



JANAÍNA MARQUES E MELO

**PROPOSTA DE UM TESTE EXATO PARA AVALIAR A
NORMALIDADE MULTIVARIADA BASEADO EM
UMA TRANSFORMAÇÃO t DE STUDENT**

**LAVRAS - MG
2016**

JANAÍNA MARQUES E MELO

**PROPOSTA DE UM TESTE EXATO PARA AVALIAR A NORMALIDADE
MULTIVARIADA BASEADO EM UMA TRANSFORMAÇÃO t DE
STUDENT**

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

Orientador
Dr. Daniel Furtado Ferreira

**LAVRAS - MG
2016**

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Melo, Janaína Marques e.

Proposta de um teste exato para avaliar a normalidade
multivariada baseado em uma transformação t de Student / Janaína
Marques e Melo. – Lavras : UFLA, 2016.

86 p. : il.

Dissertação (mestrado acadêmico)—Universidade Federal de
Lavras, 2016.

Orientador(a): Daniel Furtado Ferreira.

Bibliografia.

1. Teste de Normalidade Multivariada. 2. Gráfico de Quantil-
Quantil. 3. Distribuição t de Student. I. Universidade Federal de
Lavras. II. Título.

JANAÍNA MARQUES E MELO

**PROPOSTA DE UM TESTE EXATO PARA AVALIAR A NORMALIDADE
MULTIVARIADA BASEADO EM UMA TRANSFORMAÇÃO t DE
STUDENT**

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

APROVADA em 21 de janeiro de 2016.

Dra. Silvia Maria de Freitas

UFC

Dra. Izabela Regina Cardoso de Oliveira

UFLA

Dr. Júlio Silvio de Sousa Bueno Filho

UFLA

Dr. Daniel Furtado Ferreira
Orientador

**LAVRAS - MG
2016**

AGRADECIMENTOS

Agradeço a Deus por está presente na minha caminhada, eu não sou nada sem o seu amor, e por me proporcionar momentos de superações e alegrias.

À minha família por ser o meu alicerce. Minha mãe Núbia e o meu pai Dionízio por não deixar faltar nada, por me ajudarem com palavras de conforto e por fazerem o possível e às vezes até um pouco mais, com a intenção de me deixar confortável. Meus irmãos, João Dionízio e José Marques, por tornarem os meus dias mais felizes. À minha querida avó Consuêlo por ser uma pessoa forte e determinada. Agradeço também minhas tias, tios, primos e sobrinho por serem maravilhosos comigo e sempre presente.

Ao meu orientador Daniel Furtado que para mim é um exemplo de profissional que transborda conhecimento. Obrigada professor por passar um pouco do seu conhecimento e por transmitir a sua alegria e satisfação em trabalhar com multivariada. Sou muito grata por ter vivenciado esses momentos de sabedoria.

Aos membros da banca, Sílvia e Izabela, por aceitarem faz parte desse etapa importante para mim. À professora Silvia por está nessa minha caminhada desde a graduação, ajudando de alguma forma no meu crescimento profissional. À professora Izabela pelos conselhos, sugestões e ensinamentos.

À Nádia Ferreira por ser prestativa e atenciosa comigo.

Aos companheiros de jornada, minha turma de mestrado, pela ajuda e por compartilharem conhecimentos acadêmicos e pessoais. Obrigada pelos ensinamentos de vida e profissional Lílian, Ricardo, Débora, Renata, Marcel, Carlos e Ismael.

À Kelly e à Sidcleide por terem paciência comigo e deixarem meus dias mais divertidos e alegres.

Aos meus eternos amigos de graduação, Bruno, Laura, Raquel, Yuri, que mesmo estando longe, se fazem presentes na minha vida.

Enfim, agradeço a todos que, de alguma maneira, colaboraram com esse trabalho e com o meu crescimento acadêmico e pessoal.

RESUMO

A distribuição normal é uma das mais importantes distribuições contínuas da Estatística. Além dessa distribuição descrever uma série de fenômenos, ela é de grande uso na estatística inferencial. Vale ressaltar que o teste de normalidade influencia diretamente na qualidade e confiabilidade das pesquisas científicas, uma vez que a distribuição normal faz parte da suposição de diversos procedimentos estatísticos, e a não checagem dessa pressuposição pode conduzir a resultados e conclusões incorretas. Na multivariada, isso não é diferente. Uma maneira simples, porém subjetiva, de se verificar a normalidade de uma distribuição tanto univariada quanto multivariada é por meio de gráficos, como o gráfico quantil-quantil (*Q-Q plot*). Além disso, os *Q-Q plots* são ferramentas viáveis para a visualização de valores discrepantes. Uma desvantagem do *Q-Q plot* clássico é que os quantis observados não são independentes sendo apenas identicamente distribuídos, tornando-se independentes apenas assintoticamente. Isso compromete o *Q-Q plot* ou qualquer teste que tenha o mesmo princípio básico. O objetivo do presente trabalho é propor um teste exato baseado na distribuição *t* de Student, validar o seu desempenho por simulação Monte Carlo e desenvolver um *Q-Q plot* para fornecer provas suplementares para detectar uma possível normalidade multivariada na análise de dados em p dimensões. Este *Q-Q plot* provém de uma caracterização da distribuição normal multivariada feita por Yang *et al.* (1996) baseada em uma das propriedades da distribuição esférica (FANG *et al.*, 1990). Foi utilizado o programa R versão 3.1.0 de programação livre, e fonte aberta, para auxiliar na construção desse gráfico, bem como nas simulações de validação do teste. O desempenho, validado por simulação Monte Carlo, mostrou que o teste proposto teve sucesso no controle das taxas de erro tipo I , sendo um teste exato, porém foi pouco poderoso.

Palavras-chave: Teste de Normalidade Multivariada. Gráfico de Quantil-Quantil. Distribuição *t* de Student.

ABSTRACT

The normal distribution is one of the most important continuous probability distribution. This distribution describe several phenomena and has great hole in inferential statistics. It is noteworthy that the normality directly influences the quality and reliability of scientific research since violations of assumption can lead to incorrect results and conclusions. The same is expected for multivariate inferences. A simple manner, however subjective, to verify the univariate or multivariate normality is through quantile-quantile plots (Q-Q plots). Furthermore, the Q-Q plots are efficient tools for the visualization of outliers. A disadvantage of the classical Q-Q plot is that the quantiles are only asymptotically identically distributed, but they are not independent. This fact compromises the efficiency of the Q-Q plot or any test based on the use of the observed distance quantiles. The objective of this study is to propose an accurate test and validate its performance by Monte Carlo simulation and also provide a Q-Q plot to detect further evidence of violation of multivariate normality in p dimensions. This Q-Q plot originates from a characterization of the multivariate normal distribution made by Yang *et al.* (1996) based on the spherical distribution properties (Fang *et al.*, 1990). The *R* program version 3.1.0 was used to build this Q-Q plot normality test and to perform the validation of its performance by Monte Carlo simulations. The Monte Carlo simulation results showed that the proposed test successful controls the type I error rates being accurate, but shows lower power than any other multivariate normality test.

Keywords: Multivariate Normality Test. Quantile-Quantile graph. t-Student distribution.

Lista de Figuras

Figura 1	Regiões críticas para testes de hipóteses	15
Figura 2	Regra de decisão para testes de hipóteses bilaterais	16
Figura 3	Hemoglobina e linfócitos	65
Figura 4	Q-Q plot fornecido pelo teste Monte Carlo de normalidade multi- variada baseado em distâncias	67
Figura 5	Q-Q plot fornecido pelo teste de Henze-Zirkler	69
Figura 6	t-plot fornecido pelo teste proposto	70

Lista de Tabelas

Tabela 1	Tipos de erros em teste de hipóteses e suas probabilidades.	16
Tabela 2	Os testes de normalidade multivariados usados nas comparações, com os tamanhos amostrais n e número de variáveis p usados nas construções das matrizes de dados $\mathbf{X}_{n \times p}$, considerando os teste usados no trabalho.	39
Tabela 3	Erro tipo I dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição normal multivariada em função de n e p para $\alpha = 10\%$	42
Tabela 4	Erro tipo I dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição normal multivariada em função de n e p para $\alpha = 5\%$	43
Tabela 5	Erro tipo I dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição normal multivariada em função de n e p para $\alpha = 1\%$	45
Tabela 6	Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição <i>t-Student</i> multivariada com $\nu = 1$ grau de liberdade em função de n e p para $\alpha = 10\%$	48
Tabela 7	Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição <i>t-Student</i> multivariada com $\nu = 1$ grau de liberdade em função de n e p para $\alpha = 5\%$	49
Tabela 8	Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição <i>t-Student</i> multivariada com $\nu = 1$ grau de liberdade em função de n e p para $\alpha = 1\%$	51
Tabela 9	Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição <i>t-Student</i> multivariada com $\nu = 30$ grau de liberdade em função de n e p para $\alpha = 10\%$	53
Tabela 10	Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição <i>t-Student</i> multivariada com $\nu = 30$ grau de liberdade em função de n e p para $\alpha = 5\%$	54
Tabela 11	Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt conside-	

	rando a distribuição <i>t-Student</i> multivariada com $\nu = 30$ grau de liberdade em função de n e p para $\alpha = 1\%$	55
Tabela 12	Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição log-normal multivariada em função de n e p para $\alpha = 10\%$	57
Tabela 13	Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição normal contaminada multivariada em função de n e p para $\alpha = 10\%$	59
Tabela 14	Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição uniforme multivariada em função de n e p para $\alpha = 10\%$	61
Tabela 15	Tamanho da amostra efetiva utilizada no TNMt.	63
Tabela 16	Conjunto de dados referente à concentração de hemoglobina e contagem de linfócitos apresentado por Royston (1983).	65
Tabela 17	Resultados dos testes de normalidade multivariados fornecidos pelo programa <i>R</i>	70

SUMÁRIO

1	INTRODUÇÃO	11
2	REFERENCIAL TEÓRICO	13
2.1	Teste de Hipóteses	13
2.1.1	Estrutura do Teste de Hipóteses	14
2.1.2	Erro tipo I e II	16
2.1.3	Poder do Teste	17
2.1.4	valor-p	18
2.2	Testes de Aderência	18
2.3	Testes de Normalidade	20
2.3.1	Teste de Shapiro-Wilk de normalidade multivariada	22
2.3.2	Teste de Doornik-Hansen de normalidade multivariada	23
2.3.3	Teste de Henze-Zirkler de normalidade multivariada	26
2.4	Gráfico de Quantil-Quantil para dados Normais Multivariados usando transformação t de Student	29
2.5	Programa R	31
2.5.1	Teste de Shapiro-Wilk de normalidade multivariada do pacote 'mvnormtest'	32
2.5.2	Teste de Doornik-Hansen de normalidade multivariada do pa- cote 'asbio'	33
2.5.3	Teste de Henze-Zirkler de normalidade multivariada do pacote 'MVN'	34
2.6	Simulação Monte Carlo	34
3	MÉTODO	37
3.1	Proposta do teste de Normalidade Multivariada	37
3.2	Validação do desempenho do teste	38
4	RESULTADOS E DISCUSSÃO	40
4.1	Erro tipo I	40
4.2	Poder do teste	46
4.3	Aplicação	64
5	CONCLUSÃO	72
	REFERÊNCIAS	73
	ANEXO - Comandos usadas no R	77
	APÊNDICE - Comandos usadas no R	82

1 INTRODUÇÃO

A distribuição Normal é uma das mais importantes distribuições contínuas da estatística, conhecida também como distribuição de Gauss ou Gaussiana. Por descrever o comportamento de diversos fenômenos aleatórios, de forma exata ou aproximada, possui grande uso na estatística inferencial.

Vale ressaltar que a suposição de normalidade univariada ou multivariada é exigida em muitos modelos estatísticos e métodos de análise de dados. Por exemplo, em regressão linear testa-se a normalidade dos resíduos, caso essa premissa seja inválida, os resultados da regressão (intervalos de confiança, etc.) não podem ser usados. Destaca-se, então, a importância do teste de normalidade, que influencia diretamente na qualidade e confiabilidade das pesquisas científicas, uma vez que a não verificação da suposição de normalidade pode conduzir a resultados e conclusões equivocadas.

Outra forma de verificar a normalidade dos dados é por meio de análise gráfica, como histogramas e o gráfico de quantil-quantil (também conhecido na literatura por *Q-Q plot*), embora esse procedimento contemple uma certa subjetividade na avaliação do mesmo. Além disso, essas análises gráficas, por serem subjetivas, não são suficientes para se fazer inferência sobre a normalidade, sendo necessária a utilização de testes estatísticos. No caso multivariado, principalmente, isso é mais evidente, tendo em vista a complexa relação existente entre as variáveis envolvidas. Apesar disso, o gráfico de quantil-quantil é uma ferramenta bastante útil para a visualização de valores discrepantes da amostra. Nesses gráficos são *plotados* os percentis amostrais *versus* os percentis esperados pelo ajuste de uma distribuição normal. Se os pontos estiverem localizados próximos a uma reta de mínimos quadrados a suposição de normalidade deve ser aceita. No caso multivariado, as observações são transformadas em distâncias de Mahalanobis para o vetor de médias amostral e os quantis amostrais e esperados dessas distâncias são *plotados* para a obtenção do gráfico de quantil-quantil.

Os testes de normalidade são usados para determinar se um conjunto de dados de uma dada variável aleatória é bem modelado por uma distribuição normal ou não, que equivale a verificar se a variável aleatória subjacente é normalmente

distribuída. No caso univariado, os testes mais encontrados na literatura são: o teste de aderência qui-quadrado, Kolmogorov-Smirnov, Lilliefors e Shapiro-Wilk. No caso multivariado, em geral, os testes existentes são extensões daqueles propostos para o caso univariado (CIRILLO; FERREIRA, 2003).

Uma desvantagem do gráfico de quantil-quantil multivariado, que é uma das bases para a construção do teste que está sendo proposto nesse trabalho, é que os quantis observados não são independentes, sendo apenas identicamente distribuídos, tornando-se independentes apenas assintoticamente. Isso compromete o desempenho do procedimento gráfico quantil-quantil, bem como de qualquer teste que utilize os quantis amostrais observados.

Como já foi afirmado anteriormente, vários testes existentes de normalidade multivariada fazem uso da transformação em distâncias de cada vetor aleatório observado na amostra, que gera, por sua vez, uma amostra de distâncias dependentes. Esse trabalho visa trocar essa transformação gerando dados independentes. Essa transformação é centrada na caracterização da distribuição normal multivariada feito por Yang *et al.* (1996) baseada em uma das propriedades da distribuição esférica (FANG *et al.*, 1990).

O presente trabalho tem como principal objetivo propor um novo teste exato de normalidade multivariada, tendo como objetivos específicos: validar o desempenho do mesmo por simulação Monte Carlo e desenvolver um gráfico de quantil-quantil (*Q-Q plot*), utilizando o programa *R* como ferramenta para as simulações e para a elaboração do gráfico de quantil-quantil.

Esse trabalho está dividido em cinco seções. A Seção 2 revisa alguns conceitos sobre teste de hipóteses, teste de aderência, teste de normalidade univariado e multivariado, construção do gráfico de quantil-quantil para avaliar a normalidade multivariada e simulação de Monte Carlo. Na Seção 3 encontra-se a metodologia usada no trabalho, onde são exploradas a proposta do novo teste de normalidade multivariada e a validação do mesmo. Na Seção 4 estão apresentados os resultados da simulação como a taxa do erro tipo I, o poder do teste e um exemplo de aplicação de alguns testes de normalidade multivariado. Por fim, na Seção 5 são apresentadas considerações finais, seguida pelas referências, anexos e apêndices.

2 REFERENCIAL TEÓRICO

2.1 Teste de Hipóteses

Frequentemente encontramos situações em que são necessárias tomar decisões de aceitar ou rejeitar determinadas afirmações baseadas em um conjunto de evidências. Um exemplo comum é o caso em que um indivíduo está sendo julgado por determinado delito. Com base nas evidências (testemunhas, fatos, etc.), o júri terá que decidir pela culpa ou inocência do indivíduo. Pode-se, então, concluir que o júri formula duas hipóteses: “ H_0 : o indivíduo é inocente” e a alternativa “ H_1 : o indivíduo é culpado”. Com base nas evidências apresentadas, o júri terá que se decidir por H_0 ou por H_1 . Ao julgar, por exemplo, a decisão de não rejeitar H_1 (então, rejeita-se H_0) como verdadeira, o júri pode estar cometendo um erro, pois, apesar das evidências, o indivíduo pode ser inocente. O mesmo pode acontecer com relação a não rejeição da hipótese H_0 como verdadeira. Nesse caso, o júri estaria considerando como inocente um indivíduo culpado (BOLFARINE; SANDOVAL, 2010).

Definição 1. *Uma hipótese estatística é qualquer afirmação ou conjectura que se faça sobre a distribuição de uma ou mais variáveis aleatórias. Uma hipótese que especifique completamente a distribuição é chamada de hipótese simples; caso contrário, a hipótese é dita composta.*

Denota-se por H_0 , hipótese nula, a hipótese de interesse e por H_1 , hipótese alternativa, a hipótese que confronta a de interesse.

O objetivo de um teste de hipótese é estabelecer uma regra λ para se rejeitar, ou não, a hipótese nula em favor da hipótese alternativa, dado que se tem uma amostra aleatória (a.a.) X_1, X_2, \dots, X_n .

Então, dadas as hipóteses H_0 e H_1 , encontra-se uma estatística de teste, determina-se a condição de rejeição de H_0 (região crítica - RC) e estabelece regras de avaliação da qualidade para o teste.

