

Revista Mirante, Anápolis (GO), v. 12, n. 2, dez. 2019. ISSN 19814089

**UM ESTUDO SOBRE O TAMANHO E PODER DO TESTE *t-STUDENT*
PARA A MÉDIA VIA SIMULAÇÃO DE MONTE CARLO**

**A STUDY ABOUT THE SIZE AND POWER OF *T-STUDENT* TEST FOR
THE MEAN THROUGH A MONTE CARLO SIMULATION**

CLEBER GIUGIOLI CARRASCO

Mestre em Estatística. Docente do Campus Anápolis de Ciências Exatas e Tecnológicas
Henrique Santillo, Universidade Estadual de Goiás - UEG
cleber.carrasco@ueg.br

THIAGO SANTANA LEMES

Graduado em Matemática. Campus Anápolis de Ciências Exatas e Tecnológicas Henrique
Santillo, Universidade Estadual de Goiás - UEG / Mestre e Doutorando em Engenharia
Elétrica e de Computação pela Universidade Federal de Goiás - UFG
thiago2608santana@gmail.com

Resumo: Neste artigo avaliou-se o tamanho e o poder de um teste de hipóteses para a média conhecido como teste *t-Student*, através de um estudo de simulação de Monte Carlo, o qual foi implementado no *software free R*. Concluiu-se que o tamanho do teste está próximo do nível de significância do teste para diferentes tamanhos amostrais. Em relação ao poder do teste, concluiu-se que na medida em que se aumenta o tamanho da amostra ou o nível de significância do teste, o poder do teste aumenta, isto também ocorre quando o valor da média se distancia do valor fixado na hipótese nula. Entretanto, quanto menor é a variância dos dados, maior é o poder do teste.

Palavras-chave: Inferência estatística. Nível de significância. Teste de hipóteses.

Abstract: In this paper we evaluated the empirical size and power of the hypothesis testing for the mean known as the *t-Student* test, through a Monte Carlo simulation study, which was implemented in the software free *R*. We concluded that the empirical size of the test is close of the level of significance of the test for different sample sizes. In relation to power of the test, we concluded that the extent that increases the sample size or the level of significance the test, the power of the test increases, this is also the case when value the mean differs from the value determined in the null hypothesis. However, the greater is the variance of the data, the is lesser of the power of the test.

Keywords: Statistical inference. Level of significance. Hypothesis testing.

INTRODUÇÃO

Na inferência estatística quando se utiliza um teste de hipóteses podem ocorrer dois tipos de erros: o erro tipo I, que ocorre quando a hipótese nula é rejeitada sendo ela verdadeira, e o erro tipo II, que ocorre quando a hipótese nula não é rejeitada sendo ela falsa (CASSELA; BERGER, 2010). A probabilidade do erro tipo I, definida por α , é chamada de nível de significância ou tamanho do teste, ou seja, é a probabilidade de o teste rejeitar incorretamente a hipótese nula. A probabilidade do erro tipo II é simbolizada por β e está relacionada com o poder do teste, definido por $1 - \beta$ (MAGALHÃES; LIMA, 2010). O poder

de um teste é a probabilidade de rejeitar a hipótese nula quando ela é falsa, ou seja, a probabilidade do teste rejeitar corretamente a hipótese nula (BARROS; MAZUCHELI, 2005).

Na prática deve-se procurar um equilíbrio entre as probabilidades desses erros para que o tamanho e o poder do teste de hipóteses tenham confianças aceitáveis, uma vez que não se podem controlar ambos os erros simultaneamente, ao menos que se aumente o tamanho da amostra (CARRASCO; LEMES, 2014).

Em geral, ao realizar um teste de hipóteses controla-se o erro tipo I, ou seja, o tamanho do teste, fixando a priori o valor do nível de significância desse teste. Entretanto, a probabilidade do erro tipo II e consequentemente o poder do teste não é fixado, ficando ambos sem controle, pois o mesmo requer um procedimento mais complexo para sua determinação (BUSSAB; MORETTIN, 2011).

