Revista da Estatística UFOP, Vol III(3), 2014, ISSN 2237-8111
Edição especial: 59ª Reunião Anual da Regional Brasileira da Sociedade Internacional de Biometria - RBRAS
Trabalhos apresentados na versão poster.

AVALIAÇÃO DOS TESTES MULTIVARIADOS DA RAZÃO DE VEROSSIMILHANÇAS E T² DE HOTELLING: Um estudo por simulação de dados

Eduardo Campana Barbosa[12]
Rômulo César Manuli[2]
Patrícia Sousa[2]
Ana Carolina Campana Nascimento[2]
Moysés Nascimento[2]

1 Introdução

Quando um teste de hipótese é aplicado, seja uni ou multivariado, espera-se obter resultados eficientes e confiáveis. O ideal seria que um teste rejeitasse em 0% das vezes a hipótese nula, dado que ela é verdadeira e em 100% das vezes quando falsa. Na prática isto é impossível, e portanto, ao aplicar um teste de hipótese o pesquisador precisa conhecer as situações em que o mesmo é ou não eficiente/confiável (CANTELMO; FERREIRA, 2007).

Segundo Fonseca e Martins (2009), a avaliação de um teste de hipótese é feita fixando-se um nível de significância α. Espera-se que a taxa do Erro Tipo I, que retrata a $P[\text{Rejeitar}H_0|H_0 \text{ é verdadeira}]$, seja igual ou inferior a este valor nominal. Se isto ocorre, o teste é considerado conservativo e caso contrário, liberal. Uma abordagem complementar envolve o Poder do Teste, ou seja, $P[\text{Rejeitar}H_0|H_0 \text{ é Falsa}]$. Espera-se que um teste apresente elevados valores de Poder (> 0,80).

A avaliação de testes de hipóteses é um tema muito abordado em estudos por simulação de dados, principalmente por permitir a composição de diferentes tipos de cenários/situações, sem a necessidade de se instalar um experimento, obtendo conclusões sobre os testes, minimizando custos, evitando desperdícios de matéria-prima, mão-de-obra e tempo. Alguns resultados interessantes podem ser verificados em Cantelmo e Ferreira (2007), Cirilo, Ferreira e Sáfadi (2006) e Barroso et al., (2012).

---

[1] Agradecimentos à Fapemig, Capes e CNPq pelo apoio financeiro.
[2] DET – UFV. email: duducampana@hotmail.com

Neste sentido, o presente trabalho tem por objetivo apresentar um estudo por simulação de dados com intuito de avaliar a taxa do Erro Tipo I e o Poder de dois testes estatísticos multivariados, o T^2 de Hotelling (T^2) e o teste da Razão de Verossimilhanças (TRV) baseado na aproximação assintótica sugerida por Mood et al. (1974).

2 Materiais e Métodos

A simulação de dados foi realizada no *software* livre R e utilizou-se o pacote MASS para habilitar a função “mvrnorm”, necessária para geração de amostras aleatórias com distribuição Normal Multivariada. Optou-se por simular uma distribuição normal bivariada (p=2), considerando a matriz de variâncias e covariâncias uma identidade de ordem 2 (independência e homogeneidade de variâncias).

Desenvolveu-se uma rotina para repetir o procedimento de teste 10.000 vezes, calculando o valor da estatística e comparando com o valor tabelado (obtido pela distribuição do teste, α e os graus de liberdade). Para avaliar a taxa do Erro Tipo I (sob H_0 verdadeira) foram simuladas amostras aleatórias e computou-se a proporção de rejeição da hipótese nula. O Poder foi avaliado sob hipótese nula falsa, simulando um vetor de médias superior em 0.5 desvios padrões, isto é, $\mu = [\mu_1 + 0.5\sigma_1,\ \mu_2 + 0.5\sigma_2]^T$.

Os cenários foram compostos conforme Barroso et al. (2012), porém, com os níveis de significância 1%, 5%, e 10% e tamanhos de amostra de n = 10, 20, 30, 40, 50, 100 e 500, totalizando 21 cenários diferentes.