2.1.1 Estrutura do Teste de Hipóteses

Segundo Bussab e Morettin (2004), para a construção de um teste de hipóteses são sugeridos os seguintes passos:

Passo 1: Define-se a hipótese H_0 a ser testada e a alternativa H_1 ;

O primeiro passo é estabelecer a hipótese nula (H_0) e alternativa (H_1). A hipótese nula (H_0) é um valor suposto para o parâmetro de interesse e só será não rejeitada se os resultados da amostra não forem muito diferentes do que foi proposto por H_0 . A hipótese alternativa (H_1) é uma hipótese que contraria a hipótese nula, e só será não rejeitada se os resultados forem muito diferentes do que foi proposto por H_0 .

Passo 2: Determina-se a estatística usada no teste;

Usa-se teoria estatística e as informações disponíveis para decidir qual estatística (estimador) será usada para testar a hipótese nula, obtendo-se suas propriedades (distribuição, estimativa, erro padrão).

Passo 3: Estabelecer a área da região crítica;

A região crítica (RC) é a região em que se rejeita H_0 . A área da RC é igual ao nível de significância (α) de cometer o erro tipo I, que é a probabilidade de rejeitar hipótese nula quando ela é verdadeira, usando-se este valor para construir a RC. Lembre-se que a RC é construída para a estatística definida no **Passo 2**, usando os valores hipotetizados por H_0 . Nas figuras 1(a), 1(b) e 1(c), tem-se a região crítica para um teste unilateral à direita, um teste unilateral à esquerda e um teste bilateral, respectivamente.

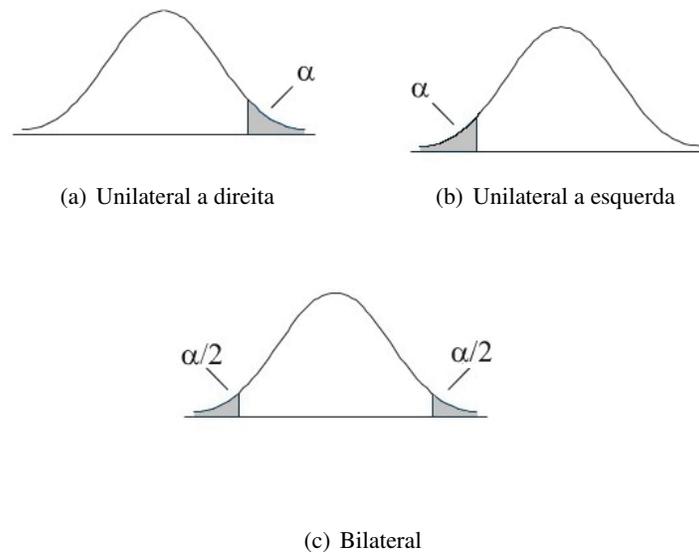


Figura 1 Regiões críticas para testes de hipóteses

Passo 4: Calcula-se a estatística do teste;

Usa-se as informações da amostra para calcular o valor da estatística do teste. A estatística do teste é o valor calculado a partir da amostra, que será usado na tomada de decisão. Uma maneira de tomar uma decisão é comparando o valor tabelado com o valor da estatística do teste.

Passo 5: Regra de decisão e conclusão;

Na regra de decisão, se o valor da estatística do teste estiver dentro da região crítica, rejeita-se H_0 . Ao rejeitar a hipótese nula existe uma forte evidência de sua falsidade. Caso contrário, quando não se rejeita H_0 , certifica-se que não houve evidência amostral significativa para permitir a rejeição de H_0 . Na Figura 2 tem-se a ilustração da área para a decisão de um teste bilateral.



Figura 2 Regra de decisão para testes de hipóteses bilaterais

2.1.2 Erro tipo I e II

Quando se realiza um teste de hipóteses pode-se cometer dois tipos de erro, pelo fato de se estar usando resultados amostrais para fazer inferência sobre a população. O erro tipo I ocorre quando se rejeita a hipótese nula, sendo ela verdadeira. Esse erro é controlado diretamente pelo pesquisador e a probabilidade de se cometê-lo é dada pelo nível de significância, α . O erro tipo II ocorre quando se não rejeita a hipótese nula, sendo ela falsa. Esse erro não é controlado diretamente pelo pesquisador e a probabilidade de se cometer esse erro é denotada por β (CASELLA; BERGER, 2001).

Formalizando a probabilidade de se cometer cada um desses erros, tem-se que:

$$\mathbb{P}[\text{erro tipo I}] = \mathbb{P}[\theta \in RC | H_0 \text{ é verdadeira}] = \alpha$$

e

$$\mathbb{P}[\text{erro tipo II}] = \mathbb{P}[\theta \notin RC | H_1 \text{ é verdadeira}] = \beta,$$

em que RC é a região crítica.

Outra forma de visualizar esses dois tipos de erros e as decisões corretas com as respectivas probabilidades está descrita na Tabela 1.

Tabela 1 Tipos de erros em teste de hipóteses e suas probabilidades.

Realidade	Decisão	
	Rejeita H_0	Não rejeita H_0
H_0 verdadeira	Erro tipo I (α)	Decisão correta ($1 - \alpha$)
H_0 falsa	Decisão correta ($1 - \beta$)	Erro tipo II (β)

Fonte: Ferreira (2005)

As probabilidades de se cometerem os erros do tipo I e II são inversamente proporcionais. Então, é necessário ficar atento para não se aplicar um teste a um nível muito baixo de probabilidade, pois isso pode causar um aumento muito expressivo na probabilidade do outro erro.

2.1.3 Poder do Teste

O poder do teste é a probabilidade de se rejeitar a hipótese nula (H_0), quando ela é realmente falsa (MOOD; GRAYBILL; BOES, 1974).

Definição 2. *O poder do teste com região crítica RC para testar $H_0 : \theta = \Theta_0$ contra $H_1 : \theta = \Theta_1$ é dado por*

$$\pi(\Theta_1) = \mathbb{P}_{H_1}[X \in RC] = \mathbb{P}[X \in RC | \Theta_1].$$

Note que $\pi(\Theta_1) = 1 - \beta$, em que β é a probabilidade de se cometer o erro tipo II.

O poder do teste informa quanto o teste de hipótese controla o erro do tipo II, ou qual a probabilidade de rejeitar a hipótese nula se realmente for falsa. Na prática, é importante que se tenham testes com nível de significância próximos do nível de significância nominal e que o poder seja alto, mesmo em situações de amostras pequenas.

O poder de um teste de hipóteses é afetado por três fatores:

- i)* **tamanho da amostra:** mantendo todos os outros parâmetros iguais, quanto maior o tamanho da amostra, maior o poder do teste;
- ii)* **nível de significância:** quanto maior o nível de significância, maior o poder do teste.
- iii)* **o verdadeiro valor do parâmetro a ser testado:** quanto maior a diferença entre o “verdadeiro” valor do parâmetro e o valor especificado pela hipótese nula, maior o poder do teste (PORTAL ACTION, 2015).

2.1.4 valor- p

O valor- p , também denominado nível descritivo do teste, é a probabilidade de que a estatística do teste (como variável aleatória) tenha valor extremo em relação ao valor observado (estatística) quando a hipótese nula é verdadeira.

A definição do valor- p por Casella e Berger (2001) é dada por:

Definição 3. *O valor- p , $\mathbb{P}[\mathbf{X}]$, é uma estatística de teste que satisfaz $0 \leq \mathbb{P}[\mathbf{X}] \leq 1$ para cada ponto amostra \mathbf{x} . Pequenos valores de $\mathbb{P}[\mathbf{X}]$ fornecem evidências de que H_1 é verdadeira. O valor- p é válido se, para cada $\theta \in \Theta_0$ e cada $0 \leq \alpha \leq 1$,*

$$\mathbb{P}_\theta[\mathbb{P}[\mathbf{X}] \leq \alpha] \leq \alpha.$$

Se $\mathbb{P}[\mathbf{X}]$ é um valor- p válido, é fácil construir um teste de nível α baseado em $\mathbb{P}[\mathbf{X}]$. O teste que rejeita H_0 , se e somente se $\mathbb{P}[\mathbf{X}] \leq \alpha$ é um teste de nível α , pois $\mathbb{P}_\theta[\mathbb{P}[\mathbf{X}] \leq \alpha] \leq \alpha$. Uma vantagem para relatar um resultado de teste por meio de um valor- p é que cada leitor pode escolher o valor de α adequado e, ao comparar com o valor relatado em $\mathbb{P}[\mathbf{X}]$, conclui-se se esses dados levam à rejeição ou não da hipótese nula (CASELLA; BERGER, 2001).

Também pode-se interpretar o valor- p como o menor valor do nível de significância para o qual se rejeita H_0 . Desta forma, se o nível de significância (α) proposto para o teste for menor que o valor- p , não se rejeita a hipótese nula.

2.2 Testes de Aderência

Os testes de aderência encontram-se dentre dos testes não-paramétricos e são chamados assim porque não é necessário especificar a distribuição da população de onde provém a amostra.

Em um teste de aderência admite-se, por hipótese, que a distribuição da variável em estudo segue o comportamento de uma distribuição teórica de probabilidade, na população.

Dentre os testes de aderência mais comuns estão o de qui-quadrado e o de Kolmogorov-Smirnov. O teste de Kolmogorov-Smirnov é baseado na função

de distribuição empírica (FDE) dos dados e apresenta vantagens sobre o teste qui-quadrado, incluindo maior poder e invariância em relação aos pontos médios dos intervalos escolhidos.

No método de Kolmogorov-Smirnov, a estatística do teste é a maior diferença observada entre a função de distribuição acumulada da distribuição teórica e a distribuição empírica da variável em estudo.

O teste consiste na obtenção do valor

$$D^+ = \max |F_{obs}(X) - F_{esp}(X)|, \quad (1)$$

em que $F_{obs}(X)$ é a função de distribuição acumulada dos dados observados e $F_{esp}(X)$ é a função de distribuição acumulada esperada. O teste fundamenta-se no valor obtido pela expressão (1) e na sua comparação com um valor crítico tabelado em função do nível de significância (α) e do tamanho da amostra (n).

O teste é unilateral, rejeitando-se a hipótese nula de que a variável em estudo segue a distribuição de probabilidade especificada se D^+ for maior que o valor crítico.

O teste qui-quadrado, para ser realizado, os dados precisam estar agrupados em k classes (intervalos ou categorias). No caso em que a distribuição \mathcal{D} é contínua, tais classes podem ser baseadas nas classes do histograma. São comparadas duas quantidades: (i) o número de valores observados em cada categoria (frequência observada, n_i); (ii) o número de valores que se teriam em cada categoria admitindo que a população tem a distribuição \mathcal{D} (frequência esperada, e_i).

Se as diferenças entre $n_i - e_i$, $i = 1, \dots, k$, forem “pequenas”, então não se rejeita a hipótese nula de que a população tem uma determinada distribuição de probabilidade. A estatística do teste é dada por:

$$\chi_c^2 = \sum_{i=1}^k \frac{(n_i - e_i)^2}{e_i}, \quad (2)$$

em que n_i é a frequência observada na i -ésima classe, e_i é a frequência esperada na i -ésima classe e k é o número de classes da distribuição de frequência. Não se rejeita a hipótese nula de que a variável em estudo segue a distribuição de probabi-

lidade especificada quando $\chi_c^2 < \chi_t^2$, em que χ_t^2 é o valor crítico obtido na tabela em função do nível de significância α e $\nu = k - p - 1$ graus de liberdade em que p é o número de parâmetros estimados.

2.3 Testes de Normalidade

Muitos procedimentos de inferência são baseados em um modelo normal da população sob amostragem (FERREIRA, 2005), pois a normalidade é a mais usual suposição feita no desenvolvimento e no uso dos procedimentos estatísticos (THODE JÚNIOR, 2002). Em geral, os testes são feitos para avaliar a veracidade de alguma hipótese de interesse, que é qualquer conjectura sobre aspectos desconhecidos da população. Estes aspectos podem ser parâmetros, ou mesmo a família de distribuições ou a forma da distribuição (PEREIRA FILHO, 2013). Os testes estatísticos usados para verificar se um conjunto de dados de uma variável aleatória pode ser modelado por uma distribuição normal são chamados de testes de normalidade.

É importante salientar que existem inúmeros testes de normalidade na literatura. No caso univariado, alguns desses testes são: Shapiro-Wilk (1965), Shapiro-Francia, qui-quadrado de Pearson, Lilliefors, Kolmogorov-Smirnov e Jarque-Bera.

Segundo Pereira Filho (2013), os procedimentos estatísticos como o teste t , o teste F na análise de variância e na verificação da homogeneidade de variâncias, os testes dos coeficientes de regressão entre outros casos possuem uma pressuposição subjacente de que os dados amostrais ou os resíduos do modelo adotado se distribuam como normais. Essa pressuposição deve ser verificada por um teste que possa efetivamente constatar se ela foi atendida ou não.

A inferência na estatística univariada é mais simples de ser realizada, pois a interpretação é feita de forma isolada para cada variável. Entretanto, apresenta algumas desvantagens como a perda de informações, uma vez que as relações existentes entre as variáveis não são consideradas, pois não basta conhecer informações das estatísticas isoladas, mas sim conhecer a totalidade destas informações fornecidas pelo conjunto de todas as variáveis. Assim, uma generalização do caso

univariado é a análise estatística multivariada, que se preocupa com os métodos e técnicas que utilizam simultaneamente todas as variáveis na interpretação teórica do conjunto de dados (BIASE, 2011).

Seja uma amostra aleatória de tamanho n no espaço p dimensional, dada por $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$, de uma normal multivariada com vetor de médias $\boldsymbol{\mu}$ com dimensão $(p \times 1)$ e matriz de covariâncias $\boldsymbol{\Sigma}$ positiva definida de dimensão $(p \times p)$, dados por:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} \quad \text{e} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}.$$

Então, sua função densidade de probabilidade conjunta $f_{\mathbf{X}}(\mathbf{x})$, é dada por:

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-\frac{p}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}. \quad (3)$$

A normalidade univariada ou multivariada, como já mencionado anteriormente, é a pressuposição básica para a maioria dos testes e procedimentos de estimação. Um dos primeiros pensamentos, ao se testar a normalidade multivariada em um banco de dados, é verificar se cada uma das marginais (as variáveis isoladamente) seguem uma distribuição normal univariada. Entretanto, não é suficiente testar as distribuições marginais univariadas, pois se as marginais são normais, não há garantia de que a conjunta seja normal multivariada. No entanto, se a conjunta for normal multivariada as marginais serão normais univariadas.

Diante desses problemas pode-se usar, como alternativa para verificar a normalidade multivariada, os testes de: Royston, Shapiro-Wilk, Shapiro-Francia, Kolmogorov-Smirnov, entre outros.

Farrell, Barrera e Naczki (2007) em seus estudos sobre testes de normalidade multivariada dividem esses testes em três principais categorias: (i) técnica da falta de ajuste, (ii) procedimentos baseados em assimetria e curtose e (iii) testes consistentes e invariantes, optando-se por escolher o teste mais poderoso de cada categoria.

Segundo Biase (2011) o teste de Royston (1992), entre os testes baseados na técnica da falta de ajuste, apresentou bom desempenho de poder. Na classe de procedimentos baseados em assimetria e curtose, o teste de Doornik e Hansen (1994) possui bom desempenho de poder, principalmente nos casos em que as amostras são maiores ou iguais a 75. O teste proposto por Henze e Zirkler (1990) para a distribuição *t-student* multivariada apresentou alto poder, sendo classificado entre os testes consistentes e invariantes.

2.3.1 Teste de Shapiro-Wilk de normalidade multivariada

Ferreira (2011) apresenta a estrutura do teste proposto por Royston (1983), que é uma generalização do teste de Shapiro-Wilk (1965) para o caso multivariado, da seguinte forma:

- 1) Seja $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ uma amostra aleatória, em que $\mathbf{X}_j \in \mathbb{R}^p$ e seja $X_{(1)k}, X_{(2)k}, \dots, X_{(n)k}$ a amostra aleatória correspondente a k -ésima variável arranjada em ordem crescente de magnitude. Royston (1983) propõem calcular para cada variável a estatística W_k , dada por:

$$W_k = \frac{\left[\sum_{j=1}^n \tilde{a}_j X_{(j)k} \right]^2}{\sum_{j=1}^n (X_{jk} - \bar{X}_{\cdot k})^2},$$

em que $\bar{X}_{\cdot k} = \sum_{j=1}^n X_{jk}/n$ e \tilde{a}_j é o estimador do coeficiente associado a j -ésima estatística de ordem. O vetor $\mathbf{a} = [a_1, a_2, \dots, a_n]^\top$ é definido por:

$$\mathbf{a} = \frac{\mathbf{V}^{-1}\mathbf{m}}{\sqrt{\mathbf{m}^\top \mathbf{V}^{-2}\mathbf{m}}},$$

em que $\mathbf{m} = [m_1, m_2, \dots, m_n]^\top$ é o vetor de médias das estatísticas de ordem da normal padrão e $\mathbf{V} = [v_{ij}]$, a matriz de covariâncias de dimensões $(n \times n)$.

- 2) Transforma-se a estatística W_k em uma normal padrão Z_k e calcula-se a quantidade κ_k , dado por:

$$\kappa_k = \left\{ \phi^{-1} \left[\frac{1}{2} \phi(-Z_k) \right] \right\}^2, \quad k = 1, 2, \dots, p$$

em que $\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\{-t^2/2\} dt$.

- 3) Calcula-se a estatística do teste definida por Royston (1993), dada por:

$$H = \frac{\nu}{p} \sum_{k=1}^p \kappa_k,$$

que sob a hipótese nula de normalidade possui distribuição aproximada de qui-quadrado com ν graus de liberdade.