Diante disso, alguns autores avaliaram o poder de um teste de hipóteses para a média mostrando que esse pode ser afetado pelo tamanho da amostra, nível de significância do teste, variabilidade e pela distância entre o valor real do parâmetro e o valor a ser testado. Para isso usaram a função poder (BUSSAB; MORETTIN, 2011) ou a curva característica de operação (COSTA NETO, 1977) para chegar a esses resultados. Carrasco e Silva (2009) avaliaram o erro tipo I e posteriormente, Carrasco e Silva (2013) avaliaram o erro tipo II, ambos na aplicação de um teste de hipóteses para a média com variância conhecida, através de um estudo de simulação. Já Carrasco e Lemes (2014) avaliaram somente o erro tipo II para um teste de hipóteses para a média com variância desconhecida.

Dessa forma, o objetivo deste trabalho é realizar um estudo de simulação utilizando o método de Monte Carlo, o qual consiste em repetir o mesmo procedimento diversas vezes (MUN, 2006), para verificar a influência do tamanho amostral, nível de significância do teste, variabilidade dos dados no tamanho e poder do teste de hipóteses para a média com variância da população desconhecida, conhecido como teste *t-Student* (MAGALHÃES; LIMA, 2010). Este estudo propicia uma percepção da dimensão do tamanho e do poder do teste *t-Student* em termos quantitativos. Para o estudo do poder do teste também é considerado o afastamento do valor do parâmetro em relação ao fixado na hipótese nula.

Para realizar esse estudo de simulação será construído um algoritmo para os cálculos necessários que será implementado computacionalmente no *software R* (VENABLES; SMITH; R CORE TEAM, 2019). A escolha pelo *software R* se dá pela facilidade na implementação de novas rotinas, além do mesmo ser gratuito e de fácil acesso (PETERNELLI; MELO, 2007).

MATERIAL E MÉTODOS

Para realizar o estudo de simulação de Monte Carlo, considerou-se o teste de hipóteses para a média com variância desconhecida, ou seja, o teste *t-Student* para uma amostra, tanto bilateral quanto unilateral à esquerda e à direita. Para o teste bilateral a hipótese nula (H_0) e a hipótese alternativa (H_A) serão definidas para testar $H_0: \mu = \mu_0$ contra $H_A: \mu \neq \mu_0$, já para os testes unilaterais à esquerda e à direita serão definidas, respectivamente, para testar $H_0: \mu = \mu_0$ contra $H_A: \mu < \mu_0$ e $H_0: \mu = \mu_0$ contra $H_A: \mu > \mu_0$. Serão geradas no *software R* amostras de tamanhos variados de uma distribuição normal com média e variância pré-definidas e fixado o nível de significância do teste (α) em alguns valores estabelecidos. Utilizar-se-á o método de Monte Carlo, repetindo os procedimentos adotados r vezes e, em cada repetição, serão efetuados os cálculos para estimar o tamanho e o poder do teste *t-Student*.

Tamanho do teste

O tamanho do teste é definido como sendo a probabilidade de o teste de hipóteses rejeitar incorretamente a hipótese nula, ou seja, a probabilidade de o teste de hipóteses rejeitar a hipótese H_0 sendo essa verdadeira.

Para calcular o tamanho do teste *t-Student* será gerado uma amostra de tamanho n de uma distribuição normal com média $\mu = \mu_0$, ou seja, sob a hipótese nula e, variância fixada em σ^2 . Em seguida, aplica-se o teste *t-Student* ao nível de significância de $\alpha\%$ e verifica se o teste rejeitou ou não a hipótese nula. Através do método de Monte Carlo esse procedimento será repetido r vezes, e para cada repetição será construído um contador que atribuirá $\delta_i = 1$ se o teste rejeitar H_0 e $\delta_i = 0$ se o teste não rejeitar H_0 , $i = 1, 2, \dots, r$. Portanto o tamanho do teste será calculado através da seguinte equação (CARRASCO; SILVA, 2009, p. 66):

$$\alpha^* = \frac{\sum_{i=1}^r \delta_i}{r}, i = 1, 2, 3, \dots, r. \quad (1)$$

Neste estudo de simulação para avaliar o tamanho do teste *t-Student* serão geradas amostras de seis tamanhos distintos e utilizados três diferentes níveis de significância para o teste. Com relação a variância, essa será fixada em σ^2 , uma vez que valores distintos para a

variância não alteram o cálculo do tamanho do teste (CARRASCO, SILVA, 2009). Ao todo serão 18 combinações entre as amostras e os níveis de significância, como exhibe a tabela 1.