2.1 Teste da Razão de Verossimilhança (TRV)

Utiliza-se o TRV quando se deseja testar a hipótese de que um parâmetro θ pertence a um subespaço de R^S, denominado como espaço restrito (Ω_0), ou a um espaço irrestrito (Ω). O espaço restrito corresponde à hipótese nula (H_0) enquanto o irrestrito a hipótese alternativa (H_a). A ideia do teste consiste em obter o máximo da função de verossimilhança para o espaço restrito e para o irrestrito, tomando a razão de verossimilhanças dada em (1), para obter uma regra de decisão sobre as hipóteses:

$$\Lambda = \frac{L_{\Omega_0}(Y;\hat{\theta})}{L_{\Omega}(Y;\hat{\theta})} \tag{1}$$

Por (1), se Λ é grande, provavelmente o espaço restrito conterá o parâmetro θ, e o contrário ocorre para Λ pequeno. Estabelecer uma região de rejeição para H_0, com base em probabilidades, é algo complexo e trabalhoso, pois seria preciso obter a distribuição de Λ.

Neste sentido, Mood et al. (1974) sugerem a aproximação assintótica de -2ln(Λ), que possui distribuição Qui-Quadrado com R – S graus de liberdade[3]. Conforme Ferreira (2011), a estatística de teste para vetores de médias de uma população e matriz de variâncias e covariâncias conhecidas, é dada em (2). Em (3) temos o resultado quando não se conhece a matriz de variâncias e covariâncias, e trabalha-se com a sua EMV (S_n), obtida com base em dados amostrais.

$$-2\ln(\Lambda) = n(\overline{Y}_{.} - \mu_0)^T \Sigma^{-1} (\overline{Y}_{.} - \mu_0) \tag{2}$$

$$-2\ln(\Lambda) = n.\ln\left[1 + (\overline{Y}_{.} - \mu_0)^T {S_n}^{-1} (\overline{Y}_{.} - \mu_0)\right] \tag{3}$$

2.2 T² de Hotelling (T²)

Segundo Ferreira (2011), sob hipótese nula verdadeira e observações amostrais com distribuição normal multivariada, independência e covariâncias homogêneas, a estatística de teste T², dada em (4), é exata e possui aproximadamente distribuição F central com p e v +1 – p graus de liberdade.

$$T^2 = n(\overline{Y}_{.} - \mu_0)^T S^{-1} (\overline{Y}_{.} - \mu_0) \sim \frac{vp}{v+1-p} F_{\alpha;p;v+1-p} \tag{4}$$

O T² é obtido via modificação do TRV, visto que (3) é uma função monótona crescente de $(\overline{Y}_{.} - \mu_0)^T {S_n}^{-1} (\overline{Y}_{.} - \mu_0)$, então após algumas operações algébricas é possível provar este resultado.

3 Resultados e Discussão

A Tabela 1 apresenta os resultados obtidos na simulação de dados quando o objetivo foi verificar a capacidade de cada teste em controlar a taxa do Erro Tipo I.

Tabela 1: Resultados para a Taxa do Erro Tipo I

Teste T² Hotelling			
n/α	**0,01**	**0,05**	**0,10**
10	0,009	0,049	0,099
20	0,009	0,050	0,100
30	0,010	0,050	0,099
40	0,010	0,049	0,100
50	0,010	0,050	0,100

Teste Razão de Verossimilhança			
n/α	**0,01**	**0,05**	**0,10**
10	0,025	0,091	0,158
20	0,016	0,067	0,126
30	0,014	0,061	0,116
40	0,013	0,058	0,112
50	0,012	0,056	0,110

[3] R é o número de parâmetros estimados no espaço irrestrito e S no espaço restrito

100	0,010	0,049	0,100		**100**	0,011	0,053	0,105
500	0,009	0,050	0,100		**500**	0,010	0,050	0,100

Verifica-se que o T² apresentou resultados satisfatórios, mostrando-se conservativo sob as condições propostas, ou seja, apresentando taxa do Erro Tipo I igual ou inferior ao valor nominal estabelecido para α em todos os cenários. O TRV não apresentou resultados satisfatórios para ambos os níveis de significância e pequenos tamanhos de amostra (n = 10, 20 e 30), tendendo a superestimar a taxa do Erro Tipo I. Destaca-se que os valores decrescem lentamente e se aproximam precisamente do valor nominal apenas para amostras suficientemente grandes (n ≥ 500), o que já era esperado, devido às propriedades assintóticas do TRV. A tabela 2 apresenta os resultados obtidos para o Poder do teste.

