$ foram exatos, da mesma forma que o TMCNMD e o TSWNM (Tabela 2), para $\alpha = 5\%$. Todos esses testes foram também exatos para amostra de tamanho $n = 200$ com $p = 10$.

3.2 Poder dos testes

Os valores de poder do TMCNMD e do TSWNM, considerando a distribuição t -Student multivariada com $\nu = 1$ para $\alpha = 5\%$ e $\alpha = 1\%$ são mostrados na Tabela 3 em função de n e p . O TMCNMD não obteve um poder satisfatório para tamanhos de amostras n pequenos, como 10 ou 20. Notou-se, considerando o TMCNMD, para tamanhos de amostrais n iguais a 10, 20, 50 e 100, que o aumento do número de variáveis causou diminuição do poder, chegando a ficar abaixo do

nível nominal de significância, em alguns casos. Isso tendia a ocorrer quando o valor de p , aproximava ou ultrapassava do limite de $n/2$. O TMCNMD apresentou bom poder para tamanhos grandes de amostras, $n = 200$ independente do número de variáveis, embora o valor máximo de p considerado foi de 100. Não se conseguiu uma explicação plausível para esse caso. Acredita-se que haja um efeito do teorema central do limite, na medida que p aumente, diminuindo assim o poder do teste.

A eficiência do teste TMCNMD para tamanho amostral $n = 10.000$ foi o melhor possível, atingindo 100% de poder, embora tenha sido avaliado somente para $p = 5$. O TSWNM apresentou poder não muito elevado para amostra $n = 10$, independentemente do número de variáveis, embora os valores fossem bastante satisfatórios. Para amostra de tamanho $n = 20$, o poder cresceu à medida que o número de variáveis p aumentava, que é um efeito inverso ao do teste proposto. Para tamanhos amostrais superiores ou iguais a 50, o poder atingiu 100%, independente do número de variáveis considerado.

É conveniente salientar que a *t-Student* multivariada com $\nu = 1$ grau de liberdade é uma distribuição bem diferenciada da normal multivariada, no sentido de ter caudas mais pesadas. Dessa forma, o TSWNM teve valores elevados de poder, e mesmo o TMCNMD, que apresentou desempenho pior em algumas situações, é ainda um teste adequado para maior parte dos casos.

As combinações de n e p foram escolhidos para que os resultados, entre outros motivos, pudessem ser comparados com os resultados de Oliveira e Ferreira (2009). Considerando, ainda, a *t-Student* multivariada com $\nu = 1$ grau de liberdade, $\alpha = 5\%$ e 2.000 simulações Monte Carlo, notou-se que o poder do TMCNMD foi semelhante ao do teste Quiquadrado de Person da proposta de Oliveira e Ferreira (2009), isto é, o poder decrescia à medida que o número de variáveis ficava próximo do tamanho da amostra. Com $n = 10$, o TMCNMD foi melhor na maior parte das vezes que o teste quiquadrado de Pearson para normalidade multivariada. Em relação ao teste conjunto de assimetria e curtose avaliado por Oliveira e Ferreira (2009), o TMCNMD e o TSWNM apresentaram maiores valores de poder principalmente em amostras com número de variáveis bem próximos do tamanho amostral.

Na Tabela 4 os valores de poder do TMCNMD e TSWNM são apresentados, considerando a distribuição *t-Student* multivariada com $\nu = 30$, para $\alpha = 1\%$ e 5% em função de n e p . Vale ressaltar que a distribuição *t-Student* multivariada com $\nu = 30$ se aproxima consideravelmente da distribuição normal multivariada. Em razão disso, verifica-se o valor do poder é quase equiparável ao nível nominal de significância para amostras pequenas nos dois testes de normalidade multivariada. Lembrando novamente que quando os graus de liberdade da distribuição *t-Student* é muito grande, ou seja, $\nu \rightarrow \infty$, essa distribuição tende à normal multivariada, motivo pelos quais os valores de poder ficam bem próximos do nível nominal de significância. À medida que o tamanho da amostra cresce o poder aumenta gradativamente.