Tabela 1: Combinações para o cálculo do tamanho do teste.

α	Tamanho da Amostra (n)					
	n_1	n_2	n_3	n_4	n_5	n_6
α_1	α_{11}^*	α_{12}^*	α_{13}^*	α_{14}^*	α_{15}^*	α_{16}^*
α_2	α_{21}^*	α_{22}^*	α_{23}^*	α_{24}^*	α_{25}^*	α_{26}^*
α_3	α_{31}^*	α_{32}^*	α_{33}^*	α_{34}^*	α_{35}^*	α_{36}^*

Fonte: Autores, 2019.

Poder do Teste

O poder de um teste é a probabilidade do teste rejeitar corretamente a hipótese nula, ou seja, a probabilidade de rejeitar a hipótese nula quando ela é falsa (BARROS; MAZUCHELI, 2005).

Para o cálculo do poder do teste *t-Student* será usado um procedimento análogo ao utilizado para o tamanho do teste, o qual consiste em gerar uma amostra de tamanho n de uma distribuição normal com média $\mu \neq \mu_0$, ou seja, sob a hipótese alternativa e , variância fixada em σ^2 . Em seguida, aplica-se o teste *t-Student* ao nível de significância de $\alpha\%$ e verifica se o teste rejeitou ou não a hipótese nula. Através do método de Monte Carlo repetir-se-á esse procedimento r vezes, e para cada repetição será construído um contador que atribuirá $\delta_i = 1$ se o teste rejeitar H_0 e $\delta_i = 0$ se o teste não rejeitar H_0 , $i = 1, 2, \dots, r$. Portanto o poder do teste pode ser calculado através da seguinte equação adaptada de Carrasco e Lemes (2014, p. 12):

$$1 - \beta^* = \frac{\sum_{i=1}^r \delta_i}{r}, i = 1, 2, 3, \dots, r. \quad (2)$$

Na avaliação do poder do teste *t-Student* será realizado um estudo de simulação onde serão geradas amostras de seis tamanhos distintos e utilizado seis valores diferentes para a média μ , três diferentes para a variância σ^2 e ainda, três valores distintos para o nível de significância do teste.

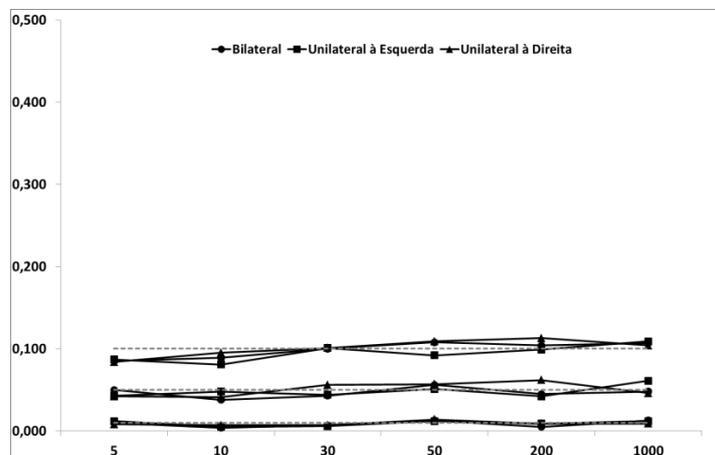
RESULTADOS E DISCUSSÃO

O procedimento descrito em material e métodos foi implementado no *software R* para calcular o tamanho e o poder do teste *t-Student*, onde a hipótese nula foi definida em $H_0: \mu = 5$ e as hipóteses alternativas em $H_A: \mu \neq 5$, $H_A: \mu < 5$ e $H_A: \mu > 5$, para os testes bilateral, unilateral à esquerda e a unilateral à direita, respectivamente. Utilizou-se amostras de tamanhos $n = 5, 10, 30, 50, 200$ e 1.000 . A escolha desses valores se deve ao fato de se considerar amostras pequenas, moderadas e grandes, neste trabalho. Os níveis de significância do teste foram definidos em $\alpha = 0,01; 0,05$ e $0,10$, os mais usuais na literatura estatística (CASSELLA; BERGER, 2010). Para o cálculo do tamanho do teste utilizou-se $\sigma^2 = 1$ e, para o poder do teste utilizou-se os valores $\sigma^2 = 0,0625; 1$ e 25 , buscando se ter baixa, média e alta variabilidade dos dados gerados. Ainda, no estudo para o poder do teste, foram geradas amostras de uma distribuição normal com média $\mu = 4,0; 4,5; 4,9; 5,1; 5,5$ e $6,0$ para os testes bilaterais; $\mu = 4,0; 4,5$ e $4,9$ para os testes unilaterais à esquerda e $\mu = 5,1; 5,5$ e $6,0$ para os testes unilaterais à direita. No método de Monte Carlo usou-se $r = 1.000$, e utilizou-se as equações (1) e (2) para estimar o tamanho e o poder do teste, respectivamente, para cada caso.