2.3.2 Teste de Doornik-Hansen de normalidade multivariada

O teste de Doornik-Hansen de normalidade multivariada (Doornik e Hansen, 1994) é baseado na assimetria e curtose dos dados multivariados. Considera-se os vetores aleatórios $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ de dimensão $1 \times p$, que formam uma matriz \mathbf{X} com dimensões $(n \times p)$, com média e matriz de covariâncias amostrais dados por:

$$\bar{\mathbf{X}} = \frac{\sum_{j=1}^n \mathbf{X}_j}{n} \quad \text{e} \quad \hat{\Sigma} = \frac{(\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})^\top}{n}.$$

Define-se a matriz \mathbf{V} , que é constituída por variâncias em sua diagonal:

$$\mathbf{V} = \text{diag}(\hat{\sigma}_1^2, \dots, \hat{\sigma}_n^2)$$

e a matriz de correlação

$$\mathbf{C} = \mathbf{V}^{-1/2} \hat{\Sigma} \mathbf{V}^{-1/2}.$$

Posteriormente é definido os vetores $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$ de dimensão $(1 \times p)$ com a seguinte transformação nos dados observados:

$$\mathbf{Y}_j = \mathbf{H}\Psi^{-1/2}\mathbf{H}^\top\Psi^{-1/2}(\mathbf{X}_j - \bar{\mathbf{X}}),$$

em que $\Psi = \text{diag}(\lambda_1, \dots, \lambda_p)$ é a matriz diagonal com os autovalores de \mathbf{C} na diagonal. As colunas de \mathbf{H} são correspondentes aos autovetores associados aos respectivos autovalores de \mathbf{C} , desde que $\mathbf{H}^\top\mathbf{H} = \mathbf{I}$ e $\Psi = \mathbf{H}^\top\mathbf{C}\mathbf{H}$, em que \mathbf{I} é a matriz identidade.

Calcula-se a estatística do teste de assimetria e curtose univariada para cada p variável transformada, de tamanho amostral n . Definindo-se:

$$\mathbf{B}_1^\top = (\sqrt{b_{11}}, \dots, \sqrt{b_{1p}}) \quad \text{e} \quad \mathbf{B}_2^\top = (b_{21}, \dots, b_{2p}),$$

em que

$$\sqrt{b_1} = \frac{m_3}{(m_2)^{3/2}} \quad \text{e} \quad b_2 = \frac{m_4}{m_2},$$

sendo m_r o momento central amostral de ordem r . A estatística do teste de Doornik e Hansen (1994) é estabelecida por:

$$\mathbb{E}_p^a = \frac{n\mathbf{B}_1^\top\mathbf{B}_1}{6} + \frac{n(\mathbf{B}_2 - 3\mathbf{1})^\top(\mathbf{B}_2 - 3\mathbf{1})}{24} \sim \chi_{(2p)}^2,$$

em que $\mathbf{1}$ é um vetor de dimensão $(p \times 1)$ de 1s. Vale ressaltar que a estatística só segue uma qui-quadrado com $2p$ graus de liberdade assintoticamente.

Os autores propuseram realizar as transformações de $\sqrt{b_{1i}}$ e b_{2i} em variáveis normais padrões, como proposto por D'Agostino (1970), invés de utilizar diretamente os coeficientes de assimetria e curtose. A estatística do teste de normalidade multivariada de Doornik e Hansen (1994) é obtida da seguinte forma

$$\mathbb{E}_p = \mathbf{Z}_1^\top\mathbf{Z}_1 + \mathbf{Z}_2^\top\mathbf{Z}_2 \sim \chi_{2p}^2,$$

em que $\mathbf{Z}_1^\top = (z_{11}, \dots, z_{1p})$ e $\mathbf{Z}_2^\top = (z_{21}, \dots, z_{2p})$ são determinados pelas

seguintes equações (4) e (5):

$$z_1 = \delta \log[y + (y^2 + 1)^{1/2}], \quad (4)$$

em que:

$$\delta = \frac{1}{[\log(\sqrt{\omega^2})]^{1/2}} \quad \text{e} \quad y = \sqrt{b_1} \left[\frac{\omega - 1}{2} \frac{(n+1)(n+3)}{6(n-2)} \right]^{1/2},$$

$$\omega^2 = -1 + [2(\beta - 1)]^{1/2} \quad \text{e} \quad \beta = \frac{3(n^2 + 27n - 70)(n+1)(n+3)}{(n-2)(n+5)(n+7)(n+9)},$$

$$z_2 = \left[\left(\frac{\chi}{2\alpha} \right)^{1/3} - 1 + \frac{1}{9\alpha} \right] (9\alpha)^{1/2}, \quad (5)$$

em que:

$$\alpha = a + b_1 c,$$

$$\chi = (b_2 - 1 - b_1) 2\kappa,$$

$$a = \frac{(n-2)(n+5)(n+7)(n^2 + 27n - 70)}{6(n-3)(n+1)(n^2 + 15n - 4)},$$

$$c = \frac{(n-7)(n+5)(n+7)(n^2 + 27n - 5)}{6(n-3)(n+1)(n^2 + 15n - 4)}$$

$$\text{e} \quad \kappa = \frac{(n+5)(n+7)(n^3 + 37n^2 + 11n - 313)}{12(n-3)(n+1)(n^2 + 15n - 4)}.$$

Segundo Biase (2011), na avaliação feita por Doornik e Hansen (1994), para o caso bivariado e tamanho amostral 50, o poder desse teste foi semelhante aquele do teste conjunto de simetria e curtose de Mardia (1970). Pelo fato da desprovida avaliação feita por Doornik e Hansen (1994) para avaliar o poder, não se sabe de fato se este teste é apropriado para testar a normalidade multivariada.

2.3.3 Teste de Henze-Zirkler de normalidade multivariada

Segundo Korkmaz, Goksuluk e Zararsiz (2015), o teste de Henze-Zirkler baseia-se numa distância funcional não negativa, que mede a distância entre as duas funções de distribuição. Se os dados são distribuídos como uma normal multivariada, a estatística de teste é aproximadamente distribuída como uma log-normalmente. Primeiro, são calculados a média, a variância e o parâmetro de suavidade. Em seguida, a média e a variância são log-normalizadas e o valor $-p$ é estimado. A estatística de teste é dada por:

$$HZ = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n e^{-\frac{\beta^2}{2} D_{ij}} - 2(1 + \beta^2)^{-\frac{p}{2}} \sum_{i=1}^n e^{-\frac{\beta^2}{2(1+\beta^2)} D_i} + n(1 + 2\beta^2)^{-\frac{p}{2}},$$

em que:

$$\beta = \frac{1}{\sqrt{2}} \left[\frac{n(2p+1)}{4} \right]^{\frac{1}{p+4}},$$

$$D_{ij} = (x_i - x_j)^\top S^{-1} (x_i - x_j),$$

$$D_i = (x_i - \bar{x})^\top S^{-1} (x_i - \bar{x}) = m_{ii}$$

e p é o número de variáveis.

Se os dados seguem uma distribuição normal multivariada, a estatística do teste (HZ) é aproximadamente log-normalmente distribuída com média μ e variância σ^2 definidas por:

$$\mu = 1 - \frac{a^{-\frac{p}{2}} [1 + p\beta^{\frac{2}{a}} + p(p+2)\beta^4]}{2a^2} e$$

$$\sigma^2 = 2(1+4\beta^2)^{-\frac{p}{2}} + \frac{2a^{-p}(1+2p\beta^4)}{a^2} + \frac{3p(p+2)\beta^8}{4a^4} - 4\omega_\beta^{-\frac{p}{2}} \left[1 + \frac{3p\beta^4}{2\omega_\beta} + \frac{p(p+2)\beta^8}{2\omega_\beta^2} \right],$$

com $a = 1 + 2\beta^2$ e $\omega_\beta = (1 + \beta^2)(1 + 3\beta^2)$. Assim, a média e a variância da

estatística HZ log-normalizada podem ser calculados por:

$$\log(\mu) = \log\left(\sqrt{\frac{\mu^4}{\sigma^2 + \mu^2}}\right) \quad \text{e} \quad \log(\sigma^2) = \log\left(\frac{\sigma^2 + \mu^2}{\sigma^2}\right).$$

Usando os parâmetros da distribuição log-normal, μ e σ , testa-se a normalidade multivariada dos dados. A estatística de teste de Wald para normalidade multivariada é dada por:

$$z = \frac{\log(HZ) - \log(\mu)}{\log(\sigma)}.$$

Além dos testes de normalidade multivariados citados anteriormente, outros procedimentos, com o mesmo objetivo de verificar a normalidade multivariada dos dados, foram baseados no Teorema 1, apresentado por Ferreira (2011).

Teorema 1 (Elipsóide de concentração). *Se o vetor aleatório $\mathbf{X} \in \mathbb{R}^p$ segue uma distribuição normal multivariada com densidade $f_{\mathbf{X}}(\mathbf{x})$, dada por:*

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

com média $\boldsymbol{\mu}$ e matriz de covariância $\boldsymbol{\Sigma}$, então

$$(\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu})$$

tem distribuição qui-quadrado com p graus de liberdade (χ_p^2) e a região

$$(\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{X} - \boldsymbol{\mu}) \leq \chi_{\alpha,p}^2$$

define o elipsóide de concentração $100(1 - \alpha)\%$ das realizações das variáveis aleatórias, em que $\chi_{\alpha,p}^2$ é o quantil superior $100\alpha\%$ da distribuição de qui-quadrado com $\nu = p$ graus de liberdade, obtido de acordo com a seguinte afirmativa probabilística $P(\chi_p^2 > \chi_{\alpha,p}^2) = \alpha$.

A prova desse teorema pode ser encontrada em Ferreira (2011). Como $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$ são desconhecidos, eles são substituídos pelos estimadores $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}}$ e

$\hat{\Sigma} = (n - 1)\mathbf{S}/n$. Vale ressaltar que nesse caso o Teorema 1 é plausível apenas assintoticamente.

Outra forma de verificar a normalidade multivariada dos dados é por meio do gráfico de quantil-quantil, também conhecido na literatura por *Q-Q plot*. Para a construção desse procedimento visual é necessário:

i) calcular cada distância de Mahalanobis D_j^2 :

$$D_j^2 = (\mathbf{X}_j - \bar{\mathbf{X}}.)^\top \mathbf{S}^{-1}(\mathbf{X}_j - \bar{\mathbf{X}}.);$$

ii) ordená-las (estatísticas de ordem):

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

iii) para cada j , obter:

$$\chi_j^2 = F_{\chi_\nu^2}^{-1} \left[\frac{j - \frac{1}{2}}{n} \right], \text{ com } j = 1, 2, \dots, n,$$

em que $F_{\chi_\nu^2}^{-1}$ é a função inversa da função de distribuição qui-quadrado com $\nu = p$ graus de liberdade;

iv) plotar as coordenadas:

$$(D_{(j)}^2, \chi_j^2).$$

O problema dos *Q-Q plots* e de qualquer teste que se fundamenta em distâncias quadráticas é que esses D_j^2 não são variáveis aleatórias independentes e identicamente distribuídas (iid), ou seja, estes dados não são independentes, mesmo se os dados originais \mathbf{X}_j s sejam independentes. Além disso, a distribuição marginal de D_j^2 sob a hipótese nula não é precisamente qui-quadrado para valores pequenos de n . Isso compromete os *Q-Q plots* ou qualquer teste que tenha o mesmo princípio básico.

Portanto, o presente trabalho apresenta uma outra forma para a construção do teste de normalidade multivariada, que não foi baseada em distâncias quadráticas.

cas, e sim, em uma transformação t de *Student*.

2.4 Gráfico de Quantil-Quantil para dados Normais Multivariados usando transformação t de Student

Sejam os vetores aleatórios contínuos $\mathbf{X}_1, \dots, \mathbf{X}_n$ com dimensão $(p \times 1)$ independentes e identicamente distribuídos (iid) de acordo com a função densidade de probabilidade (fdp) normal multivariada $f_{\mathbf{X}}(\mathbf{x})$, com $\mathbf{X} \in \mathbb{R}^p$ (no espaço euclidiano p -dimensional). Escrevendo de uma forma simplificada tem-se que $\mathbf{X}_1, \dots, \mathbf{X}_n \stackrel{iid}{\sim} f_{\mathbf{X}}(\mathbf{x})$, em que a fdp $f_{\mathbf{X}}(\mathbf{x})$, dado por (3), tem média zero e momentos finitos de segunda ordem. Obtêm-se:

$$\mathbf{S}_k = \sum_{i=1}^k \mathbf{X}_i \mathbf{X}_i^{\top}, \mathbf{Y}_k = \mathbf{S}_k^{-1/2} \mathbf{X}_k, \mathbf{Z}_k = \frac{\mathbf{Y}_k}{\sqrt{1 - \mathbf{Y}_k \mathbf{Y}_k^{\top}}}, k = p+1, \dots, n \quad (6)$$

em que $\mathbf{S}_k^{-1/2} = (\mathbf{S}_k^{1/2})^{-1}$ e $\mathbf{S}_k^{1/2}$ representa a raiz quadrada positiva definida de \mathbf{S}_k . Yang *et al.* (1996) deram a seguinte caracterização para a distribuição normal multivariada $N_p(\mathbf{0}, \Sigma)$.

Teorema 2. *Sejam $\mathbf{X}_1, \dots, \mathbf{X}_n \stackrel{iid}{\sim} N_p(\mathbf{0}, \Sigma)$. Definindo-se os vetores aleatórios \mathbf{Y}_k e \mathbf{Z}_k por (6), então,*

1. $\mathbf{Y}_{p+1}, \dots, \mathbf{Y}_n$ são independentes entre si e \mathbf{Y}_k ($k \geq p+1$) tem uma distribuição simétrica multivariada de Pearson tipo II com a seguinte fdp

$$f_k(\mathbf{y}) = \frac{\Gamma(k/2)}{\pi^{p/2} \Gamma(k-p)/2} (1 - \mathbf{Y}^{\top} \mathbf{Y})^{(k-p-2)/2}, \mathbf{Y} \in \mathbb{R}^p, \mathbf{Y}^{\top} \mathbf{Y} < 1; \quad (7)$$

2. $\mathbf{Z}_{p+1}, \dots, \mathbf{Z}_n$ são independentes entre si e \mathbf{Z}_k ($k \geq p+1$) tem uma distribuição simétrica multivariada de Pearson tipo VII com a seguinte fdp

$$h_k(\mathbf{z}) = \frac{\Gamma(k/2)}{\pi^{p/2} \Gamma(k-p)/2} (1 - \mathbf{Z}^{\top} \mathbf{Z})^{-k/2}, \mathbf{Z} \in \mathbb{R}^p;$$

3. Sejam $\mathbf{X}_1, \dots, \mathbf{X}_n$ (iid) em \mathbb{R}^p com uma fdp $v(\mathbf{x})$ que é contínua em $v(\mathbf{0}) > 0$ e os vetores aleatórios \mathbf{Z}_k definidos em (6). Se \mathbf{Z}_k e \mathbf{Z}_{k-1} ($k \geq$

$p + 2$) tem fdps $h_k(z)$ e $h_{k-1}(z)$ definidas por (2), respectivamente, então \mathbf{X}_i ($i = 1, \dots, n$) tem uma distribuição normal multivariada.

A partir do Teorema 2 proposto por Yang *et al.* (1996), pode-se deduzir o seguinte corolário.

Corolário 1. *Suponha que $\mathbf{X}_1, \dots, \mathbf{X}_n \stackrel{iid}{\sim} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Definindo-se os vetores aleatórios*

$$\mathbf{U}_i = \frac{\mathbf{X}_1 + \dots + \mathbf{X}_i - i \mathbf{X}_{i+1}}{\sqrt{i(i+1)}}, \quad i = 1, 2, \dots, n-1$$

e

$$\mathbf{S}_k = \sum_{i=1}^k \mathbf{U}_i \mathbf{U}_i^\top, \quad \mathbf{Y}_k = \mathbf{S}_k^{-1/2} \mathbf{U}_k, \quad k = p+1, p+2, \dots, n-1,$$

em que $\mathbf{U}_i \stackrel{iid}{\sim} N_p(\mathbf{0}, \boldsymbol{\Sigma})$, $\mathbf{Y}_k = (y_{k1}, \dots, y_{kp})^\top$ tem dimensão $p \times 1$ e

$$\bar{Y}_k = \frac{1}{p} \sum_{i=1}^p y_{ki}, \quad S_k^2 = \frac{1}{p-1} \sum_{i=1}^p (y_{ki} - \bar{Y}_k)^2,$$

$$T_k = \frac{\bar{Y}_k \sqrt{p}}{S_k}, \quad (8)$$

então, $T_k \stackrel{iid}{\sim} t_\nu$, em que $\nu = p-1$ e $k = p+1, p+2, \dots, n-1$, ou seja, T_k tem distribuição t de Student com $p-1$ graus de liberdade (gl).

O $Q-Q$ plot é baseado no corolário 1 e nas seguintes relações:

$$\{\mathbf{X}_i\}_{i=1}^n \Rightarrow \{\mathbf{U}_i\}_{i=1}^{n-1} \stackrel{iid}{\sim} N_p(\mathbf{0}, \boldsymbol{\Sigma}) \Rightarrow \{T_k(\mathbf{y}_k)\}_{k=p+1}^{n-1} \stackrel{iid}{\sim} t_{\nu=p-1}.$$

Em seguida é definida a estatística de ordem da estatística t de T_k s:

$$t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(n-p-1)}.$$

Seja $F^{-1}(\cdot; p-1)$ a inversa da função de distribuição de t_ν com $\nu = p-1$,

e

$$t_{q_j} = F^{-1}\left(\frac{2j-1}{2(n-p-1)}; p-1\right),$$

para $j = 1, 2, \dots, n - p - 1$. Estes são os quantis das distribuições $t_{\nu=p-1}$. De acordo com o princípio do *Q-Q plot*, para se testar a normalidade multivariada dos dados, deve-se *plotar* os pontos $\{(t_{q_j}, t_{(j)}) : j = 1, \dots, n - p - 1\}$, que é chamado de *t - plot*.

Haverá evidências de não normalidade multivariada, se o *t-plot* desviar-se visivelmente do equi-angular da reta $y = x$.