Para $\alpha = 1\%$, o TMCNMD apresentou os mais baixos valores de poder em relação ao TSWNM quando o tamanho da amostra foi de $n = 20$ e $n = 50$ para valores de $p > 2$. O TMCNMD foi melhor do que o TSWNM nas amostras de tamanho $n = 20$, $n = 50$ e $n = 100$ para $p = 2$. De forma geral, o TMCNMD

Tabela 3 - Poder dos testes TMCNMD e TSWNM considerando a distribuição *t-Student* multivariada com $\nu = 1$ grau de liberdade em função de n e p para $\alpha = 1\%$ e 5%

n	p	1%		5%	
		TMCNMD	TSWNM	TSWNM	TSWNM
10	2	0,1600	0,6070	0,3930	0,7185
	4	0,0245	0,6990	0,1450	0,7945
	6	0,0055	0,7495	0,0605	0,8425
	8	0,0250	0,8190	0,0970	0,8855
20	2	0,7430	0,8955	0,8635	0,9375
	5	0,5035	0,9715	0,8010	0,9850
	10	0,0625	0,9895	0,2705	0,9940
	18	0,1070	0,9950	0,2550	0,9980
50	2	0,9875	1,0000	0,9970	1,0000
	10	0,9985	1,0000	1,0000	1,0000
	25	0,7180	1,0000	0,9275	1,0000
	48	0,2920	1,0000	0,5250	1,0000
100	2	1,0000	1,0000	1,0000	1,0000
	10	1,0000	1,0000	1,0000	1,0000
	50	0,9980	1,0000	1,0000	1,0000
	98	0,4890	1,0000	0,7590	1,0000
200	2	1,0000	1,0000	1,0000	1,0000
	10	1,0000	1,0000	1,0000	1,0000
	50	1,0000	1,0000	1,0000	1,0000
	100	1,0000	1,0000	1,0000	1,0000
10.000	5	1,0000	-	1,0000	-

apresentou maior poder do que o TSWNM em amostras de tamanho 10 e 200. O padrão do comportamento TMCNMD e TSWNM para $\alpha = 5\%$ foi similar ao observado para $\alpha = 1\%$. Não houve um teste que superasse os demais em todas as configurações de n e p consideradas. Para n pequeno, os valores de poder se assemelham ao nível nominal de significância. O TMCNMD em grandes amostras ($n = 200$) apresentou maior poder quando p era intermediário (10 e 50) e seu poder reduziu na medida que p se aproximava de n , ao contrário do TSWNM. Mesmo para $n = 200$, os valores de poder foram muito pequenos. Para $n = 10.000$ e $p = 5$, o novo teste apresentou poder de 100%, mesmo com uma distribuição com grande semelhança em relação à normal multivariada.

O poder do novo teste foi comparado com os dos testes avaliados por Oliveira e Ferreira (2009) para normalidade multivariada em relação a *t-Student* multivariada com $\nu = 30$ e $\alpha = 5\%$. Notou-se que o TMCNMD apresentou maior poder que o do teste conjunto de assimetria e curtose multivariado e do teste quiquadrado de Pearson multivariado para circunstâncias de tamanhos amostrais relativamente pequenos, ou seja, 10 e 20, sem exceções quanto aos número de variáveis consideradas. O TMCNMD apresentou também maior poder quando