Análise do tamanho do teste

A figura 1 apresenta os resultados do tamanho dos testes *t-Student* obtidos através do estudo de simulação de Monte Carlo. As linhas tracejadas na figura representam os níveis de significância do teste, fixados em $\alpha = 0,01; 0,05$ e $0,10$, respectivamente. Pode-se observar que não houve diferença relevante entre o tamanho dos testes e os níveis de significância.

Figura 1: Resultados da simulação de Monte Carlo para o tamanho dos testes.



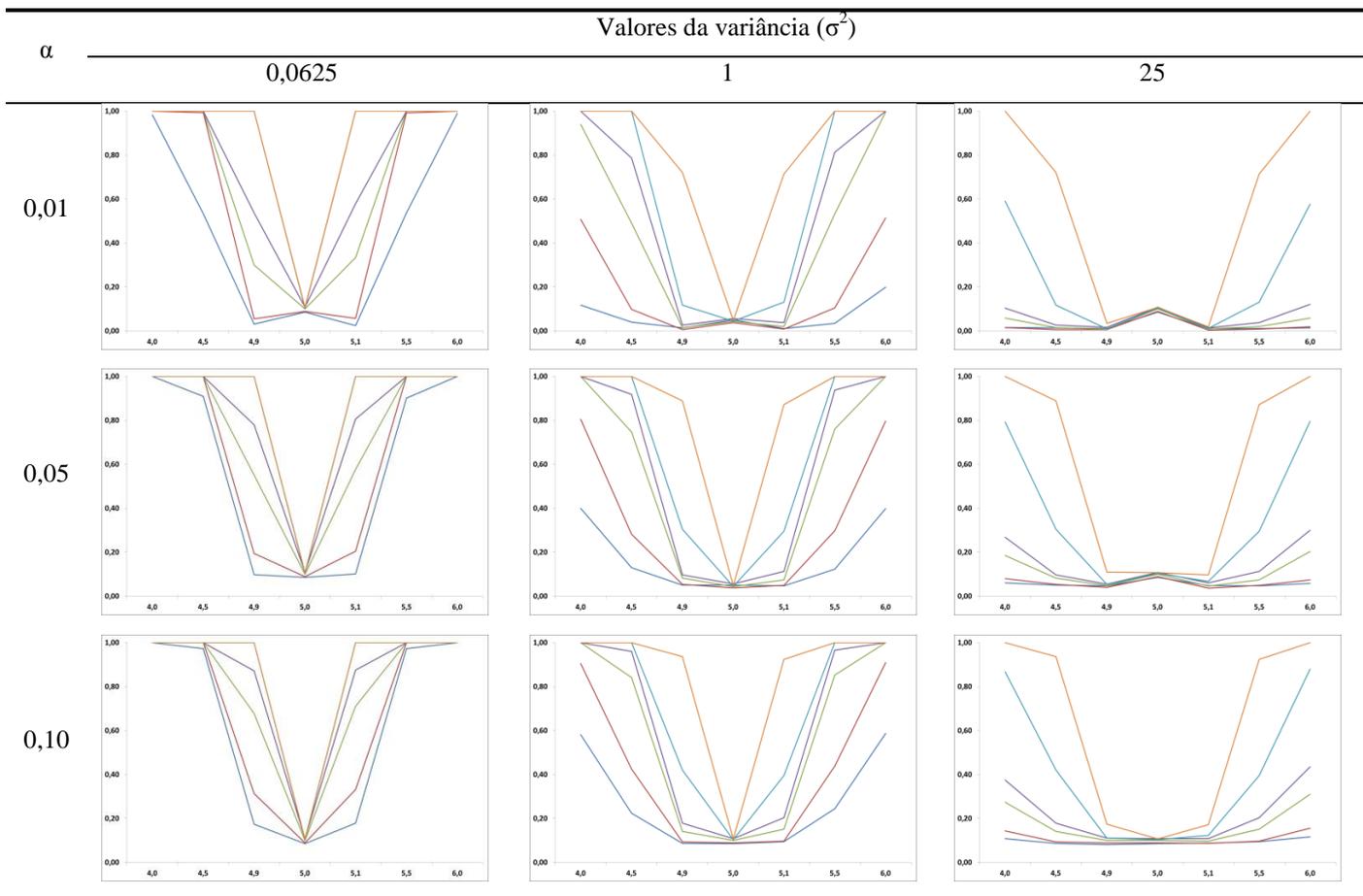
Fonte: Autores, 2019.