2.5 Programa R

Existe um conjunto de ferramentas computacionais que podem ser usadas para explorar, avaliar e compreender as técnicas estatísticas, bem como a parte inferencial. Nesse estudo, foi utilizado o programa *R* versão 3.1.0.

O programa *R* começou a ser desenvolvido por Robert Gentleman e Ross Ihaka do Departamento de Estatística da Universidade de Auckland, em Nova Zelândia, mais conhecidos por “R & R”, apelido do qual originou-se o nome *R* do programa. O objetivo inicial de “R & R”, em 1991, era produzir um programa para as suas aulas de laboratório baseado na linguagem *S*, utilizada pelo programa comercial *S-Plus* criado por Jonh M. Chambers da AT&T, que atualmente vem contribuindo para o aperfeiçoamento e ampliação das análises estatísticas do *R*. O primeiro relato da distribuição do *R* foi em 1993, quando algumas cópias foram disponibilizadas no *StatLib*, um sistema de distribuição de programas estatísticos.

Em 1997 foi formado um grupo de profissionais que têm acesso ao código fonte do *R*, possibilitando, assim, atualizações mais rápidas do programa. Desde então o *R* vem ganhando cada vez mais adeptos em todo o mundo, em parte devido ao fato de ser totalmente gratuito e também por ser um programa que necessita do usuário o conhecimento das análises que estão sendo feitas, diminuindo assim as chances de uma interpretação errada dos resultados (R CORE TEAM, 2015).

Vale ressaltar que a estimação de parâmetros de um modelo pode ser feita utilizando o programa, por meio de uma biblioteca ou com linguagem de programação.

2.5.1 Teste de Shapiro-Wilk de normalidade multivariada do pacote ‘mvnormtest’

O teste proposto por Royston (1983), que é uma generalização do teste de Shapiro-Wilk (1965) para o caso multivariado, entre os testes baseados na técnica da falta de ajuste, apresentou bom desempenho de poder, Biase (2011). O pacote ‘mvnormtest’ possui uma função que calcula a estatística do teste e o valor- p do teste de Shapiro-Wilk para normalidade multivariada. Porém, uma desvantagem deste teste é a respeito da limitação quanto ao tamanho da amostra ($3 < n < 5000$).

Pacote:

```
mvnormtest
```

Comando:

```
mshapiro.test(U)
```

em que:

U é uma matriz numérica ($p \times n$) com os valores dos dados multivariados, cujos números devem ser, para cada amostra entre, 3 e 5000.

Descrição do comando:

Executa o teste de Shapiro-Wilk para normalidade multivariada.

2.5.2 Teste de Doornik-Hansen de normalidade multivariada do pacote ‘asbio’

O teste Doornik-Hansen de normalidade multivariada é baseado na assimetria e curtose dos dados multivariados. O teste de Doornik e Hansen (1994) possui bom desempenho de poder, principalmente nos casos em que as amostras são maiores ou iguais a 75. O pacote ‘asbio’ possui uma função que calcula a estatística do teste, os graus de liberdade e os valores- p relacionados ao caso de normalidade multivariada (multi), e outro relacionado aos casos univariados (univ) de cada variável p da matriz de dados Y .

Pacote:

```
asbio
```

Comando:

```
DH.test(Y, Y.names = NULL)
```

em que:

Y é uma matriz numérica ($n \times p$) com os valores dos dados multivariados;

$Y.names$ é um vetor de nomes das variáveis Y .

Descrição do comando:

Executa o teste de Doornick-Hansen para normalidade multivariada que é uma poderosa alternativa ao teste de Shapiro-Wilk.

2.5.3 Teste de Henze-Zirkler de normalidade multivariada do pacote ‘MVN’

O teste proposto por Henze e Zirkler (1990) para a distribuição t -student multivariada apresentou alto poder. Esse teste classifica-se entre os testes consistentes e invariantes. O pacote ‘MVN’ existe uma função que calcula as estatísticas do teste e o valor- p .

Pacote:

```
MVN
```

Comando:

```
hzTest(data, cov = TRUE, qqplot = FALSE)
```

em que:

`data` é uma matriz numérica ($n \times p$) com os valores dos dados multivariados;
`cov = TRUE` significa que a matriz de covariâncias dos dados é normalizada por n e se colocar
`cov = FALSE` é normalizada por $n - 1$;
`qqplot` gera um $Q-Q$ plot quando for especificado TRUE.

Descrição do comando:

Essa função avalia a normalidade multivariada pelo teste de Henze-Zirkler.

2.6 Simulação Monte Carlo

A simulação é um processo que tenta reproduzir, por meio de programas de computadores, o comportamento de um sistema real, com a finalidade de estudar seu funcionamento sob condições alternativas (DACHS, 1988). O uso corrente

da simulação é encontrado, por exemplo, em esculturas e pinturas da Renascença, escala de aviões supersônicos, modelos analíticos de processos mentais, na área médica, sistemas industriais, na economia e em outras aplicações.

O uso de simulações tornou-se importante, pois é um processo que serve para projetar um modelo computacional de um sistema real e conduzir experimentos com este modelo com o propósito de entender seu comportamento e/ou avaliar estratégias para sua operação. Seu uso ainda é extremamente ativo nos dias atuais, principalmente com o uso de recursos computacionais, e muita pesquisa tem sido desenvolvida por estatísticos. Assim sendo, Silva (2009) relatam que o aparecimento do computador permitiu a um número crescente de pessoas a possibilidade de realizar simulações, desde simples jogos, até modelos na economia em escala mundial.

Por diversas épocas o homem tem usado processamentos de realizações de variáveis aleatórias, mas a teoria de probabilidade só se formalizou no século XVII, devido ao desenvolvimento matemático que surgiu de pesquisas em jogos de azar. No entanto, o nome Monte Carlo está relacionado com a cidade de mesmo nome, localizada no Principado de Mônaco, onde ocorriam jogos de azar.

Além disso, é importante enfatizar que outro grande impulso para o desenvolvimento de simulações foi durante o século XX, na década de 40, durante trabalhos bélicos sobre a bomba atômica. Os cientistas não conseguiam, por meio de análise, resolver alguns problemas matemáticos não-probabilísticos complexos. A solução desse problema só foi possível de ser resolvidos pela combinação de processamentos ao acaso com a teoria de probabilidade.

O método de Monte Carlo tem seus princípios na década de 1940, quando Ferni, Ulam, von Neumann, Metropolis e outros começaram a considerar o uso de números aleatórios para analisar diferentes problemas em Física a partir de uma perspectiva estocástica (LANDAU; BINDER, 2009).

Silva (2009) mencionaram de maneira simplificada que o método de Monte Carlo é usado para que sejam simuladas realizações de variáveis aleatórias a partir de uma sequência pseudoaleatória, baseada na distribuição uniforme (0,1). Além do mais, todo processo simulado que envolve um componente aleatório de qualquer distribuição é considerado como pertencente ao método Monte Carlo.

Ferreira (2010) determina que a simulação é uma técnica que usamos para a solução de problemas. Se a solução alcançada for mais rápida, de menor custo e de fácil interpretação em relação a outro método qualquer, o uso de simulação é justificável.

Landau e Binder (2009) afirmam que com o aumento das potências dos computadores que estamos vivenciando agora, juntamente com a queda constante no preço, as simulações de computador serão capazes de rapidamente aumentar em sofisticação para permitir que comparações mais perspicazes sejam feitas.

Uma grande vantagem em utilizar a simulação computacional é por ser um procedimento de custo baixo, evita prejuízos por não utilizar procedimentos inadequados e otimiza a decisão e o funcionamento do sistema real, sob a visão de Ferreira (2010).

Métodos simples de Monte Carlo foram criados para fornecer meios para estimar respostas para os problemas analiticamente intratáveis. Muitas dessas técnicas simples mantem a sua importância por causa do aumento de acessos e conseqüente poder dos computadores que levaram a aplicação de métodos simples para um amplo espectro de problemas interessantes, descrevem Landau e Binder (2009).

Segundo Ferreira (2010) os modelos multivariados ganharam grande aceitação no meio científico em função das facilidades computacionais e do desenvolvimento de programas especializados na área.

Os testes de normalidade, tanto para os casos univariados como para os casos multivariados, utilizam o método Monte Carlo com o propósito de validar testes quanto ao desempenho do poder e controle da taxa de erro tipo I.

3 MÉTODO

3.1 Proposta do teste de Normalidade Multivariada

O novo teste proposto é baseado na transformação de Liang *et al.* (2004) para a distribuição t de *Student*.

Para o novo teste proposto, que será chamado de teste de normalidade multivariada t (TMNt), as hipóteses são:

$$\begin{cases} H_0 : \text{Os dados amostrais seguem uma dist. normal multivariada } f_{\mathbf{X}}(\mathbf{x}), \\ H_1 : \text{Os dados amostrais não seguem uma dist. normal multivariada } f_{\mathbf{X}}(\mathbf{x}), \end{cases}$$

em que:

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

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

Para determinar a estatística do teste, foi considerada uma amostra p -variada $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_j, \dots, \mathbf{X}_n$ de uma distribuição qualquer com vetor de médias $\boldsymbol{\mu}$ com dimensão $(p \times 1)$ e matriz de covariâncias Σ positiva definida de dimensão $(p \times p)$. Sob a hipótese de normalidade multivariada de \mathbf{X}_j , $j = 1, 2, \dots, n$, foi realizada a seguinte transformação:

$$\mathbf{U}_i = \frac{\mathbf{X}_1 + \dots + \mathbf{X}_i - i \mathbf{X}_{i+1}}{\sqrt{i(i+1)}}, \quad i = 1, 2, \dots, n-1.$$

Posteriormente, foram computados

$$\mathbf{S}_k = \sum_{i=1}^k \mathbf{U}_i \mathbf{U}_i^\top \quad \text{e} \quad \mathbf{Y}_k = \mathbf{S}_k^{-1/2} \mathbf{U}_k, \quad \text{com} \quad k = p+1, \dots, n-1.$$

Finalmente, obteve-se

$$\bar{Y}_k = \frac{1}{p} \sum_{i=1}^p Y_{ki}, \quad S_k^2 = \frac{1}{p-1} \sum_{i=1}^p (Y_{ki} - \bar{Y}_k)^2 \quad \text{e} \quad T_k = \frac{\bar{Y}_k \sqrt{p}}{S_k},$$

para $k = p + 1, p + 2, \dots, n - 1$.

As variáveis aleatórias T_k são $\overset{iid}{\sim} t_\nu$ sendo $\nu = p - 1$ graus de liberdade e $k = p + 1, \dots, n - 1$ (LIANG *et al.*, 2004). Sendo assim, foi computado como estatística do teste de normalidade a estatística de Kolmogorov-Smirnov:

$$D^+ = \max[F_{T_k}(t_k) - F_T(t)], \quad (9)$$

sendo T_k a variável aleatória obtida em (8), T uma variável aleatória da distribuição t de *Student* com $\nu = p - 1$ graus de liberdade, $F_{T_k}(t_k)$ e $F_T(t)$ as funções de distribuição empírica amostral e t de *Student*. O valor $-p$, sob H_0 , é obtido conforme descrito em Marsaglia, Tsang, e Wang (2003). A não rejeição da hipótese de que $T_k \sim t_{p-1}$ equivale a não rejeição da hipótese de que $\mathbf{X}_j \sim N_p(\mu, \Sigma)$.

Para a construção do t -plot foi plotado os pontos $(tq_j, t_{(j)})$, em que $t_{(j)}$ é a estatística de ordem da estatística t de T_k s, dada por:

$$t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(n-p-1)}$$

e tq_j é igual a inversa da função de distribuição t_ν com $\nu = p - 1$ aplicada em um ponto, dada por:

$$tq_j = F^{-1} \left(\frac{2j - 1}{2(n - p - 1)}; p - 1 \right),$$

para $j = 1, 2, \dots, n - p - 1$.

3.2 Validação do desempenho do teste

Foi usada simulação Monte Carlo em dois cenários para avaliar o desempenho do TNMt, mensurando-se as taxas de erro tipo I e o poder do teste.

Inicialmente, foram simuladas amostras normais p -variadas considerando diferentes números de variáveis p e tamanhos amostrais n . Para cada configuração foram gerados $N = 2000$ amostras. A proporção de rejeições nas N simulações de cada configuração correspondeu à taxa de erro tipo I, e os níveis de significância nominais foram fixados em $\alpha = 10\%$, $\alpha = 5\%$ e $\alpha = 1\%$.

Posteriormente, foram simuladas amostras de outras distribuições p variadas não-normais: a distribuição *t-Student* multivariada com $\nu = 1$ gl e $\nu = 30$ gl, a distribuição log-normal multivariada, a distribuição normal contaminada multivariada e a distribuição uniforme multivariada. As mesma configurações de p e n anteriores foram utilizadas. A proporção de rejeições em cada configuração foi computada em $N = 2000$ simulações Monte Carlo e correspondeu à estimativa empírica do poder do teste.

Para fins de comparação, foram usados os resultados de Biase (2011), que aplicou o teste de Shapiro-Wilk de normalidade multivariada (TSWNM) de Royston (1983), o teste Monte Carlo de normalidade multivariada baseado em distâncias (TMCNMD) e o teste de normalidade multivariada baseado em *bootstrap* paramétrico (TNMBP) propostos por ela.

Foram considerados os valores na combinações de n e p para a dimensão da matriz de dados $\mathbf{X}_{n \times p}$, de acordo com a Tabela 2.

Tabela 2 Os testes de normalidade multivariados usados nas comparações, com os tamanhos amostrais n e número de variáveis p usados nas construções das matrizes de dados $\mathbf{X}_{n \times p}$, considerando os teste usados no trabalho.

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

O teste estatístico será considerado exato, quando o nível de significância real (estimado por simulação) for igual ao nível de significância nominal (fixado previamente pelo pesquisador), conservador, quando o nível de significância real for inferior ao nível de significância nominal, ou liberal, quando o nível de significância real for superior ao nível de significância nominal. Os procedimentos para essa classificação são os mesmos adotados por Basto (2013).

4 RESULTADOS E DISCUSSÃO

Nessa seção, os resultados das simulações são apresentados em duas etapas: uma referente ao erro tipo I e outra ao poder do teste. Além disso, foi utilizado um exemplo com o propósito de ilustrar alguns testes existentes de normalidade multivariada e o teste proposto.

A taxa de erro tipo I e o poder do teste proposto (TNMt) foram computados e comparados com três testes de normalidade multivariada: o teste de Shapiro-Wilk (TSWNM), proposto por Royston (1983); o teste Monte Carlo baseado em distâncias (TMCNMD), proposto por Biase (2011); o teste baseado em *bootstrap* paramétrico (TNMBP), também proposto por Biase (2011). Os níveis de significância nominais considerados para o erro tipo I e o poder do teste foram $\alpha = 10\%$, $\alpha = 5\%$ e $\alpha = 1\%$.

Vale ressaltar que para o poder dos testes foram consideradas cinco distribuições não-normais multivariadas: a *t-Student* (com diferentes graus de liberdades, $\nu = 1$ e $\nu = 30$), log-normal, normal contaminada e uniforme.

4.1 Erro tipo I

As probabilidades de se cometer os erros do tipo I e II são inversamente proporcionais. Assim, um teste adequado deve apresentar um controle adequado das taxas de erro tipo I e um alto poder relativamente aos seus concorrentes. Se um teste apresenta taxas de erro tipo I muito inferiores aos níveis nominais de significância adotados, então ele tende a apresentar valores de poder relativamente muito baixos.

As Tabelas 3, 4 e 5, referem-se ao erro tipo I dos testes TMCNMD, TNMBP, TSWNM e TNMt, considerando os níveis de significância nominais de 10%, 5% e 1%, respectivamente. Foram consideradas as combinações de n e p , previamente apresentadas na Tabela 2.

Na Tabela 3 apresentam-se as taxas de erro tipo I referente aos testes TMCNMD, TNMBP, TSWNM e TNMt, considerando a distribuição normal multivariada em função das combinações de n e p , para $\alpha = 10\%$. O TMCNMD apre-

sentou, na maior parte dos resultados, os valores dos níveis de significância reais próximos do nível de significância nominal de $\alpha = 10\%$. Em nenhuma situação ele foi considerado liberal ou conservativo. Para checar isso, aplicou-se um teste binomial exato no nível nominal de significância de 1%. Nesse caso, valores estimados da proporção de rejeição de H_0 , inferiores a 8,34% e superiores a 11,85%, são considerados significativamente diferentes ($p < 0,01$) do nível nominal de 10%. Da mesma forma, para os níveis nominais de significância de 5% e 1%, os limites para as rejeições são 3,83% e 6,39% e 0,52% e 1,73%, respectivamente.

O TNMBP, na maioria das vezes, foi considerado exato, mas nas situações $n = 10$ com $p = 4$ e 6, $n = 20$ com $p = 18$, o teste foi conservativo. Observa-se no TSWNM, que para valores de p grande, as taxas de erro tipo I foram consideradas abaixo do nível de significância nominal (conservativo). Isso ocorreu nos casos em que os valores de p se aproximam dos valores de n , ou seja quando, $n = 20$ e $p = 18$, $n = 100$ e $p = 50$ e 98, $n = 200$ com $p = 50$ e 100.

O TNMt apresentou em todos os seus resultados os valores dos níveis de significância real próximos do nível de significância nominal fixado, ou seja, o teste pode ser classificado como exato. Vale ressaltar que os valores das taxas de erro tipo I estimadas para $n = 10$ e $p = 8$, $n = 20$ e $p = 18$, $n = 50$ e $p = 48$, $n = 100$ e $p = 98$, não puderam ser computadas, em virtude de o tamanho da amostra efetiva ser igual $n - p - 1 = 1$, não havendo assim como aplicar o teste proposto.