Tabela 4 - Poder dos testes TMCNMD e TSWNM considerando a distribuição *t-Student* multivariada com $\nu = 30$ graus de liberdade em função de n e p para $\alpha = 1\%$ e 5%

n	p	1%		5%	
		TMCNMD	TSWNM	TMCNMD	TSWNM
10	2	0,0135	0,0110	0,0500	0,0515
	4	0,0125	0,0105	0,0555	0,0465
	6	0,0080	0,0120	0,0470	0,0575
	8	0,0105	0,0070	0,0530	0,0455
20	2	0,0160	0,0145	0,0700	0,0605
	5	0,0110	0,0125	0,0555	0,0630
	10	0,0070	0,0220	0,0405	0,0795
	18	0,0190	0,0265	0,0720	0,0805
50	2	0,0245	0,0210	0,0840	0,0780
	10	0,0270	0,0420	0,0920	0,1020
	25	0,0045	0,0680	0,0380	0,1490
	48	0,0215	0,0730	0,0780	0,1585
100	2	0,0395	0,0320	0,1160	0,0995
	10	0,0685	0,0640	0,1805	0,1465
	50	0,0080	0,1115	0,0460	0,2275
	98	0,0215	0,1450	0,0770	0,2535
200	2	0,0470	0,0385	0,1450	0,1130
	10	0,1855	0,1005	0,3730	0,2070
	50	0,2165	0,2160	0,4200	0,3480
	100	0,0110	0,2475	0,0690	0,3960
10.000	5	1,0000	-	1,0000	-

os tamanhos amostrais foram $n = 50$, $n = 100$ e $n = 200$ em relação ao teste quiquadrado de Pearson multivariado proposto por Oliveira e Ferreira (2009), exceto para $n = 50$ com $p = 25$ e $n = 100$, 200 quando $p = 50$ e $n = 200$ com $p = 100$. Para amostras pequenas, o TMCNMD desempenhou melhor do que o teste multivariado conjunto de assimetria e curtose.

Os valores de poder para TMCNMD e TSWNM, considerando a distribuição log-normal multivariada para $\alpha = 1\%$ e $\alpha = 5\%$ são mostrados na Tabela 5 com os relativos valores de n e p adotados. O TMCNMD não obteve desempenho satisfatório para tamanhos de amostras n com valores iguais a 10 e 20. Para amostras de tamanho 50 e número de variáveis com valores iguais a 2 e 10, o desempenho foi considerado bom. Já considerando esse mesmo tamanho de amostra com o número de variáveis igual a 25 e 48, o poder do TMCNMD reduz. Essa tendência de redução com o aumento do número de variáveis ocorreu também para valores de n maiores que 50, como já relatado anteriormente em outras distribuições. Para grandes amostras, com p pequeno ou intermediário em relação ao n , os valores de poder foram considerados altos.

O comportamento em relação ao poder do TSWNM para ambos os níveis

Tabela 5 - Poder dos testes TMCNMD e TSWNM considerando a distribuição log-normal multivariada em função de n e p para $\alpha = 1\%$ e 5%

n	p	1%		5%	
		TMCNMD	TSWNM	TMCNMD	TSWNM
10	2	0,1495	0,6335	0,4010	0,8050
	4	0,0265	0,8455	0,1480	0,9400
	6	0,0050	0,9490	0,0740	0,9815
	8	0,0095	0,9790	0,0610	0,9945
20	2	0,6005	0,9715	0,7705	0,9900
	5	0,4115	1,0000	0,7650	1,0000
	10	0,0480	1,0000	0,2315	1,0000
	18	0,1090	1,0000	0,2845	1,0000
50	2	0,9260	1,0000	0,9810	1,0000
	10	0,9890	1,0000	1,0000	1,0000
	25	0,3360	1,0000	0,6915	1,0000
	48	0,3850	1,0000	0,5995	1,0000
100	2	0,9990	1,0000	1,0000	1,0000
	10	1,0000	1,0000	1,0000	1,0000
	50	0,8420	1,0000	0,9800	1,0000
	98	0,6775	1,0000	0,8760	1,0000
200	2	1,0000	1,0000	1,0000	1,0000
	10	1,0000	1,0000	1,0000	1,0000
	50	1,0000	1,0000	1,0000	1,0000
	100	1,0000	1,0000	1,0000	1,0000
10.000	5	1,0000	-	1,0000	-

nominais de significância é semelhante. O que se observa é uma redução do poder, como esperado pela redução de α . Para ambos os níveis nominais de significância e tamanhos de amostras maiores que 10, o TSWNM atingiu poder máximo de 100% independente do número de variáveis consideradas, exceto para $n = 20$ com $p = 2$.