O tamanho amostral não influencia o resultado, visto que mesmo para amostras menores, os tamanhos dos testes mantiveram-se próximos de α . Esses resultados são observados tanto para o teste bilateral quanto para os testes unilaterais à direita e à esquerda.

Análise do poder do teste

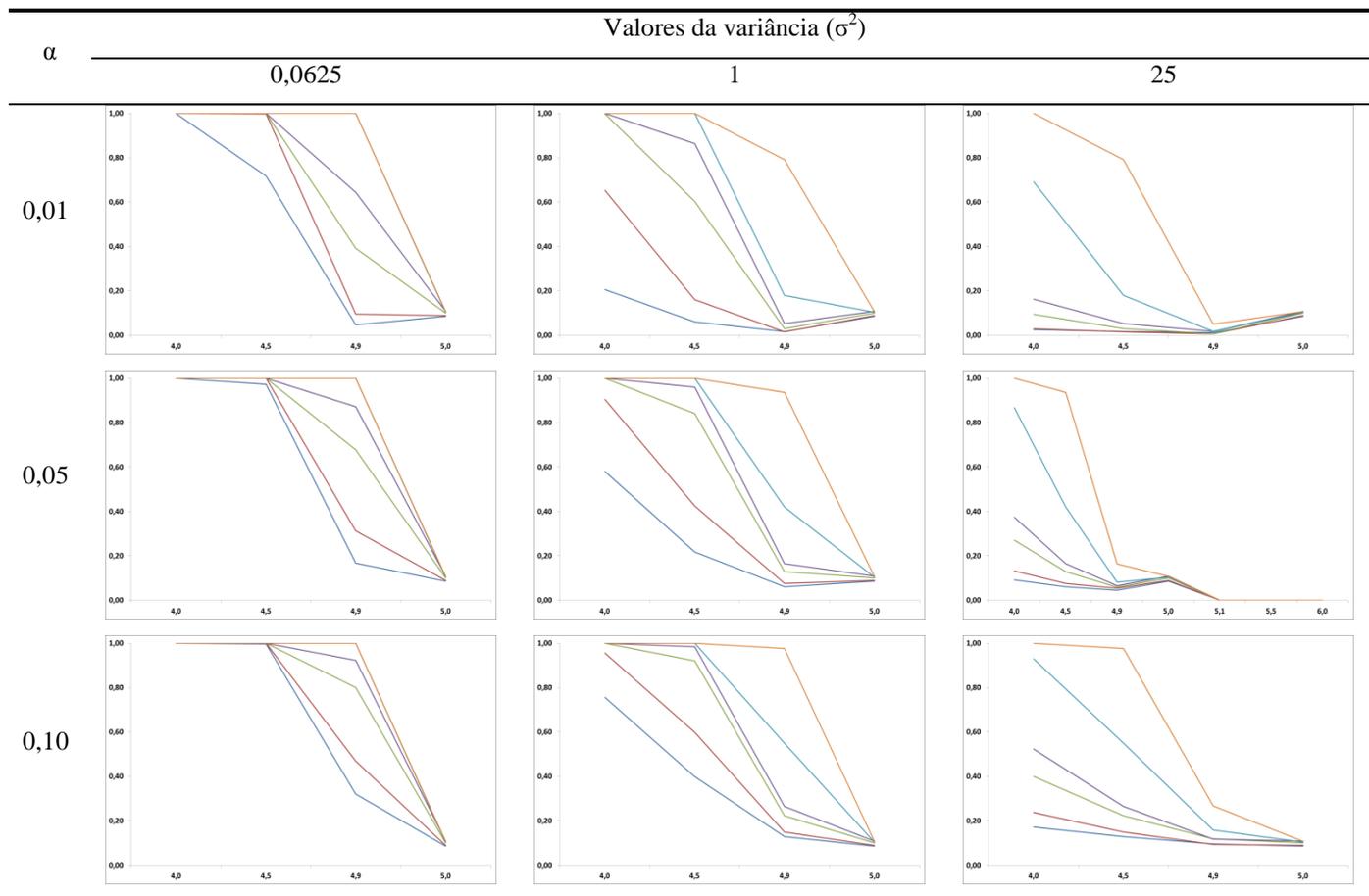
As tabelas 2, 3 e 4 apresentam através de gráficos os resultados do estudo de simulação de Monte Carlo para o poder dos testes *t-Student*, respectivamente para os testes bilateral, unilateral à esquerda e unilateral à direita. Nestes gráficos os tamanhos amostrais são representados por linhas de diferentes cores nas quais: (—) $n = 5$, (—) $n = 10$, (—) $n = 30$, (—) $n = 50$, (—) $n = 200$ e (—) $n = 1.000$.