Tabela 2: Resultados para o Poder do Teste

Teste T² Hotelling					**Teste Razão de Verossimilhança**			
n/α	**0,01**	**0,05**	**0,10**		**n/α**	**0,01**	**0,05**	**0,10**
10	0,130	0,364	0,517		**10**	0,241	0,495	0,632
20	0,466	0,742	0,846		**20**	0,544	0,789	0,876
30	0,751	0,917	0,960		**30**	0,790	0,932	0,966
40	0,905	0,977	0,991		**40**	0,919	0,981	0,992
50	0,968	0,994	0,998		**50**	0,973	0,995	0,998
100	1,000	1,000	1,000		**100**	1,000	1,000	1,000
500	1,000	1,000	1,000		**500**	1,000	1,000	1,000

Os dois testes não foram poderosos para pequenos tamanhos de amostra (n=10 e 20) e níveis de significância de 1%, 5% e 10%, com exceção em n = 20 e α = 10%, obtendo valores inferiores a 0,80. Para os demais tamanhos de amostra, com exceção do cenário em que n = 30 e α = 1%, os dois testes foram poderosos, apresentando valores iguais ou superiores a 0,80. Segundo Hair et al. (2010), quando o tamanho da amostra é pequeno o teste de hipótese é pouco sensível, o que torna difícil a detecção de efeitos significativos. Este autor afirma ainda que o aumento do tamanho da amostra implicará no aumento do Poder do teste, o que corrobora os resultados obtidos.

Ainda na Tabela 2, nota-se que o TRV mostrou-se superior ao T² em relação ao Poder. Porém, este não é um resultado real/confiável, visto que o TRV não controlou a taxa de Erro Tipo I para pequenos tamanhos de amostra, mostrando-se um teste liberal. Com o aumento de n os valores tendem a se tornar estatisticamente iguais, o que pode ser verificado, por exemplo, via intervalos de confiança. Cantelmo e Ferreira (2007) obtiveram resultados semelhantes a este quando compararam o teste de normalidade multivariado de Shapiro-Wilk com os testes de assimetria e curtose de Mardia. Esta mesma situação ocorreu em Cirilo, Ferreira e Sáfadi (2006), avaliando o teste multivariado de Levene por simulação Monte Carlo e abordagem *Bootstrap*.

4 Conclusões

Os resultados demonstraram que o T^2 obteve melhores resultados em relação ao TRV quando o objetivo foi controlar a taxa do Erro Tipo I, o que já era esperado devido as propriedades assintóticas do segundo. Quanto ao Poder do Teste, o TRV foi superior, inclusive para pequenos tamanhos de amostra, porém, este resultado não deve ser considerado, visto que o teste mostrou-se liberal e não foi capaz de controlar o Erro Tipo I principalmente nestes cenários. Para futuros trabalhos sugere-se que sejam incorporados novos cenários com a influência de diferentes matrizes de variâncias e covariâncias além do aumento do número de variáveis ($p > 2$).

Referências

[1] BARROSO, L. M. A.; NASCIMENTO, M.; SILVA, F. F.; NASCIMENTO, A. C. C.; PETERNELLI, L. A. Avaliação do teste generalizado de Durbin-Watson. **Revista Brasileira de Biometria**, v. 30, p. 432-441, 2012.

[2] CANTELMO, N. F.; FERREIRA, D. F. Desempenho de testes de normalidade multivariados avaliados por simulação monte carlo. **Ciência e Agrotecnologia**, Lavras, v. 31, n. 6, p. 1630-1636, nov./dez., 2007.

[3] CIRILO, M. A.; FERREIRA, D. F.; SÁFADI, T. Estudo do Poder e tamanho do teste de Lavene multivariado via simulação Monte Carlo e Booststrap. **Acta Scientiarum. Technology,** Maringa, v. 28, n.1, p. 105-112, 2006.

[4] FERREIRA, D. F. **Estatística Multivariada**. 2 ed. Lavras. Ed: UFLA, 2011.

[5] FONSECA, J. S.; MARTINS, G. A. **Curso de Estatística**. 6 ed. São Paulo: Atlas, 2009.

[6] HAIR JR, J.F.; BLACK, W.C.; BABIN, B. J.; ANDERSON, R. E. **Multivariate data analysis**. 7 ed. New Jersey: Prentice Hall. 2010.

[7] MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction of the theory of statistics**. New York. McGraw-Hill, 1974. 564 p.