O TSWNM, assim como o TNMBP, na maioria das vezes, foi considerado exato e em outras ocasiões foi considerado conservativo. O TMCNMD e o TNMt para estas configurações (Tabela 3) obtiveram o controle do erro tipo I melhor do que os outros dois testes.

Em comparação a outros testes de normalidade multivariada existentes na literatura, buscou-se fazer um paralelo em algumas circunstâncias apropriadas da Tabela 3 com os testes comparados por Cirillo e Ferreira (2003). Notou-se, por exemplo, que o teste de normalidade baseado no coeficiente de correlação Quantil-Quantil multivariado para amostras de tamanho $n = 100$ com $p = 2$ e $p = 10$ tiveram controle adequado das taxas de erro tipo I, dados respectivamente por 0,1038 e 0,0965, o mesmo ocorre para os quatro testes comparados na Tabela 3.

Tabela 3 Erro tipo I dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição normal multivariada em função de n e p para $\alpha = 10\%$.

n	p	TMCNMD	TNMBP	TSWNM	TNMt
10	2	0,1000	0,1065	0,1010	0,0975
	4	0,1030	0,0785	0,0850	0,1075
	6	0,0980	0,0765	0,0895	0,1020
	8	0,1175	0,0915	0,0950	-
20	2	0,0965	0,1010	0,1045	0,0925
	5	0,0975	0,0900	0,0935	0,1010
	10	0,1080	0,0890	0,0875	0,1050
	18	0,1000	0,0770	0,0665	-
50	2	0,1020	0,0960	0,0905	0,1080
	10	0,1050	0,0905	0,0830	0,0995
	25	0,0975	0,0955	0,0940	0,0960
	48	0,0875	0,0960	0,0865	-
100	2	0,0915	0,1035	0,0980	0,0930
	10	0,1010	0,0880	0,0845	0,1070
	50	0,1160	0,0880	0,0725	0,1030
	98	0,1075	0,0975	0,0765	-
200	2	0,0895	0,0840	0,0870	0,0970
	10	0,1035	0,1030	0,0885	0,0875
	50	0,0900	0,0885	0,0760	0,0980
	100	0,0915	0,0935	0,0645	0,1085
10000	5	0,0991	0,1032	-	0,1010

Na Tabela 4 são apresentadas as taxas de erro tipo I referente aos testes TMCNMD, TNMBP, TSWNM e TNMt, considerando a distribuição normal mul-

tivariada em função das combinações de n e p , para $\alpha = 5\%$. O TMCNMD controlou a taxa de erro tipo I em todas as situações, independente do tamanho da amostra e do número de variáveis, ou seja, este teste foi considerado exato em todas as situações.

Para o TNMBP, na maioria das situações, o teste foi considerado exato, exceto para $n = 10$ com $p = 6$ e $n = 20$ com $p = 18$, em que o teste foi conservativo. O TSWNM foi considerado conservativo nas situações em que $n = 20$ com $p = 10$, $n = 20$ com $p = 18$ e $n = 100$ com $p = 50$, apresentando exato para os restantes das combinações de n e p .

O TNMt foi conceituado como sendo um teste exato, tal como o TMCNMD, por apresentar o controle da taxa de erro tipo I em todas as configurações de n e p .

Novamente procurou-se comparar algumas das taxas de erro tipo I dos testes da Tabela 4 com outros testes já avaliados na literatura. Os testes avaliados por Farrell, Barrera e Naczki (2007) para 10000 repetições Monte Carlo, como o teste de Doornik e Hansen (1994) e o teste de Henze e Zilkler (1990), para amostras de tamanho $n = 50$ com $p = 2$, $n = 50$ com $p = 10$ e $n = 200$ com $p = 2$ foram exatos, da mesma forma que os quatro testes avaliados na Tabela 4. Os testes foram também exatos para amostra de tamanho $n = 200$ com $p = 10$, evento que não ocorre para o teste de Shapiro-Francia multivariado proposto por Silva (2009), que foi considerado liberal.

Tabela 4 Erro tipo I dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição normal multivariada em função de n e p para $\alpha = 5\%$.

n	p	TMCNMD	TNMBP	TSWNM	TNMt
10	2	0,0450	0,0525	0,0535	0,0530
	4	0,0525	0,0430	0,0495	0,0595
	6	0,0515	0,0350	0,0440	0,0520
	8	0,0510	0,0410	0,0450	-

Tabela 4, Continua

20	2	0,0555	0,0530	0,0460	0,0450
	5	0,0510	0,0430	0,0405	0,0525
	10	0,0550	0,0390	0,0355	0,0490
	18	0,0505	0,0285	0,0340	-
50	2	0,0550	0,0420	0,0465	0,0550
	10	0,0505	0,0390	0,0405	0,0470
	25	0,0455	0,0485	0,0475	0,0485
	48	0,0430	0,0395	0,0420	-
100	2	0,0435	0,0440	0,0460	0,0465
	10	0,0525	0,0455	0,0470	0,0560
	50	0,0530	0,0380	0,0350	0,0525
	98	0,0475	0,0440	0,0435	-
200	2	0,0455	0,0430	0,0440	0,0475
	10	0,0495	0,0450	0,0455	0,0470
	50	0,0455	0,0400	0,0410	0,0480
	100	0,0455	0,0440	0,0385	0,0560
10000	5	0,0418	0,0477	-	0,0530

No último caso, foi considerando a distribuição normal multivariada em função das combinações de n e p , para $\alpha = 1\%$. As taxas de erro tipo I para os testes de normalidade multivariadas TMCNMD, TNMBP, TSWNM e TNMt encontram-se na Tabela 5. Novamente, o TMCNMD obteve controle adequado da taxa de erro tipo I em todas as situações, independente do tamanho da amostra, n , e do número de variáveis, p , ou seja, o teste foi considerado exato.

O TNMBP, quando avaliado em relação as taxas de erro tipo I observadas, foi considerado conservador, para os casos em que a amostra foi pequena, como em $n = 10$ com valores de p iguais a 6 e 8 e em $n = 20$ com $p = 18$. Para as demais situações, o teste foi exato. O TSWNM apresentou-se exato, com exceção

dos casos em que $p = 6$ com $n = 10$, sendo o teste conservativo e em que p foi igual a 98, com n igual a 100, em que esse teste foi considerado liberal. O TNMt, em sua totalidade dos casos, foi considerado um teste exato.

Para amostras $n = 100$ com $p = 2$ e $n = 100$ com $p = 10$, os testes de normalidade multivariada avaliados por Cirillo e Ferreira (2003), o teste de normalidade multivariada de simetria e o teste de normalidade multivariada de curtose, como também, os testes referentes à Tabela 5 e o teste de Shapiro-Francia multivariado proposto por Silva (2009) foram todos exatos, havendo um excelente controle das taxas de erro tipo I.

Tabela 5 Erro tipo I dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição normal multivariada em função de n e p para $\alpha = 1\%$.

n	p	TMCNMD	TNMBP	TSWNM	TNMt
10	2	0,0070	0,0110	0,0125	0,0090
	4	0,0085	0,0065	0,0110	0,0170
	6	0,0100	0,0035	0,0050	0,0100
	8	0,0110	0,0035	0,0055	-
20	2	0,0115	0,0115	0,0110	0,0085
	5	0,0085	0,0095	0,0080	0,0115
	10	0,0085	0,0060	0,0070	0,0065
	18	0,0085	0,0050	0,0080	-
50	2	0,0105	0,0095	0,0110	0,0110
	10	0,0080	0,0065	0,0095	0,0130
	25	0,0085	0,0070	0,0115	0,0105
	48	0,0090	0,0110	0,0165	-
100	2	0,0080	0,0105	0,0100	0,0055
	10	0,0095	0,0085	0,0090	0,0115
	50	0,0135	0,0060	0,0095	0,0085
	98	0,0095	0,0090	0,0175	-

Tabela 5, Continua

	2	0,0090	0,0085	0,0075	0,0120
200	10	0,0115	0,0085	0,0110	0,0140
	50	0,0110	0,0060	0,0085	0,0080
	100	0,0135	0,0070	0,0135	0,0100
10000	5	0,0082	0,0091	-	0,0080

Vale evidenciar que, conforme pode ser visto nas Tabelas 3, 4 e 5, o TNMt foi exato em praticamente todas as combinações, n e p , para $\alpha = 10\%$ e $\alpha = 5\%$ e $\alpha = 1\%$, respectivamente. Obviamente isso é um fato interessante, uma vez que a estatística tem uma distribuição nula que é uma aproximação assintótica qui-quadrado. Logo, esperaria-se que o teste não fosse exato em pequenas amostras.

4.2 Poder do teste

O estudo do poder do teste tem como objetivo conhecer o quanto o teste de hipóteses controla o erro tipo II ou o quanto ele é poderoso, ou seja, qual a probabilidade de rejeitar a hipótese nula se realmente for falsa, considerando-se vários cenários sob H_1 . Na prática, é importante que se tenham testes com nível de significância próximos do nível de significância nominal (erro tipo I) e que o poder seja alto. O poder de um teste de hipótese pode ser afetado negativamente por três fatores: (i) quando o tamanho da amostra é pequena; (ii) se o nível de significância for baixo; e (iii) no caso em que a diferença entre o “verdadeiro” valor do parâmetro e o valor especificado pela hipótese nula seja pequeno. No caso em questão, essa distância não pode ser quantificada, mas a sua ideia é de que quanto mais afastada do modelo normal multivariada for a distribuição, maior deverá ser o poder do teste. A quantificação de quão afastada está a distribuição dos dados em relação ao modelo normal não pode ser feita diretamente de forma quantitativa.

Nas tabelas a seguir estão apresentados os resultados de poder dos quatro testes de normalidade multivariada, TMCNM, TNMBP, TSWNM e TNMt. Para avaliar o poder do teste foram escolhidas as distribuições não-normais multivari-

adas: *t-Student* multivariada com $\nu = 1$ e $\nu = 30$, log-normal multivariada, normal contaminada multivariada e uniforme multivariada. Os níveis de significâncias foram os mesmos fixados para o estudo das taxas de erro tipo I anteriormente apresentado, ou seja, $\alpha = 10\%$, $\alpha = 5\%$ e $\alpha = 1\%$.

O poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição *t-Student* multivariada com $\nu = 1$ grau de liberdade em função de n e p , para $\alpha = 10\%$, está indicado na Tabela 6. O TMCNMD apresentou baixos valores de poder nas situações em que a amostra era pequena como, por exemplo, para $n = 10$ e quando o valor de n se aproxima do valor do p ($n = 20$ com $p = 10$ e 18 , $n = 50$ com $p = 48$ e $n = 100$ com $p = 98$).

O TNMBP apresentou os melhores desempenhos em relação aos demais testes, atingindo 100% a partir da configuração de $n = 20$ com $p = 18$. O TSWNM só atingiu 100% de desempenho em poder quando a amostra foi superior a $n = 50$, para qualquer valor de p . Para as demais configurações de n e p , ele também teve um excelente desempenho.

Já o TNMt apresentou poder muito baixo em todas as combinações de n e p avaliadas, com exceção, de quando a amostra foi muito grande, $n = 10000$. Em relação à amostra de tamanho 10000, o TMCNMD, TNMBP e o TNMt apresentaram 100% de poder, o que é uma grande vantagem em relação ao teste TSWNM, pois esse é restrito às condições de amostra menores que 5000.

É conveniente salientar que a *t-Student* multivariada com $\nu = 1$ grau de liberdade é uma distribuição bem diferenciada da normal multivariada. Dessa forma o ideal é se tenha valores elevados de poder, o que não aconteceu com o teste proposto (TNMt). Portanto, o TNMt não é considerado um teste adequado para a maior parte das situações apresentadas.

Tabela 6 Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição *t-Student* multivariada com $\nu = 1$ grau de liberdade em função de n e p para $\alpha = 10\%$.

n	p	TMCNMD	TNMBP	TSWNM	TNMt
10	2	0,5645	0,8040	0,7690	0,1585
	4	0,2530	0,8675	0,8345	0,1565
	6	0,1250	0,9065	0,8790	0,1165
	8	0,2000	0,9380	0,9120	-
20	2	0,9060	0,9690	0,9555	0,2765
	5	0,8955	0,9940	0,9900	0,1770
	10	0,4280	0,9975	0,9955	0,1185
	18	0,3580	1,0000	0,9985	-
50	2	1,0000	1,0000	1,0000	0,3875
	10	1,0000	1,0000	1,0000	0,2125
	25	0,9675	1,0000	1,0000	0,1025
	48	0,6370	1,0000	1,0000	-
100	2	1,0000	1,0000	1,0000	0,4440
	10	1,0000	1,0000	1,0000	0,3980
	50	1,0000	1,0000	1,0000	0,1055
	98	0,8455	1,0000	1,0000	-
200	2	1,0000	1,0000	1,0000	0,4980
	10	1,0000	1,0000	1,0000	0,5730
	50	1,0000	1,0000	1,0000	0,1785
	100	1,0000	1,0000	1,0000	0,1155
10000	5	1,0000	1,0000	-	0,9990

Os desempenhos para o poder dos TMCNMD, TNMBP, TSWNM e TNMt, considerando a distribuição *t-student* multivariada com $\nu = 1$ para $\alpha = 5\%$ e

$\alpha = 1\%$ são mostrados, respectivamente, nas Tabelas 7 e 8 em função de n e p .

Notou-se, para o TMCNMD, com tamanhos de amostrais n iguais a 10 e 20, que o aumento do número de variáveis fizeram o poder diminuir, chegando a ficar abaixo do nível nominal de significância, em alguns casos. O TMCNMD apresentou bom desempenho para tamanhos grandes de amostras, $n = 200$, independentemente do número de variáveis (Tabelas 7 e 8).

O TNMBP apresentou melhor desempenho que o TMCNMD e o TNMt nas situações em que $n = 10$ e $n = 20$, para todas as combinações de p consideradas e, inclusive, o TNMBP mostrou-se melhor que o TSWNM para os mesmos tamanhos amostrais de 10 e 20, com exceção somente dos casos do $\alpha = 1\%$, para $n = 10$ com $p = 8$ e $n = 20$ com $p = 10$ (Tabela 8), porém, essas diferenças foram relativamente pequenas chegando a ser insignificantes. O TSWNM apresentou ótimo desempenho do teste em relação ao poder, atingindo 100% a partir da situação de $n = 50$, para qualquer valor de p .

O TNMt para o caso de $\alpha = 5\%$ e $\alpha = 1\%$ não teve um desempenho satisfatório, exceto quando a amostra foi de 10000. Retirando esse caso, o poder do teste oscilou de 5,35% ($n = 50$ com $p = 25$) a 48,70% ($n = 200$ com $p = 10$) para o caso de $\alpha = 5\%$ (Tabela 7). No caso de $\alpha = 1\%$ (Tabela 8) esse intervalo varia de 0,90 ($n = 200$ com $p = 100$) a 41,80 ($n = 200$ com $p = 2$) pontos percentuais.

Tabela 7 Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição *t-Student* multivariada com $\nu = 1$ grau de liberdade em função de n e p para $\alpha = 5\%$.

n	p	TMCNMD	TNMBP	TSWNM	TNMt
10	2	0,3930	0,7495	0,7185	0,0990
	4	0,1450	0,8210	0,7945	0,0890
	6	0,0605	0,8690	0,8425	0,0645
	8	0,0970	0,9100	0,8855	-

Tabela 7, Continua

20	2	0,8635	0,9555	0,9375	0,2010
	5	0,8010	0,9905	0,9850	0,1135
	10	0,2705	0,9955	0,9940	0,0685
	18	0,2550	0,9990	0,9980	-
50	2	0,9970	1,0000	1,0000	0,3415
	10	1,0000	1,0000	1,0000	0,1385
	25	0,9275	1,0000	1,0000	0,0535
	48	0,5250	1,0000	1,0000	-
100	2	1,0000	1,0000	1,0000	0,4040
	10	1,0000	1,0000	1,0000	0,2995
	50	1,0000	1,0000	1,0000	0,0580
	98	0,7590	1,0000	1,0000	-
200	2	1,0000	1,0000	1,0000	0,4640
	10	1,0000	1,0000	1,0000	0,4870
	50	1,0000	1,0000	1,0000	0,1050
	100	1,0000	1,0000	1,0000	0,0610
10000	5	1,0000	1,0000	-	0,9985

O TNMBP mostrou-se equivalente ao desempenho do TSWNM para tamanhos de amostras maiores ou iguais a 50, em que os valores de poder desses dois testes atingiram 100%.

É importante ressaltar que para a distribuição *t-Student* multivariada com $\nu = 1$ para $\alpha = 5\%$ e $\alpha = 1\%$, o TNMBP, foi melhor ou igual ao TSWNM. Já o TNMt não é eficiente para se testar normalidade multivariada de um banco de dados com amostras menores que 10000 (Tabela 7 e 8).