Tabela 2: Resultados do poder do teste *t-Student* para o teste bilateral.



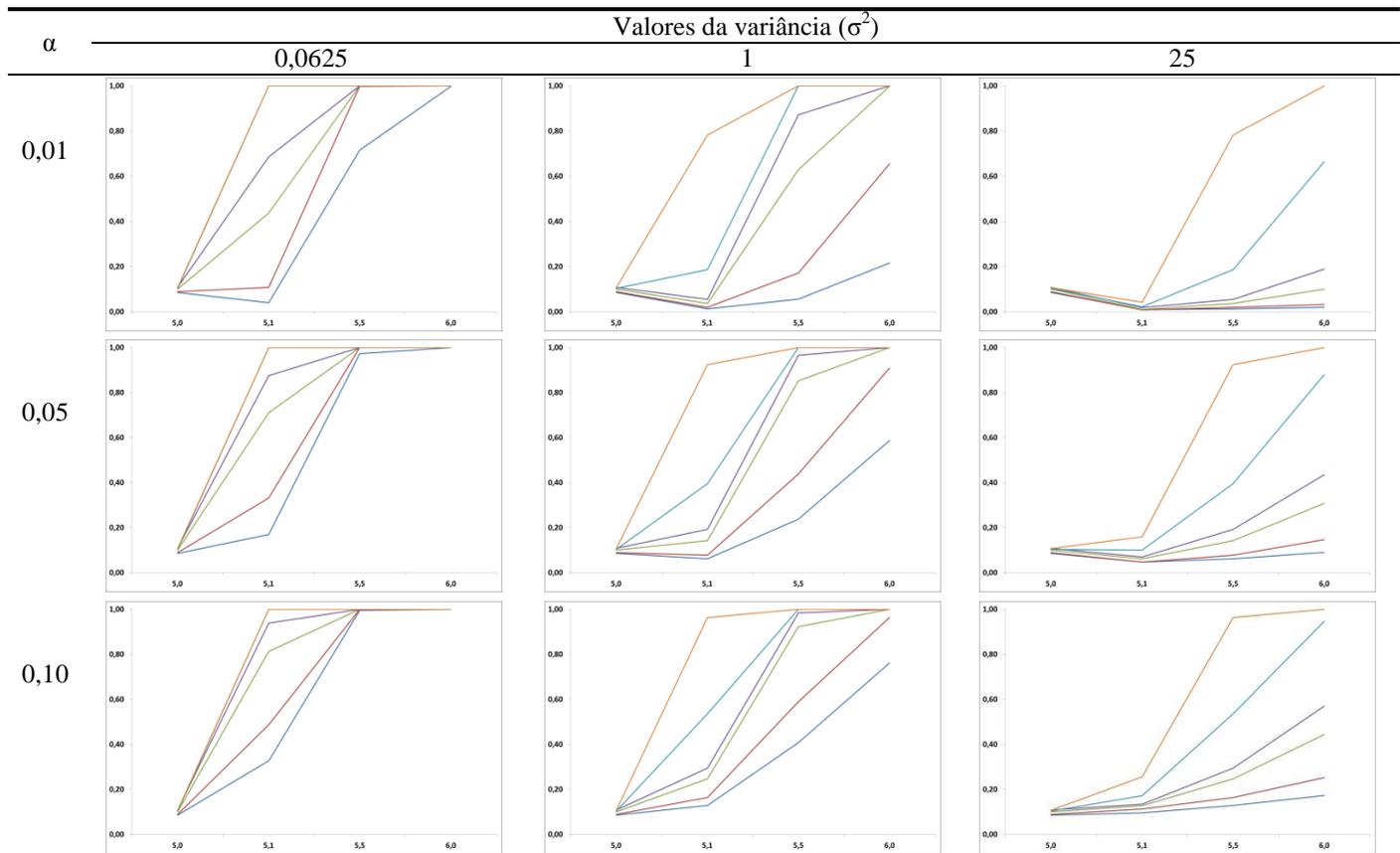
Fonte: Autores, 2019.