Os resultados do poder para distribuição log-normal com $\alpha = 5\%$ foram comparados com os dos testes avaliados por Oliveira e Ferreira (2009), a única diferença consiste nos valores dos parâmetros adotados por esses autores que fizeram com que a distribuição apresentasse características bem diferenciadas da normal multivariada em relação aos dos parâmetros adotados nesse trabalho. No entanto, para amostras pequenas, 10, o teste de normalidade multivariado de Royston (1983, 1993), avaliado por Oliveira e Ferreira (2009) tiveram mais rejeição da hipótese H_0 do que os valores do TSWNM observados na tabela 5.

Notou-se, para amostras pequenas, que o TMCNMD foi melhor que o teste quiquadrado de Pearson multivariado, como, por exemplo, para amostras $n = 10$ com $p = 4$ e $n = 10$ com $p = 6$. Para $n = 10$ com $p = 6$ e $n = 10$ com $p = 8$, o TMCNMD foi melhor que o teste conjunto de assimetria e curtose multivariado. O TMCNMD foi inferior aos testes avaliados por Oliveira e Ferreira (2009) para tamanhos amostrais com $n = 20$ relativos a todos os tamanhos de p , exceto quando

$n = 20$ com $p = 18$, para o teste conjunto de assimetria e curtose multivariado. O TMCNMD foi superior em poder em relação ao teste conjunto de assimetria e curtose multivariado para $n = 50$ com $p = 48$ e $n = 100$ com $p = 98$. Para amostra $n = 200$ tanto os testes avaliados por Oliveira e Ferreira (2009) quanto os testes avaliados nesse trabalho tiveram poder máximos ou próximos do máximo.

Os valores de poder para os TMCNMD e TSWNM considerando a distribuição dada pela mistura de normais multivariadas para $\alpha = 1\%$ e 5% são mostrados na Tabela 6 em função de n e p . O padrão relativo de resposta para $\alpha = 5\%$ quanto ao poder foi similar ao obtido para $\alpha = 1\%$. O TMCNMD não obteve um desempenho satisfatório para tamanhos de amostras n com valores iguais a 10 e 20. Para $n = 50$, o teste mostrou-se melhor quando o número de variáveis foi igual a $p = 25$, para os demais valores de variáveis o teste mostrou com baixo poder. Para $n = 200$, o TMCNMD apresentou menor poder que o TSWNM para o número de variáveis iguais a 2 e 10. Já para o número de variáveis iguais a 50 e 100, o teste atingiu poder máximo. O TSWNM mostrou-se melhor do que novo teste proposto principalmente nas situações em que as amostras eram pequenas. No caso em que foi contemplado o tamanho amostral 10.000, o TNMDCM apresentou poder 100% tanto para $\alpha = 5\%$ quanto para $\alpha = 1\%$.

Considerando $\alpha = 5\%$ e os mesmos valores paramétricos da mistura de normais multivariadas adotados nesse trabalho e por Oliveira e Ferreira (2009), verifica-se que o TMCNMD apresentou na maioria das vezes melhor desempenho ou desempenho equivalente quanto ao poder para amostras entre $n = 10$ e $n = 200$ (Tabela 6), em relação ao teste conjunto de assimetria e curtose multivariado e ao teste quiquadrado de Pearson multivariado da proposta de Oliveira e Ferreira (2009).