Tabela 8 Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição *t-Student* multivariada com $\nu = 1$ grau de liberdade em função de n e p para $\alpha = 1\%$.

n	p	TMCNMD	TNMBP	TSWNM	TNMt
10	2	0,1600	0,6255	0,6070	0,0370
	4	0,0245	0,7110	0,6990	0,0335
	6	0,0055	0,7590	0,7495	0,0125
	8	0,0250	0,8165	0,8190	-
20	2	0,7430	0,9095	0,8955	0,0995
	5	0,5035	0,9760	0,9715	0,0345
	10	0,0625	0,9890	0,9895	0,0160
	18	0,1070	0,9965	0,9950	-
50	2	0,9875	1,0000	1,0000	0,2485
	10	0,9985	1,0000	1,0000	0,0545
	25	0,7180	1,0000	1,0000	0,0125
	48	0,2920	1,0000	1,0000	-
100	2	1,0000	1,0000	1,0000	0,3365
	10	1,0000	1,0000	1,0000	0,1675
	50	0,9980	1,0000	1,0000	0,0125
	98	0,4890	1,0000	1,0000	-
200	2	1,0000	1,0000	1,0000	0,4180
	10	1,0000	1,0000	1,0000	0,3415
	50	1,0000	1,0000	1,0000	0,0340
	100	1,0000	1,0000	1,0000	0,0090
10000	5	1,0000	1,0000	-	0,9945

Nas Tabelas 9, 10 e 11 os desempenhos de poder para os testes TMCNMD, TNMBP, TSWNM e TNMt são mostrados, considerando a distribuição *t-Student*

multivariada com $\nu = 30$ para $\alpha = 10\%$, $\alpha = 5\%$ e $\alpha = 1\%$, respectivamente, em função de n e p . Vale ressaltar que quando os graus de liberdade da distribuição *t-Student* são muito grandes ($\nu \rightarrow \infty$) essa distribuição tende à normal multivariada, por isso, os desempenhos de poder ficam bem próximos do nível nominal de significância. A medida que o tamanho da amostra cresce o poder tende a aumentar gradativamente. Portanto, a distribuição *t-Student* multivariada com $\nu = 30$ assemelha-se à distribuição normal multivariada. Em razão disso, verifica-se que os desempenhos do poder são quase equiparáveis ao nível nominal para amostras pequenas nos quatro testes de normalidade multivariada que estão sendo comparados.

Na Tabela 9 estão apresentados os desempenhos de poder para os TMCNMD, TNMBP, TSWNM e TNMt, considerando a distribuição *t-Student* multivariada com $\nu = 30$ para $\alpha = 10\%$ em função de n e p . Tem-se que o TMCNMD apresentou melhor desempenho em relação aos demais testes em três situações, duas em que a amostra era pequena com $n = 10$ e com números de variáveis p iguais a 4 e 8, embora praticamente desprezível, e a outra quando a amostra era grande $n = 200$ com $p = 10$.

O TNMt só apresentou desempenho melhor do que os demais testes quando $n = 10$ e $p = 2$, sendo que, a diferença foi pouco expressiva em relação ao TNMBP. Nas outras situações, o TNMt apresentou menor poder em relação aos outros três testes comparados. Enquanto os demais testes apresentaram um aumento de poder com o aumento de p , o TMCNMD tem um comportamento diferente. Para o TMCNMD à medida que p aumenta, o poder deste teste aumenta até um certo ponto e depois decresce. Quanto mais próximo de n fica p , menor é o poder.

No caso do TNMt os valores do poder variaram de 0,0865 a 0,1220, em todas as combinações de n com p , até mesmo quando $n = 10000$. Portanto, o TNMt não é poderoso. Assim, ele não é recomendado para julgar se uma determinada amostra segue uma distribuição normal multivariada, por não conseguir distinguir uma situação de normalidade de outra de não-normalidade.

Tabela 9 Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição *t-Student* multivariada com $\nu = 30$ grau de liberdade em função de n e p para $\alpha = 10\%$.

n	p	TMCNMD	TNMBP	TSWNM	TNMt
10	2	0,1030	0,1105	0,1065	0,1155
	4	0,1125	0,1015	0,1030	0,1055
	6	0,0950	0,1215	0,1070	0,0990
	8	0,1030	0,0995	0,0880	-
20	2	0,1215	0,1325	0,1220	0,1015
	5	0,0985	0,1380	0,1200	0,1040
	10	0,0940	0,1570	0,1350	0,0910
	18	0,1270	0,1610	0,1405	-
50	2	0,1395	0,1630	0,1385	0,0945
	10	0,1465	0,2165	0,1750	0,0995
	25	0,0745	0,2805	0,2180	0,1010
	48	0,1360	0,2965	0,2315	-
100	2	0,1765	0,2055	0,1665	0,1110
	10	0,2710	0,2980	0,2265	0,1015
	50	0,0905	0,4215	0,3170	0,0865
	98	0,1280	0,4565	0,3340	-
200	2	0,2225	0,2305	0,1865	0,1070
	10	0,4875	0,3975	0,3020	0,0960
	50	0,5395	0,5870	0,4525	0,0955
	100	0,1305	0,6535	0,4910	0,1150
10000	5	1,0000	1,0000	-	0,1220

Os desempenhos quanto ao poder para os testes TMCNMD, TNMBP, TSWNM e TNMt, considerando a distribuição *t-Student* multivariada com $\nu = 30$

para $\alpha = 5\%$, são mostrados na Tabela 10, em função de n e p .

O TMCNMD mostrou-se melhor do que o TNMBP, TSWNM e TNMt, na situação $n = 200$ com $p = 10$. O TNMBP obteve melhor desempenho em quase todas as situações em relação aos outros testes. O TNMBP só foi considerado com desempenho intermediário nas situações em que o tamanho da amostra foi pequeno $n = 10$, com número de variáveis $p = 4$ (o TNMt obteve o mesmo resultado) e $p = 8$. Também ocorreu isso na situação em que $n = 200$ e $p = 10$, condição em que o TMCNMD foi considerado melhor.

O TNMt apresentou os mais baixos valores de poder em relação aos outros testes quando $n > 20$ para qualquer valor de p escolhido, com exceção, de $p = 5$. Na amostra de tamanho $n = 10000$, o TMCNMD e o TNMBP atingiram 100% do desempenho, já o TNMt só obteve 12,20% e o TSWNM é limitado para amostras de tamanha 5000.

Tabela 10 Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição *t-Student* multivariada com $\nu = 30$ grau de liberdade em função de n e p para $\alpha = 5\%$.

n	p	TMCNMD	TNMBP	TSWNM	TNMt
10	2	0,0500	0,0560	0,0515	0,0655
	4	0,0555	0,0510	0,0465	0,0510
	6	0,0470	0,0585	0,0575	0,0455
	8	0,0530	0,0460	0,0455	-
20	2	0,0700	0,0765	0,0605	0,0480
	5	0,0555	0,0730	0,0630	0,0525
	10	0,0405	0,0920	0,0795	0,0515
	18	0,0720	0,0850	0,0805	-
50	2	0,0840	0,0965	0,0780	0,0480
	10	0,0920	0,1360	0,1020	0,0560
	25	0,0380	0,1750	0,1490	0,0475
	48	0,0780	0,1775	0,1585	

Tabela 10, Continua

	2	0,1160	0,1265	0,0995	0,0525
100	10	0,1805	0,1990	0,1465	0,0540
	50	0,0460	0,2960	0,2275	0,0410
	98	0,0770	0,3060	0,2535	-
		2	0,1450	0,1575	0,1130
200	10	0,3730	0,2885	0,2070	0,0515
	50	0,4200	0,4540	0,3480	0,0495
	100	0,0690	0,4940	0,3960	0,0645
		2	0,1450	0,1575	0,1130
10000	5	1,0000	1,0000	-	0,0610

Na Tabela 11, são apresentados os valores de poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição *t-Student* multivariada com $\nu = 30$ graus de liberdade em função de n e p para $\alpha = 1\%$.

O padrão do comportamento dos testes quanto ao poder foi similar ao observado para $\alpha = 10\%$ e $\alpha = 5\%$. 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, só no caso do TNMt que esse valores permaneceram baixos para todos as situações, até mesmo com amostra muito grande ($n = 10000$).

Tabela 11 Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição *t-Student* multivariada com $\nu = 30$ grau de liberdade em função de n e p para $\alpha = 1\%$.

n	p	TMCNMD	TNMBP	TSWNM	TNMt
10	2	0,0135	0,0100	0,0110	0,0135
	4	0,0125	0,0115	0,0105	0,0085
	6	0,0080	0,0095	0,0120	0,0100
	8	0,0105	0,0015	0,0070	-

Tabela 11, Continua

20	2	0,0160	0,0160	0,0145	0,0105
	5	0,0110	0,0160	0,0125	0,0110
	10	0,0070	0,0185	0,0220	0,0120
	18	0,0190	0,0170	0,0265	-
50	2	0,0245	0,0275	0,0210	0,0115
	10	0,0270	0,0415	0,0420	0,0130
	25	0,0045	0,0515	0,0680	0,0115
	48	0,0215	0,0550	0,0730	-
100	2	0,0395	0,0405	0,0320	0,0105
	10	0,0685	0,0760	0,0640	0,0120
	50	0,0080	0,1015	0,1115	0,0110
	98	0,0215	0,1080	0,1450	-
200	2	0,0470	0,0495	0,0385	0,0140
	10	0,1855	0,1350	0,1005	0,0115
	50	0,2165	0,2140	0,2160	0,0110
	100	0,0110	0,2225	0,2475	0,0125
10000	5	1,0000	1,0000	-	0,0130

Os desempenhos de poder para os testes TMCNMD, TNMBP, TSWNM e TNMt, considerando a distribuição log-normal multivariada para $\alpha = 10\%$, encontram-se na Tabela 12. O TMCNMD e o TNMt não obtiveram bom desempenho para tamanhos de amostras n com valores iguais a 10. Para amostras de tamanho 20, o TMCNMD apresentou baixo desempenho com números de variáveis iguais a 10 e 18, para os demais números de variáveis consideradas nesse caso, o teste TMCNMD apresentou poder superior 83,00%.

O desempenho do TNMBP aproximou-se do desempenho do TSWNM, ocorrendo o mesmo valor de poder para $n = 10$ com $p = 6$ com uma taxa de 99,05%. Para tamanhos amostrais iguais ou maiores que 20, o desempenho foi de

100%, exceto para $n = 20$ com $p = 2$. Os desempenhos do TSWNM mostraram-se superiores aos demais testes em algumas situações, embora praticamente equivalente ao do TNMBP.

O TNMt obteve o seu melhor desempenho quando o valor de p é intermediário, ou seja, em valores centrais de p para cada combinação de n , com exceção de $n = 10$. Para a amostra de 10000, os desempenhos dos testes TMCNMD, TNMBP e TNMt, foram máximos, ou seja, rejeitaram 100% das vezes a hipótese nula.

Tabela 12 Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição log-normal multivariada em função de n e p para $\alpha = 10\%$

n	p	TMCNMD	TNMBP	TSWNM	TNMt
10	2	0,5475	0,8525	0,8635	0,0345
	4	0,2630	0,9605	0,9655	0,0385
	6	0,1490	0,9905	0,9905	0,1175
	8	0,1380	0,9975	0,9985	-
20	2	0,8395	0,9920	0,9945	0,0145
	5	0,8685	1,0000	1,0000	0,2215
	10	0,3855	1,0000	1,0000	0,1270
	18	0,3840	1,0000	1,0000	-
50	2	0,9915	1,0000	1,0000	0,0005
	10	1,0000	1,0000	1,0000	0,4195
	25	0,8420	1,0000	1,0000	0,1180
	48	0,7080	1,0000	1,0000	-
100	2	1,0000	1,0000	1,0000	0,0015
	10	1,0000	1,0000	1,0000	0,8970
	50	0,9955	1,0000	1,0000	0,1375
	98	0,9335	1,0000	1,0000	-

Tabela 12, Continua

	2	1,0000	1,0000	1,0000	0,0010
200	10	1,0000	1,0000	1,0000	0,9985
	50	1,0000	1,0000	1,0000	0,6470
	100	1,0000	1,0000	1,0000	0,1665
10000	5	1,0000	1,0000	-	1,0000

O poder dos testes TMCNMD, TNMBP, TSWNM e TNMt, considerando a distribuição log-normal multivariada em função de n e p para $\alpha = 5\%$, apresentou padrão de comportamento semelhante ao de 10% e os resultados não serão apresentados.

O TMCNMD junto com o TNMt apresentou os piores valores de poder nas situações em que a amostra é pequena, casos em que n possui os valores de 10 e 20, sendo que o TNMt ainda teve os resultados inferiores ao do TMCNMD. O TMCNMD atingiu 100% de desempenho nas situações em que $n = 50$ e $n = 100$ com número de variáveis $p = 10$ e, também, nas situações em que o tamanho da amostra foi $n = 200$, para todos os números de variáveis consideradas. Já o TNMt só obteve uma melhora em seu desempenho para os casos: $n = 20$ com $p = 5$, $n = 50$ com $p = 10$, $n = 100$ com $p = 10$, $n = 200$ com $p = 10$ e $p = 50$.

O TNMBP foi considerado com um comportamento intermediário entre o TMCNMD e o TSWNM, nas situações $n = 10$, em que a amostra é pequena. No entanto, essas estimativas dos desempenhos do poder para esse tamanho amostral no TNMBP foram relativamente próximas das estimativas dos desempenhos do poder do TSWNM. O poder foi de 100% com $n = 10000$ e $p = 5$ para o TMCNMD, o TNMBP e o TNMt (resultados não apresentados).

Os desempenhos quanto ao poder para TMCNMD, TNMBP e TSWNM, considerando a distribuição log-normal multivariada para $\alpha = 1\%$ novamente apresentaram o mesmo padrão e não serão apresentados.

O TMCNMD e o TNMt não conseguiram um desempenho satisfatório para tamanhos de amostras n com valores iguais a 10 e 20. O TNMt para os diversos valores de amostras combinado com os das variáveis sugeridos, teve desempenho ruim, zerando nas situações em que $n = 50$ com $p = 2$ e $n = 100$

com $p = 2$. Para amostras de tamanho $n = 100$, o TMCNMD apresentou bom desempenho para $p = 10$, atingindo a eficiência máxima de 100% (resultados não apresentados).

Em relação ao TNMBP e ao TSWNM, destaca-se que o comportamento foi semelhante ao já detalhado para $\alpha = 10\%$ e $\alpha = 5\%$. O que se observa é uma redução do poder, como esperado pela redução de α .

Os desempenhos de poder para os testes TMCNMD, TNMBP, TSWNM e TNMt, considerando a distribuição normal contaminada multivariada para $\alpha = 10\%$ são apresentados na Tabela 13 em função de n e p . O TMCNMD não obteve um desempenho adequado para tamanhos de amostras n com valores iguais a 10 e 20. Para tamanho de amostras iguais a $n = 50$, o teste mostrou-se melhor quando o número de variáveis foi igual a $p = 25$, para os demais valores de variáveis o teste mostrou baixo desempenho. Já para os valores de amostras acima de 200 e $p > 2$ o desempenho desse teste foi satisfatório.

O TNMBP, assim como o TMCNMD, TSWNM e TNMt, apresentou baixo desempenho para amostras de tamanho $n = 10$, embora o TNMBP e o TSWNM apresentaram valores de poder bem superiores ao TMCNMD e o TNMt. Para amostras de tamanho $n = 20$, houve melhoras quanto ao desempenho, mas ainda o seu desempenho não superou o do TSWNM.

Para amostras de tamanho 100 e 200 o desempenho obtido no TNMBP e no TSWNM foram máximos, mostrando-se excelente. O TNMt só apresentou um bom desempenho para amostra igual a 10000, sendo que o TMCNMD e o TNMBP também o obtiveram. O TSWNM mostrou-se melhor do que todos os testes apresentados com as configurações escolhidas para n e p .

Tabela 13 Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição normal contaminada multivariada em função de n e p para $\alpha = 10\%$

n	p	TMCNMD	TNMBP	TSWNM	TNMt
10	2	0,2475	0,3385	0,4380	0,0525
	4	0,1770	0,3770	0,5000	0,0460
	6	0,1220	0,4010	0,5315	0,0970
	8	0,1260	0,4235	0,5560	-

Tabela 13, Continua

	2	0,2905	0,6720	0,8070	0,0145
20	5	0,2515	0,7925	0,8945	0,1200
	10	0,2685	0,8575	0,9395	0,0985
	18	0,3805	0,8765	0,9560	-
50	2	0,3930	0,9945	0,9975	0,0030
	10	0,4435	1,0000	1,0000	0,1025
	25	0,9990	1,0000	1,0000	0,0920
	48	0,6730	1,0000	1,0000	-
100	2	0,5125	1,0000	1,0000	0,0005
	10	0,7255	1,0000	1,0000	0,1070
	50	1,0000	1,0000	1,0000	0,0880
	98	0,8755	1,0000	1,0000	-
200	2	0,7365	1,0000	1,0000	0,0000
	10	0,9675	1,0000	1,0000	0,1165
	50	1,0000	1,0000	1,0000	0,1050
	100	1,0000	1,0000	1,0000	0,1030
10000	5	1,0000	1,0000	-	1,0000

O desempenho dos dois testes propostos por Biase (2011), TMCNMD e TNMBP, e o TSWNM proposto por Royston (1983) em comparação ao TNMt, proposto nesse trabalho, quando considerada a distribuição normal contaminada multivariada em função de n e p para o nível de significância de $\alpha = 5\%$ e para $\alpha = 1\%$, apresentou o mesmo padrão de resposta observado para $\alpha = 10\%$. É evidente que a redução de α provoca uma redução nos valores de poder dos testes.

Os quatro testes comparados entre si, de forma geral, não apresentaram desempenho satisfatório para amostras pequenas. A partir de amostras maiores ou iguais a 50, os testes TNMBP e TSWNM foram melhores quanto à detecção de não normalidade dos dados. No caso em que foi contemplado o tamanho amostral

10000 para o TNMDC, o TNMBP e o TNMt, o desempenho foi excelente, atingindo os 100% tanto para $\alpha = 10\%$, $\alpha = 5\%$ e $\alpha = 1\%$.

A Tabela 14 apresenta os valores de poder do novo teste proposto, TNMt, e dos testes TMCNMD, TNMBP e TSWNM, para a distribuição uniforme multivariada em função de n e p para o nível de significância de $\alpha = 10\%$.

De maneira geral, o TMCNMD e o TNMt apresentaram desempenhos inferiores aos dos outros dois concorrentes. A partir de amostras de tamanho $n = 50$ com $p = 25$, o TSWNM atingiu 100% de desempenho, assim como o TSWNM.

Para amostras pequenas ($n = 10$), os quatro testes apresentaram baixos valores de poder. Para o tamanho amostral 10000 os testes obtiveram resultados máximos de 100% em relação ao poder. O TNMt teve o mesmo comportamento, em relação as outras distribuições não normais, referente ao poder do teste (baixo poder).