Na Tabela 7 são apresentados os valores de poder do TMCNMD e do TSWNM para a distribuição uniforme multivariada em função de n e p , para os níveis de significância $\alpha = 5\%$ e $\alpha = 1\%$. De maneira geral o TMCNMD apresentou pior desempenho que o outro concorrente. Seu poder foi muito insatisfatório para $n \leq 50$. Mesmo assim, o TMCNMD apresentou maior poder que o teste quiquadrado de Pearson multivariado proposto por Oliveira e Ferreira (2009) e o teste conjunto de assimetria e curtose multivariado, também avaliado pelos mesmos autores para amostras com tamanhos amostrais $n = 10$ e $n = 20$ referente a todos os números de variáveis apresentados na Tabela 7 para $\alpha = 5\%$. Para $n = 50$ e $n = 100$, relativos a todos os números de variáveis, o TMCNMD apresentou maior poder que o teste quiquadrado de Pearson multivariado, exceto para o número de variáveis iguais a 2 e 10. Para $n = 50$ e $n = 100$, o TMCNMD foi melhor em todas as situações relativos a todos os números de variáveis em comparação com o teste conjunto de assimetria e curtose multivariado ($\alpha = 5\%$). Observou-se que os valores de poder do TMCNMD bem como dos testes quiquadrado multivariado e conjunto de assimetria e curtose multivariado, apresentaram uma queda quando $p = 10$, para todos os tamanhos amostrais apresentados na Tabela 7 e $\alpha = 5\%$.

Como considerações importantes observa-se que para todas as distribuições não-normais consideradas, o poder do TMCNMD foi muito pequeno para amostras

Tabela 6 - Poder dos testes TMCNMD e TSWNM considerando a distribuição mistura de normais multivariadas em função de n e p para $\alpha = 1\%$ e 5%

n	p	1%		5%	
		TNMDMC	TSWNM	TNMDMC	TSWNM
10	2	0,0415	0,0590	0,1490	0,2785
	4	0,0140	0,0580	0,0880	0,3195
	6	0,0095	0,0675	0,0655	0,3285
	8	0,0070	0,0835	0,0500	0,3640
20	2	0,0955	0,2880	0,1985	0,6650
	5	0,0195	0,4725	0,1325	0,7990
	10	0,0265	0,5705	0,1455	0,8640
	18	0,0905	0,6330	0,2640	0,8925
50	2	0,0980	0,9565	0,2640	0,9940
	10	0,0545	0,9980	0,2670	1,0000
	25	0,9830	0,9995	0,9975	1,0000
	48	0,2380	1,0000	0,5240	1,0000
100	2	0,1390	1,0000	0,3495	1,0000
	10	0,0940	1,0000	0,4765	1,0000
	50	1,0000	1,0000	1,0000	1,0000
	98	0,3965	1,0000	0,7725	1,0000
200	2	0,1990	1,0000	0,5310	1,0000
	10	0,2850	1,0000	0,8570	1,0000
	50	1,0000	1,0000	1,0000	1,0000
	100	1,0000	1,0000	1,0000	1,0000
10.000	5	1,0000	-	1,0000	-

pequenas ($n \leq 20$). Para amostras maiores, o desempenho para distribuição assimétrica (lognormal) foi o mesmo observado para distribuição de cauda pesada (t-Student com $\nu = 1$ gl), isto é, o poder decresceu a medida que o número variáveis aumentava e tornava-se próximo do tamanho amostral. Com o aumento do tamanho da amostra o poder do teste proposto estabilizou-se, em geral, em 100%. Para distribuições com cauda muito pesadas, como o caso da uniforme, baixos valores de poder foram observados em geral. O TNMSW apresentou um excelente desempenho em relação ao poder para distribuições assimétrica à direita (lognormal), de cauda pesada e simétrica (t-Student com $\nu = 1$ gl). Para a uniforme, de cauda muito pesada e simétrica, o TSWNM teve menor poder em amostras pequenas e um excelente desempenho em amostras de tamanhos maiores ou iguais a 50. Esse mesmo padrão de resposta para o teste foi observado para a mistura de normais multivariadas. Para o caso da t-Student com $\nu = 30$, tanto o TMCNMD quanto o TSWNM obtiveram valores pequenos de poder. Porém, para $n \geq 30$, o TMCNMD mostrou ter, em muitas configurações em relação a p , melhor desempenho do que o TSWNM.

Tabela 7 - Poder dos testes TMCNMD e TSWNM considerando a distribuição uniforme multivariada em função de n e p para $\alpha = 1\%$ e 5%

n	p	1%		5%	
		TMCNMD	TSWNM	TMCNMD	TSWNM
10	2	0,0210	0,0145	0,0865	0,0970
	4	0,0170	0,0220	0,0760	0,1275
	6	0,0120	0,0280	0,0510	0,1515
	8	0,0075	0,0290	0,0395	0,1505
20	2	0,0320	0,0660	0,1335	0,2740
	5	0,0200	0,1865	0,0780	0,4935
	10	0,0095	0,4280	0,0530	0,7485
	18	0,0165	0,6645	0,0765	0,8840
50	2	0,0665	0,7025	0,2870	0,9295
	10	0,0075	0,9995	0,0425	1,0000
	25	0,0560	1,0000	0,1765	1,0000
	48	0,0745	1,0000	0,1740	1,0000
100	2	0,1355	1,0000	0,5655	1,0000
	10	0,0045	1,0000	0,0385	1,0000
	50	0,7975	1,0000	0,9180	1,0000
	98	0,1855	1,0000	0,2990	1,0000
200	2	0,3915	1,0000	0,9085	1,0000
	10	0,0060	1,0000	0,0620	1,0000
	50	0,9865	1,0000	0,9965	1,0000
	100	1,0000	1,0000	1,0000	1,0000
10.000	5	1,0000	-	1,0000	-

3.3 Exemplo

Para o conjunto de dados, foram estimados o vetor de médias e a matriz de covariâncias amostrais, dados por

$$\bar{\mathbf{X}} = \begin{bmatrix} 15,035 \\ 20,950 \end{bmatrix} \quad \text{e} \quad \mathbf{S} = \begin{bmatrix} 0,8824 & 2,8492 \\ 2,8492 & 37,8395 \end{bmatrix}.$$

Com esses resultados, as distâncias de Mahalanobis cada observação para a média amostral foram computadas e estão apresentadas na Tabela 8. Nesse caso, tem-se $p = 2$ e $n = 20$, então $a = 0$, $b = 0,4412$, $\alpha = 1$ e $\beta = 8,5$. Assim, foram obtidos as estatísticas de ordem observadas da distribuição beta $b_{(j)}$ e os seus respectivos valores esperados b_j^* (Tabela 8). O coeficiente de determinação entre $b_{(j)}$ e b_j^* é $r_c^2 = 0,9384$.

A distribuição nula de r_c^2 foi obtida por simulação Monte Carlo, em um total de 2.000 repetições. O valor- p foi computado, confrontando os valores da distribuição nula Figura 1a e o valor original, resultando em 0,2854. Assim, não se tem evidências significativas ao nível nominal de significância para rejeitar a hipótese de normalidade dos dados. O Q-Q plot originário dos quantis esperados da

Tabela 8 - Valores ordenados das distâncias de Mahalanobis $D_{(j)}^2$ de cada observação para a média amostral, estatísticas de ordem observados relativos a transformação para uma beta $b_{(j)}$ e valores esperados da beta correspondentes a cada estatística de ordem b_j^*

j	$D_{(j)}^2$	$b_{(j)}$	b_j^*
1	0,1648	0,00913	0,00585
2	0,4407	0,02441	0,01197
3	0,4696	0,02602	0,01839
4	0,5691	0,03153	0,02513
5	0,6090	0,03374	0,03225
6	0,6759	0,03745	0,03979
7	0,7433	0,04118	0,04779
8	0,9359	0,05185	0,05634
9	1,0937	0,06059	0,06550
10	1,3547	0,07505	0,07540
11	1,4138	0,07833	0,08616
12	1,4220	0,07878	0,09796
13	1,5961	0,08842	0,11105
14	2,7746	0,15371	0,12577
15	2,9859	0,16542	0,14262
16	3,1270	0,17324	0,16239
17	3,9787	0,22043	0,18644
18	4,2516	0,23554	0,21741
19	4,4920	0,24886	0,26174
20	4,9019	0,27157	0,34567