Tabela 14 Poder dos testes TMCNMD, TNMBP, TSWNM e TNMt considerando a distribuição uniforme multivariada em função de n e p para $\alpha = 10\%$

n	p	TMCNMD	TNMBP	TSWNM	TNMt
10	2	0,1555	0,1130	0,1860	0,0840
	4	0,1295	0,1270	0,2400	0,0855
	6	0,0940	0,1380	0,2800	0,1070
	8	0,0885	0,1270	0,2790	-
20	2	0,2215	0,2215	0,4450	0,0790
	5	0,1420	0,3140	0,6715	0,0915
	10	0,0980	0,4550	0,8605	0,1055
	18	0,1370	0,5740	0,9405	-
50	2	0,4905	0,8565	0,9710	0,0940
	10	0,0905	0,9995	1,0000	0,0915
	25	0,2710	1,0000	1,0000	0,1060
	48	0,2395	1,0000	1,0000	-

Tabela 14, Continua

	2	0,7655	1,0000	1,0000	0,1150
100	10	0,0850	1,0000	1,0000	0,0930
	50	0,9455	1,0000	1,0000	0,0975
	98	0,3750	1,0000	1,0000	-
200	2	0,9840	1,0000	1,0000	0,1280
	10	0,1505	1,0000	1,0000	0,0845
	50	0,9995	1,0000	1,0000	0,0975
	100	1,0000	1,0000	1,0000	0,0990
10000	5	1,0000	1,0000	-	1,0000

Os resultados de poder dos testes TMCNMD, TNMBP, TSWNM e TNMt, considerando a distribuição uniforme multivariada em função de n e p para $\alpha = 5\%$ e $\alpha = 1\%$, respectivamente, apresentaram o mesmo padrão em relação aos resultados observados para $\alpha = 10\%$ quanto ao poder dos quatro testes.

O que se pode destacar é que o desempenho do TSWNM em relação aos demais testes, para amostras pequenas $n = 20$, foi bem superior ao que já havia sido observado para a distribuição normal contaminada (resultados não apresentados).

O TNMBP tem os valores próximos dos resultados obtidos pelo TSWNM. O TMCNMD e o TNMt foram os que tiveram desempenhos mais precários. Em relação ao poder, em nenhum caso o TNMt foi avaliado como satisfatório. Sendo assim, não é indicado como um teste apropriado para normalidade multivariada.

Como considerações importantes, observa-se que para todas as distribuições não-normais consideradas, o poder do TMNt foi muito pequeno para todos os tamanhos de amostras, com exceção de $n = 10000$, situação em que o teste atingiu desempenho máximo de 100%, menos no caso em que a distribuição foi *t-Student* com $\nu = 30$ graus de liberdade. Isso é justificado, pois, como os graus de liberdade utilizados ($\nu = 30$) foram grandes, essa distribuição tende à normal, por isso, os desempenhos de poder ficaram bem próximos do nível nominal de significância. Vale destacar, também, que quando a amostra efetiva ($n - p - 1$), isto

é, a amostra usada de fato pelo TNMt depois de feita as devidas transformações, é próxima da unidimensionalidade ou quando o tamanho da amostra efetiva for a maior possível (Tabela 15), o poder do teste apresenta-se pior do que quando os tamanhos da amostra efetiva são intermediários.

Tabela 15 Tamanho da amostra efetiva utilizada no TNMt.

n	p	AE	n	p	AE
10	2	7	100	2	97
	4	5		10	89
	6	3		50	49
	8	1		98	1
20	2	17	200	2	197
	5	14		10	189
	10	9		50	149
	18	1		100	99
50	2	47	10000	5	9994
	10	39			
	25	24			
	48	1			

AE = Amostra efetiva

Para tamanhos de amostras grandes, o desempenho do TMCNMD para distribuição log-normal multivariada foi o mesmo observado para distribuição *t-Student* multivariado com $\nu = 1$, isto é, o poder decrescia à medida que o número de variáveis aumentava e tornava-se próximo do tamanho amostral. Com o aumento do tamanho da amostra o poder do teste estabilizou-se, em geral, em 100%. Para a distribuição uniforme multivariada, baixos valores de poder foram observados em geral.

O comportamento do TNMBP e do TSWNM foram semelhantes para as distribuições não normais comparadas. Observou-se para esses testes um excelente desempenho em relação ao poder para distribuições assimétricas à direita

(log-normal) e de cauda pesada e simétrica (*t-Student* com $\nu = 1$). Para a uniforme, de cauda muito pesada e simétrica, o TNMBP e o TSWNM tiveram menor poder em amostras pequenas e um excelente desempenho em amostras de tamanhos maiores ou iguais a 50. Esse mesmo padrão de resposta para os dois testes foram observados para a normal contaminada.

Para o caso da distribuição *t-Student* multivariada com $\nu = 30$, tanto o teste proposto, TNMt, quanto os testes propostos por Biase (2011), TMCNMD e TNMBP, e o teste proposto por Royston (1983), TSWNM, obtiveram valores pequenos de poder.

4.3 Aplicação

Para exemplificar e ilustrar o teste proposto e outros testes de normalidade multivariada apresentado nesse trabalho, foi utilizado um conjunto de dados fornecido por Royston (1983), referente à concentração de hemoglobina e contagem de linfócitos. Foi utilizado o programa *R* como ferramenta de análise, sendo necessário instalar e carregar as bibliotecas: “asbio”, para o teste de Doornik-Hansen de normalidade multivariada e o “MVN” para executar o teste de Henze-Zirkler de normalidade multivariada. Para a utilização do teste proposto, TNMt, foram usadas funções elaboradas pela própria autora. Para o teste de Shapiro-Wilk para normalidade multivariada foram usadas as funções elaboradas por Silva (2009) e para o teste Monte Carlo de normalidade multivariada baseado em distâncias e o teste de normalidade multivariada baseado em *bootstrap* paramétrico foram usadas as funções elaboradas por Biase (2011).

Os dados referem-se às informações de 20 pintores de carro sobre a concentração de hemoglobina (X_1) e contagem de linfócitos (X_2).

Os valores de referência do hemograma completo, geralmente, variam de acordo com o sexo e idade do paciente, no entanto, também é possível observar diferenças nos valores dependendo do laboratório onde foi feita a colheita. O hemograma é utilizado para avaliar determinados aspectos sanguíneos como o número de glóbulos vermelhos, leucócitos e plaquetas, sendo uma boa forma para identificar a presença de infecções, excesso de ferro ou de anemia, por exemplo.

Vale destacar que a hemoglobina (Figura 3(a)) é o pigmento que dá a cor aos glóbulos vermelhos (eritrócitos) e tem a função vital de transportar o oxigênio dos pulmões aos tecidos e linfócito (Figura 3(b)) que é um tipo de leucócito (glóbulo branco) presente no sangue e que realiza a defesa do organismo contra agentes infecciosos (vírus, bactérias, substâncias alergênicas e outros corpos hostis que podem penetrar no corpo humano).

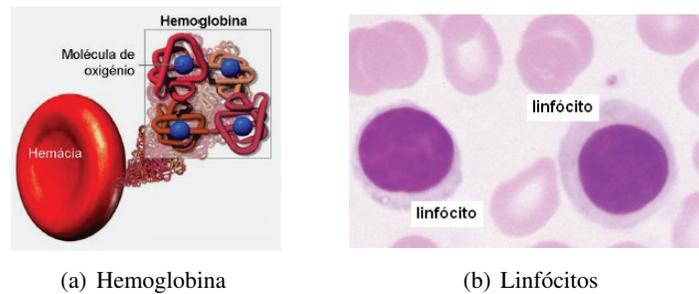


Figura 3 Hemoglobina e linfócitos

Na Tabela 16 foram apresentados os dados amostrais referentes às variáveis em estudo (X_1 e X_2), hemoglobina e número de linfócitos, que foram utilizados em uma aplicação em Royston (1983).

Tabela 16 Conjunto de dados referente à concentração de hemoglobina e contagem de linfócitos apresentado por Royston (1983).

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

Para a aplicação do teste de Shapiro-Wilk de normalidade multivariada é

necessário carregar a função `Shapiro.Royston` implementada por Silva (2009) (anexo).

Depois de carregar a função, estabelecer o banco de dados:

```
# X1: concentração de hemoglobina e
# X2: contagem de linfócitos.
X1 <- c(13.4, 14.6, 13.5, 15.0, 14.6, 14.0, 16.4, 14.8,
        15.2, 15.5, 15.2, 16.9, 14.8, 16.2, 14.7, 14.7,
        16.5, 15.4, 15.1, 14.2)
X2 <- c(14,15,19,23,17,20,21,16,27,34,26,28,24,26,23,9,
        18,28,17,14)
```

```
dados <- data.frame(X1,X2)
```

Executando a função fornecida pela autora:

```
# Teste de Shapiro-Wilk normalidade multivariado
x <- cbind(X1,X2)
Shapiro.Royston(x)
```

É importante lembrar que esse teste tem uma limitação a respeito do tamanho para cada amostra (entre 3 e 5000). Nas colunas da matriz utilizada na função `Shapiro.Royston` estão as variáveis em estudo ($p = 2$) e nas linhas estão as observações de cada variável (a amostra tem tamanho $n = 20$).

A saída do programa *R* é dada por:

```
\$W
[1] 0.9706747
```

```
\$G
[1] 0.1788431
```

```
\$H
[1] 0.3547996
```

```
\$dfeq
[1] 1.983859
```

```
\$p.value
```

```
[1] 0.8345404
```

Para a aplicação do teste Monte Carlo de normalidade multivariada baseado em distâncias (TMCNMD) proposto por Biase (2011) é necessário carregar a função `testNMMC` elaborada pela a mesma (anexo).

```
# Teste Monte Carlo de normalidade multivariada baseado em
# distâncias
> X <- cbind(X1,X2)
> testNMMC(X)
```

A saída gerada pelo programa *R* é dada por:

```
$r2
[1] 0.9384081
```

```
$valor.p
[1] 0.2753623
```

A função `testNMMC` além de apresentar a estatística do teste e o valor-*p*, ela também, fornece o *Q-Q plot* (Figura 4).

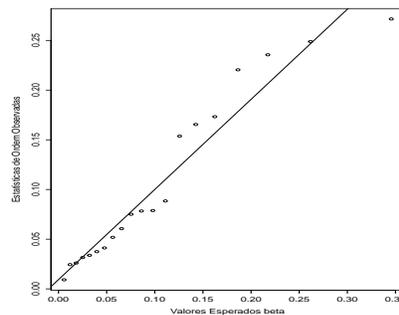


Figura 4 Q-Q plot fornecido pelo teste Monte Carlo de normalidade multivariada baseado em distâncias

O teste de normalidade multivariada baseado em *bootstrap* paramétrico (TNMBP), proposto por Biase (2011), exige, para a sua utilização no programa *R*, que seja carregada a função `testHNMMC`, elaborada pela a mesma (anexo).

```
# Teste de normalidade multivariada baseado em bootstrap
# paramétrico
> X <- cbind(X1,X2)
> testHNMMC(X)
```

A função `testHNMMC` tem como argumento a matriz de dados X e obteve a seguinte saída no programa *R*:

```
$H
[1] 0.9720267

$valor.p
[1] 0.8330835
```

O teste de Doornik-Hansen de normalidade multivariada requer o uso da biblioteca `asbio`:

```
install.packages("asbio")
library(asbio)
```

A função para a aplicação desse teste é dada por:

```
Y <- dados
DH.test(Y)
```

A saída do programa:

```
$multi
      E df P(Chi > E)
1 10.49981  4 0.03279965

$univ
      E df P(Chi > E)
Y1  9.203391  2 0.01003481
Y2  1.296417  2 0.52298198
```

A função `DH.test` fornece o teste de normalidade multivariado e univariado dos dados.

Para o teste de Henze-Zirkler de normalidade multivariada, é necessário que se instale e carregue o pacote “`MVN`”, da seguinte forma:

```
install.packages("MVN")
library(MVN)
```

Sua utilização se dá da seguinte forma:

```
hzTest(dados, qqplot = TRUE)
```

A função `hzTest`, além de disponibilizar o valor da estatística do teste e do valor- p , ela também apresenta o gráfico de quantil-quantil (Figura 5). A saída do programa é:

```
Henze-Zirkler's Multivariate Normality Test
-----
data : data
HZ      : 0.3624304
p-value : 0.5505191
Result  : Data are multivariate normal.
-----
```

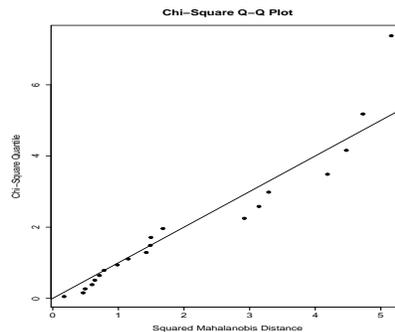


Figura 5 Q-Q plot fornecido pelo teste de Henze-Zirkler

A função do teste proposto (TNMt), baseado na transformação de Liang *et al.* (2004), foi implementada e o código está disponível no apêndice. Para sua utilização é necessário carregar a função `NMt.test` e sua aplicação é dada por:

```
X <- cbind(X1,X2)
NMt.test(X, tplot = TRUE)
```

Tem-se a seguinte saída do programa *R*:

```
One-sample Kolmogorov-Smirnov test
```

```

data: t
D^+ = 0.27547, p-value = 0.06222
alternative hypothesis: the CDF of x lies above the null
hypothesis

```

Na Figura 6 tem-se o t -plot. Pelo gráfico pode-se observar que há indício de não normalidade multivariada dos dados.

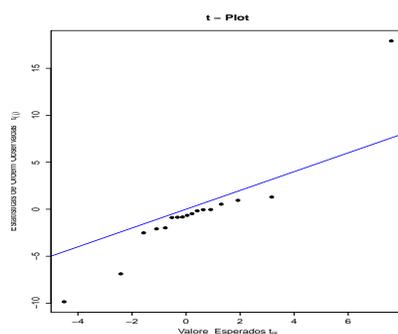


Figura 6 t-plot fornecido pelo teste proposto

Na Tabela 17 encontram-se as estatísticas de cada teste e seus respectivos valores- p obtidos através do programa R .

Tabela 17 Resultados dos testes de normalidade multivariados fornecidos pelo programa R .

Teste Normalidade Multivariado	Estatística do Teste	Valor- p
Shapiro-Wilk	0,97067	0,8345
TMCNMD	0,93840	0,2753
TNMBP	0,97202	0,8331
Doornik-Hansen	10,49981	0,0328
Henze-Zirkler	0,36243	0,5505
TNMt	0,27547	0,0622

Considerando, o nível de significância de 5%, todos os testes, exceto o de Doornik-Hansen, foram não significativos (valor- p maior que 5%). Assim, a hipótese de normalidade multivariada (bivariada) não deve ser rejeitada. Entretanto, verificou-se que o teste de Doornik-Hansen de normalidade multivariada implementado no programa R resultou em uma conclusão oposta à dos demais testes,

considerando o nível nominal de significância de 5%. O teste proposto (TNMt), apresentou resultados concordantes com a maioria dos testes considerados na aplicação, embora, se fosse considerado o nível nominal de 10% de significância, ele, juntamente com o teste de Doornik-Hansen, teria também resultados discrepantes dos demais testes considerados. Esse resultado é até surpreendente, haja vista, que os resultados simulados do teste proposto, em geral, mostraram que o teste é menos poderoso que os demais, o que contrasta com os resultados da presente situação.

5 CONCLUSÃO

Os testes de normalidade, univariada e multivariada são importantes, visto que a normalidade é uma premissa necessária para vários modelos estatísticos usados na análise de dados. Com isso, é necessário que se busque um teste exato e mais poderoso, uma vez que não existe na literatura um teste uniformemente poderoso para verificar a normalidade de dados multivariados.

Neste trabalho foi proposto um teste de normalidade multivariada que é baseado na transformação de Liang *et al.* (2004) para a distribuição *t* de Student. O desempenho do teste proposto foi validado por simulação Monte Carlo e observou-se que ele teve grande sucesso no controle das taxas de erro tipo *I*, possibilitando concluir, assim, que o TNMt é um teste exato. Porém, o TMNt foi analisado como sendo um teste pouco poderoso por apresentar o desempenho baixo em relação ao poder do teste, principalmente em amostras pequenas. Junto com o novo teste proposto também foi realizada a construção do gráfico *t*-plot para auxiliar na detecção de possíveis dados discrepantes. Infelizmente, o teste proposto teve poder muito baixo relativamente aos seus concorrentes e não deve ser recomendado. Embora o resultado seja negativo, considera-se o trabalho realizado como importante, pois cumpre a finalidade primordial de apresentar os resultados para a comunidade científica. Dessa forma, outros pesquisadores não teriam o trabalho e evitariam todos os esforços aqui empregados para descobrir que o procedimento não é eficiente. Ademais, o presente trabalho permite que pesquisadores de outras áreas, interessados em um procedimento gráfico de avaliação de normalidade utilizem com muita restrição os *t*-plots. Esse tipo de gráfico pode ser usado como um indicativo de não-normalidade, mas nunca como de normalidade, haja vista o baixo poder do teste proposto.

Recomenda-se a utilização do teste proposto por Royston (1983), caso a amostra esteja entre 3 a 5000, para se verificar a normalidade multivariada de uma amostra aleatória. Caso o tamanho da amostra seja superior a 5000 é aconselhável o uso do teste proposto por Biase (2011), baseado em *bootstrap* paramétrico (TN-MNP). Além disso, este último teste pode ser ainda usado em situações em que $p \geq n$.

REFERÊNCIAS

BASTOS, R. L. **Proposição de testes bootstrap para o índice de qualidade sensorial**. Dissertação de mestrado em Estatística e Experimentação Agropecuária. Universidade Federal de Lavras - UFLA. Lavras-MG, 2013.