beta contra os quantis observados está apresentado na Figura 1b. Observa-se que os pontos dos quantis observados e esperados da distribuição beta se situam de forma razoável sobre uma reta de mínimos quadrados ajustada entre eles, cujo coeficiente de determinação é 0,9384, indicando que os dados possam ter sido originários de uma normal bivariada.

Da mesma forma foi aplicado o teste de Shapiro-Wilk de normalidade multivariado de Royston. A estatística, que é uma qui-quadrado, resultou em um valor observado de 0,3548 para os dados dos pintores, com $\nu = 1,9839$ grau de liberdade associado. O valor- p associado, foi de 0,8345, indicando que os dados são normais bivariados, se for considerado o nível nominal de significância de 5%. Ambos os testes apresentaram resultados concordantes nesse exemplo.

Conclusões

O novo teste de normalidade multivariado foi proposto com sucesso.

Não houve um teste uniformemente mais poderoso em todos os casos considerados, sendo que em geral houve um melhor desempenho do TSWNM.

O teste de normalidade multivariada Monte Carlo baseado em distância apresentou controle adequado das taxas de erro tipo I em relação às taxas nominais de significância e poder praticamente equivalente ao teste de normalidade multivariada de Royston (1983, 1993), para grandes amostras. Além do mais, esse

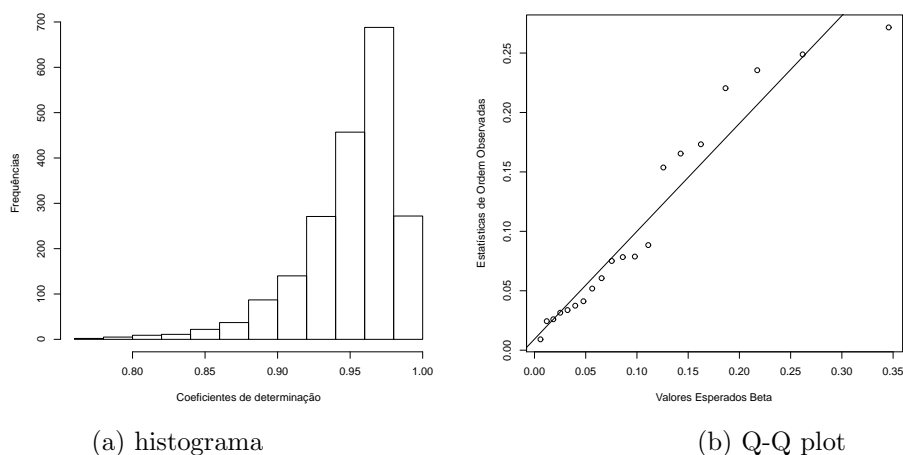


Figura 1 - Histograma da distribuição nula dos coeficiente de determinação gerados sob a hipótese nula de distribuição bivariada (a) e Q-Q plot obtido entre os quantis observados e esperados da distribuição beta para os dados dos pintores de carros.

teste não possui restrição quanto aos tamanhos amostrais que possam limitar a sua aplicação, ao contrário do TSWNM, que é limitado a $n \leq 5.000$.

BIASE, A. G.; FERREIRA, D. F. Computationally intensive multivariate normality test based on Mahalanobis distance. *Rev. Bras. Biom.*, São Paulo, v.30, n.1, p.1-22, 2012.