BIASE, A. G. **Proposição de testes computacionalmente intensivos de normalidade multivariada**. Dissertação de mestrado em Estatística e Experimentação Agropecuária. Universidade Federal de Lavras - UFLA. Lavras-MG, 2011.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística básica**. 5ª ed. São Paulo. Editora Saraiva, 2004.

BOLFARINE, H.; SANDOVAL, M.C. **Introdução à inferência estatística**. 2ª ed. SBM. 2010.

CASELLA G., BERGER, R.L. **Statistical inference**. 2ª ed. Pacific Grove: Duxbury/Thomson Learning. 2001.

CIRILLO, M. A.; FERREIRA, D. F. **Extensão do teste para normalidade univariado baseado no coeficiente de correlação quantil-quantil para o caso multivariado**. Revista de Matemática e Estatística, São Paulo, v. 21, n. 3, 2003.

DACHS, J.N. **Estatística computacional: uma introdução em turbo pascal**. Rio de Janeiro: Livros Técnicos e Científicos, 1988.

D'AGOSTINO, R. B. **Transformation to normality of the null distribution of g_1** . Biometrika, London, v. 57, p. 679-681, 1970.

DOORNIK, J. A.; HANSEN, D. **An omnibus test for univariate and multivari-**

ate normality. Working Paper, Oxford: Nuffield College, 1994.

FANG, K.T.; KOTZ, S.; NG, K.W. **Symmetric multivariate and related distributions.** Chapman & Hall, London, New York. 1990.

FARRELL, P. J.; BARRERA, M. S.; NACZK, K. **On tests for multivariate normality and association simulation studies.** Journal of Statistical Computation and Simulation, New York, v. 77, n. 12, 2007.

FERREIRA, D. F. **Estatística básica.** 2ª ed. Editora: UFLA. Lavras-MG. 2005.

FERREIRA, D. F. **Estatística computacional usando R.** Universidade Federal de Lavras, 2010. Disponível em: <<http://www.dex.ufla.br/danielff/dex523.htm>>. Acesso em: 11 jul. 2015.

FERREIRA, D. F. **Estatística multivariada.** 2ª ed. Editora: UFLA. Lavras-MG. 2011.

HENZE, N.; ZIRKLER, B. **A class of invariant consistent tests for multivariate normality.** Communications in Statistics: Theory and Methods, New York, v. 19, n. 10, 1990.

KORKMAZ, S.; GOKSULUK, D.; ZARARSIZ, G. **MVN: An R package for assessing multivariate normality, 2015** MVN version 4.0. Disponível em: <<https://cran.r-project.org/web/packages/MVN/vignettes/MVN.pdf>>. Acesso em: 24 ago. 2015.

LANDAU, D. P.; BINDER, K. **A guide to Monte-Carlo simulations in statistical physics.** New York: Cambridge university press, v.1, 3ª edição, 2009.

LIANG, J.; LI, R., FANG, H.; FANG, K.T. **Testing multinormality based on low-dimensional projection.** J. Statist. Plann. Inference 86, 129-141. 2000.

LIANG, J.; PAN, W.S.Y.; YANG, Z.H. **Characterization-based Q-Q plots for testing multinormality.** Statistics & Probability Letters 70, 183-190. 2004.

MARDIA, K. V. **Measures of multivariate skewness and kurtosis with applications.** Biometrika, London, v. 57, n. 3, p. 519-530, Dec. 1970.

MARSAGLIA G.; TSANG W.W.; WANG J. **Evaluating Kolmogorov's distribution.** Journal of Statistical Software, Vol. 8, Issue 18, 1-4. 2003. Disponível em: <<http://www.jstatsoft.org/v08/i18/>>. Acesso em: 12 jan. 2015.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics.** 3ª ed. Singapore: McGraw-Hill, 1974.

PEREIRA FILHO, N. de A. **Teste Monte Carlo de normalidade univariado.** Tese de doutorado em Estatística e Experimentação Agropecuária. Universidade Federal de Lavras - UFLA. Lavras-MG, 2013.

PORTAL ACTION. **Ambiente virtual de aprendizado.** Disponível em: <<http://www.portalaction.com.br/inferencia/513-poder-do-teste>>. Acesso em: 10 jul. 2015.

R CORE TEAM. **R: A Language and Environment for Statistical Computing.** R Foundation for Statistical Computing. Vienna, Austria. 2015. Disponível em: <<http://www.r-project.org/>>. Acesso em: 10 jan. 2015.

ROYSTON, J. B. **Some techniques for assessing multivariate based on the Shapiro-Wilk W.** Applied Statistics, London, v. 32, n. 2, 1983.

ROYSTON, J. B. **Approximating the Shapiro-Wilk W-Test for non-normality.** Statistics and Computing, London v. 2, n. 1, 1992.

ROYSTON, J. P. **A toolkit for testing for non-normality in complete and censored samples.** The Statistician, London, v. 42, n. 1, 1993.

SHAPIRO, S. S.; WILK, M. B. **An analysis of variance test for normality.** Journal Biometrika, London, v. 52, n. 3/4, p. 591-611, Nov. 1965.

SILVA, R. B. V. **Extensão do teste de normalidade de Shapiro-Francia para o caso multivariado.** Tese de Doutorado em Estatística e Experimentação Agropecuária. Universidade Federal de Lavras - UFLA. Lavras-MG, 2009.

THODE JÚNIOR, H. C. **Testing for Normality.** New York: Marcel Dekker, 2002.

YANG, Z.H.; FANG, K.T.; LIANG, J. **A characterization of multivariate normal distribution and its application.** Statist. Probab. Lett. 30, 347-352. 1996.

ANEXO - Comandos usadas no R

```
##-----##
#      Função para gerar dados da normal contaminada      #
##-----##
mvrnormCT <- function(n,mu1, Sigma1, mu2, Sigma2,delta)
{
  library(mvtnorm)
  u <- runif(n)
  if(u[1] <= delta) X <- mvrnorm(1, mu1, Sigma1) else
    X <- mvrnorm (1, mu2, Sigma2)
  for (i in 2:n)
  {
    if(u[i] <= delta) y <- mvrnorm(1,mu1,Sigma1 ) else
      y <- mvrnorm(1, mu2, Sigma2)
    X <- rbind(X, y)
  }
  return(X)
}

##-----##
# Implementação da rotina realizada por Silva (2009)      #
# para o teste Shapiro-Wilk de normalidade multivariado  #
# proposto por Royston (1983). Em que x é a matriz de    #
# dados (n x p) e as funções auxiliares ki, fz, fe,      #
# grhon foram baseadas no artigo:                        #
# ROYSTON, J.P.A toolkit for testing for non-normality   #
# incomplete and censored samples. The Statistician,    #
# London, v.42, n.1, p.37-43, 1993.                    #
##-----##
Shapiro.Royston = function(x) {
  # auxiliar functions
  ki = function(zj) {
    ki = (qnorm(0.5*pnorm(-zj)))^2
    return(ki)
  }
  fz = function(wj, n){
```

```

if(n <= 11) {
  lamb = -2.273+0.459*n
  u = log(1-wj)
  y = -log (lamb-u)
  muy = 0.544-0.39978*n+0.025054*n^2-0.0006714*n^3
  sigy = exp(1.3822-0.77857*n+0.062767*n^2-
             0.0020322*n^3)
} else {
  y = log(1-wj )
  xx = log(n)
  muy = -1.5861-0.31082*xx-0.083751*xx^2+0.0038915*xx^3
  sigy = exp(-0.4803-0.082676*xx + 0.0030302*xx^2)
}
zc = (y-muy)/sigy
return(zc)
}
fe = function(cbar, p) {
  e = p/(1+(p-1)*cbar)
  return(e)
}
grhon = function(rho, n) {
  mu = 0.715
  lamb = 5
  xx = log(n)
  nu = 0.21364 + 0.015124*xx^2 - 0.0018034*xx^3
  #nu =0.35
  #print(nu)
  grijn = rho^lamb*(1-mu/nu*rho*(1-rho)^mu)
  grijn
}
# main procedures
if(!is.matrix(x))
  stop("x[] is not a matrix with number of rows
       (sample size) between 3 and 5000")
p = ncol(x); p
n = nrow(x); n
if(n < 3 || n > 5000)

```

```

    stop("sample size must be between 3 and 5000" )
rx = cor(x)
# Calculating the univariate W statistics
# using R function
Wis = matrix (0, p, 1)
for( i in 1:p ) {
  xj = x[1:n,i]
  wi = shapiro.test(xj)
  Wis[i,1] = wi $ statistic
}
# calculating Cij matrix
Cij = matrix (1, p, p)
if(p > 1){
  pp =1: p-1
  # print(pp)
  for(i in pp){
    ii = i+1
    seq = ii:p
    for(j in seq) {
      # print(i); print(j); print(seq)
      cij = grhon(abs(rx[i,j]), n)
      Cij[i,j] = cij
      # print(cij)
      Cij[j,i] = cij
    }
  }
}
# test statistics and p-value
if(p > 1) cbar = sum(Cij-diag(p))/(p*(p-1)) else
  cbar = 0
e = fe(cbar, p)
Kis = ki(fz(Wis, n))
G = sum(Kis)/p
W.mean = mean(Wis)
H=e*G
prH = 1-pchisq(H,e)
return(list(W=W.mean, G=G, H=H, dfeq=e, p.value=prH))

```

```

}

##-----##
# Teste Monte Carlo de normalidade multivariada baseado #
# em distâncias, proposto por Biase(2011) entrar com: #
# X: matriz de dados n x p e com #
# NSM: número de simulações Monte Carlo #
##-----##
testNMMC <- function (X, NSM = 2000)
{
  n <- nrow(X); p <- ncol(X)
  Xb <- apply(X, 2, mean)
  S <- var(X); SI <- solve(S)
  D2 <- function(x,med,SI)
  {
    return(t(x - med)\%*\%SI\%*\%(x - med))
  }
  D2j <- apply(X, 1, D2, Xb, SI)
  D2j <- sort(D2j)
  bj <- n*D2j/(n - 1)^2
  a <- (p - 2)/(2*p)
  b <- (n - p - 3)/(2 * (n - p - 1))
  pj <- ((1:n) - a)/(n - a - b + 1)
  alpha <- p/2
  beta <- (n - p - 1)/2
  bjs <- qbeta(pj, alpha, beta)
  r2c <- cor(bj, bjs)^2
  r2MC <- r2c
  for (i in 1:NSM)
  {
    Z <- matrix(rnorm(n*p), n, p)
    Zb <- apply(Z, 2, mean)
    Sz <- var(Z)
    SIz <- solve(Sz)
    D2jz <- apply(Z, 1, D2, Zb, SIz)
    D2jz <- sort(D2jz)
    bjz <- n*D2jz/(n - 1)^2
  }
}

```

```

    r2MC <- c(r2MC, cor(bjz, bjs)^2)
  }
  plot(bjs, bj, xlab = paste("Valores Esperados beta",
    sep = ""), ylab = paste("Estatísticas de Ordem
    Observadas", sep = " "))
  reg <- lm(bj ~ bjs)
  abline(reg$coefficients[1], reg$coefficients [2])
  valor.p <- length (r2MC[r2MC <= r2c])/(NSM + 1)
  return(list(r2 = r2c, valor.p = valor.p))
}

##-----##
# Teste de normalidade multivariada baseado em bootstrap #
# paramétrico proposto por Biase(2011) entrar com:      #
# X: matriz de dados n x p e com                        #
# NSM: número de replicações do bootstrap paramétrico  #
##-----##
testHNMMC <- function(X, NSM = 2000)
{
  n <- nrow(X); p <- ncol(X); S <- var(X)
  pj <- ((1 : n) - 3/8)/(n + 1/4)
  mj <- qnorm(pj)
  X <- apply(X, 2, sort)
  Hc <- sum(apply(X, 2, cor, mj)^2)/p
  H <- Hc
  library(MASS)
  for (i in 1:NSM)
  {
    Z <- mvrnorm(n, rep(0, p), S)
    Z <- apply(Z, 2, sort)
    HMC <- sum(apply(Z, 2, cor, mj)^2)/p
    H <- c(H, HMC)
  }
  valor.p <- length(H[H <= Hc])/(NSM + 1)
  return(list(H = Hc, valor.p = valor.p))
}

```

APÊNDICE - Comandos usadas no R

```
##-----##
#  Teste de normalidade multivariada baseado na transf.  #
#  de Liang et al. (2004) para a distribuição t-Student, #
#  em que:                                             #
#  X: matriz de dados de dimensões n x p             #
#  tplot: se colocar TRUE cria o gráfico t-plot.     #
##-----##

NMt.test <- function(X, tplot = FALSE)
{
  n <- nrow(X)
  p <- ncol(X)
  i <- 1:(n-1)
  U <- (apply(X,2,cumsum)[i,] - i*X[i+1,])/(i*(i+1))^0.5
  ki <- p + 1
  k <- 1
  sk <- U[k,]%*%t(U[k,])
  for (k in 2:ki) sk <- sk + U[k,]%*%t(U[k,])
  skinroot <- t(solve(chol(sk)))
  Y <- t(skinroot %*% U[ki,])
  for (k in (ki+1):(n-1))
  {
    sk <- sk + U[k,]%*%t(U[k,])
    skinroot <- t(solve(chol(sk)))
    Y <- rbind(Y,t(skinroot %*% U[k,]))
  }
  t <- p^0.5*apply(Y[1:(n-p-1)],,1,mean)/
    diag(var(t(Y)))^0.5
  # construção do t-plot
  if(tplot){
    nu <- p - 1
    nn <- n - p - 1
    tj <- sort(t)
    tq <- qt(((1:nn)+1/4)/(nn+1),nu)
    plot(tq,tj, pch = 19, main = "t - Plot",
```

```

        xlab = expression(paste('Valore Esperados',
            t[qj])),
        ylab = expression(paste('Estatísticas de
            Ordem Observadas', t[(j)])))
abline(0, 1, lwd=2, col="blue")
}
# pacote LambertW se df > 2
install.packages("LambertW")
library(LambertW)
if (p-1>2)
{
    beta.true <- c(location = 0, scale = 1, df= p - 1)
    ks <- ks.test.t(t,beta.true)
} else ks <- ks.test(t,"pt",p-1,alternative ="greater")
return(ks)
}

##-----##
# Função para Validação do teste proposto #
# SMC = Simulacao Monte Carlo #
# dist = 1: Normal Multivariada #
# dist = 2: t multivariada 1 GL #
# dist = 3: t multivariada 30 GL #
# dist = 4: Log-normal multivariada #
# dist = 5: Uniforme Multivariada #
# dist = 6: Normais Multivariadas Contaminada #
# N = Número de simulações Monte Carlo #
# Obs.: resultados com taxas de rejeições de H0 #
# para alpha = 10%, 5% e 1% #
##-----##
SMC <- function(n, p, N = 2000, dist = 1)
{
    library(MASS)
    library(mvtnorm)
    resultado <- matrix(0,1,3)
    rownames(resultado) <- c("TNMt")
    colnames(resultado) <- c("0.10", "0.05", "0.01")

```

```

mu                <- rep(0, times = p)
rho               <- 0.5
sigma2 <- 1
Sigma  <- sigma2*((1-rho)*diag(p)+rho*matrix(1, p, p))
# valores referentes da normal contaminada
mu2                <- mu + 100
Sigma2             <- 100^(1/p) * Sigma
delta              <- 0.7
for(i in 1:N)
{
  if(dist == 1) X <- rmvnorm(n, mu, Sigma) else
  if(dist == 2) X <- rmvt(n, Sigma, 1, mu) else
  if(dist == 3) X <- rmvt(n, Sigma, 30, mu) else
  if(dist == 4) X <- exp(rmvnorm(n, mu, Sigma)) else
  if(dist == 5) X <- pnorm(rmvnorm(n, mu, Sigma)) else
  if(dist == 6) X <- mvrnormCT(n,mu,Sigma,mu2,Sigma2,delta)
  res <- TNMt(X)
  res
  if(res$p.value <= 0.10) resultado[1,1] <- resultado[1,1]+1/N
  if(res$p.value <= 0.05) resultado[1,2] <- resultado[1,2]+1/N
  if(res$p.value <= 0.01) resultado[1,3] <- resultado[1,3]+1/N
}
return(resultado)
}

##-----##
#                               Exemplo                               #
##-----##

# X1: concentração de hemoglobina e
# X2: contagem de linfócitos.
X1 <- c(13.4, 14.6, 13.5, 15.0, 14.6, 14.0, 16.4, 14.8,
        15.2, 15.5, 15.2, 16.9, 14.8, 16.2, 14.7, 14.7,
        16.5, 15.4, 15.1, 14.2)
X2 <- c(14,15,19,23,17,20,21,16,27,34,26,28,24,26,23,9,
        18,28,17,14)

```

```
dados <- data.frame(X1,X2)

# Teste de Shapiro-Wilk normalidade multivariado
# Carregar a função que se encontra no anexo
x <- cbind(X1,X2)
Shapiro.Royston(x)

# Teste Monte Carlo de normalidade multivariada
# baseado em distâncias (TMCNMD)
# Carregar a função que se encontra no anexo
X <- cbind(X1,X2)
testNMMC(X)

# Teste de normalidade multivariada baseado em
# bootstrap paramétrico
X <- cbind(X1,X2)
testHNMMC(X)

# Teste de Doornik-Hansen normalidade multivariado
install.packages("asbio")
library(asbio)
Y <- dados
DH.test(Y)

# Teste de Henze-Zirkler normalidade multivariado
install.packages("MVN")
library(MVN)
data <- dados
hzTest(data, cov = TRUE, qqplot = TRUE)

# Teste proposto (TNMt)
# Carregar a função que se encontra no começo
# do apêndice
X <- cbind(X1,X2)
NMt.test(X, tplot = TRUE)
```