- **ABSTRACT:** *The multivariate normality tests have direct influence on the quality and reliability of most scientific research that apply multivariate analysis, since the procedures for inference consider the multivariate normal the appropriate model for the data or for the error distribution. Some multivariate normality tests have limitations and for large samples the best of them does not apply. Therefore, this work aimed to propose a new test for multivariate normality without limitation concerning the sample size and to evaluate its performance. The new procedure was a Monte Carlo test for multivariate normality based on distance. The performance was compared to the Royston's test for multivariate normality, considered the best test. To evaluate the properties of the new test Monte Carlo simulation was used. The type I error rates and power were evaluated. All procedures were implemented in the software R. The Monte Carlo multivariate normality test based on distance had great success in controlling the type I error rates and showed power roughly equivalent to the Royston's multivariate normality test for large samples and it also has the advantage of being unlimited regarding the sample size.*
- **KEYWORDS:** *Type I error; multivariate normality; power.*

Referências

- ALVA, J. A. V.; ESTRADA, E. G. A generalization of Shapiro-Wilk's test for multivariate normality. *Commun. Stat.: Theory Methods*, New York, v.38, n.11, p.1870-1883, 2009.
- CIRILLO, M. A.; FERREIRA, D. F. Extensão do teste para normalidade univariado baseado no coeficiente de correlação quantil-quantil para o caso multivariado. *Rev. Mat. Estat.*, São Paulo, v.21, n.3, p.57-75, 2003.
- DOORNIK, J. A.; HANSEN, D. An omnibus test for univariate and multivariate normality. *Working Paper*, Oxford: Nuffield College, 1994. p.1-16.
- DOORNIK, J. A.; HANSEN, H. An Omnibus test for univariate and multivariate normality. *Oxf. Bull. Econ. Stat.*, New York, v.70, p.927-939, 2008.
- FARRELL, P. J.; BARRERA, M. S.; NACZK, K. On tests for multivariate normality and association simulation studies. *J. Stat. Comput. Simul.*, New York, v.77, n.12, p.1053- 1068, 2007.
- GNANADESIKAN, R.; KETTENRING, J.R. Robust estimates, residuals, and outlier detection with multiresponse data. *Biometrics*, Washington, v.28, p.81-124, 1972.
- HENZE, N.; ZIRKLER, B. A class of invariant consistent tests for multivariate normality. *Commun. Stat.: Theory Methods*, New York, v.19, n.10, p.3595-3617, 1990.
- LIANG, J.; TANG, M. L.; CHAN, P. S. A generalized Shapiro-Wilk W statistic for testing high-dimensional normality. *Comput. Stat. Data Anal.*, Amsterdam, v.53, n.1, p.3883-3891, 2009.
- MARDIA, K. V. Applications of some measures of multivariate skewness and kurtosis for testing normality and robustness studies. *Sankhya: Indian J. Stat.*, Calcutta, v.36, p.115-128, 1974.
- MARDIA, K. V. Measures of multivariate skewness and kurtosis with applications. *Biometrika*, London, v.57, n.3, p.519-530, 1970.
- MCNEIL, A. J.; FREY, R. F.; EMBRECHTS, P. *Quantitative risk management: concepts, techniques, and tools*. London: Princeton University Press, 2005. 538p.
- OLIVEIRA, I. R. C.; FERREIRA, D. F. Multivariate extension of chi-squared univariate normality test. *J. Stat. Comput. Simul.*, New York, v.80, n.5, p.513-526, 2009.
- R DEVELOPMENT CORE TEAM. *R: a language and environment for statistical computing*. Vienna, 2011. Disponível em: <<http://www.R-project.org>>. Acesso em: 20 mar. 2011.
- ROYSTON, J. P. A toolkit for testing for non-normality in complete and censored samples. *The Statistician*, London, v.42, n.1, p.37-43, 1993.

ROYSTON, J. P. Some techniques for assessing multivariate based on the Shapiro-Wilk W. *J. Royal Stat. Soc. - Ser C, Appl. Stat.* London, v.32, n.2, p.121-133, 1983.

THODE JR, H. C. *Testing for normality*. New York: Marcel Dekker, 2002. 479p.

Recebido em 09.12.2011.

Aprovado após revisão em 09.04.2012.