Tabela 3: Resultados do poder do teste *t-Student* para o teste unilateral à esquerda.



Fonte: Autores, 2019.

Tabela 4: Resultados do poder do teste *t-Student* para o teste unilateral à direita.



Fonte: Autores, 2019.

No estudo de simulação de Monte Carlo foi constatado que em relação aos valores atribuídos para a média μ , o poder aumenta na medida em que o valor da média é afastado da média sob a hipótese nula, neste caso, $\mu = 5$. Com relação ao tamanho amostral, é visto que o poder aumenta para o teste realizado em amostras maiores, ou seja, quanto maior a amostra maior será a probabilidade de se tomar uma decisão correta ao rejeitar a hipótese nula. Analisando a variabilidade dos dados foi constatado que o poder é maior quando a variância é menor, assim maior variabilidade implica em menor poder do teste. O nível de significância α também influencia o poder do teste, pois na medida em que se aumenta este, aumenta-se o poder do teste. Portanto, quanto maior for o poder expresso em cada uma dessas análises, maior será a probabilidade de se tomar a decisão correta ao rejeitar H_0 . Esses resultados são válidos para os testes bilaterais e unilaterais.

Nos gráficos em que o valor da variância é 0,0625, a linha que corresponde aos valores da amostra $n = 200$ está sobreposta pela linha que corresponde aos valores da amostra $n = 1.000$. E no gráfico em que a variância é 25 e o valor de α é 0,01 a linha que corresponde aos valores da amostra $n = 5$ está sobreposta pela linha que corresponde aos valores da

amostra $n = 10$. Essas sobreposições ocorrem devido ao fato de os resultados obtidos serem muito próximos.

Pode-se observar na tabela 3 que os gráficos nos quais o valor da variância é 0,0625, a linha que corresponde aos valores da amostra $n = 200$ está sobreposta pela linha que corresponde aos valores da amostra $n = 1.000$. E no gráfico em que a variância é 25 e o valor de α é 0,01 a linha que corresponde aos valores da amostra $n = 5$ está sobreposta pela linha que corresponde aos valores da amostra $n = 10$. Assim como na Tabela 2, os resultados são próximos, e devido a isso ocorrem essas sobreposições.

Assim como nas tabelas 2 e 3 em que os resultados são muito próximos, os gráficos da tabela 4 nos quais o valor da variância é 0,0625 tem a linha que corresponde aos valores da amostra $n = 200$ sobreposta pela linha que corresponde aos valores da amostra $n = 1.000$.

CONCLUSÃO

Com o estudo de simulação de Monte Carlo constatou-se que o tamanho do teste *t-Student* está próximo do nível de significância do teste, mesmo para diferentes tamanhos amostrais. Em relação ao poder do teste, observa-se que este aumenta quando o tamanho amostral e o nível de significância aumentam, e quando a média alternativa se distancia da média sob hipótese nula e ainda, quando a variabilidade diminui. Os resultados encontrados neste trabalho corroboram com os resultados encontrados na literatura estatística por outros autores.

Portanto, ao realizar o teste *t-Student* para média o tamanho do teste se mostrou controlável ao fixar o nível de significância. Entretanto, para o poder do teste é necessário a preocupação com alguns fatores, uma vez que, o poder é sensível ao tamanho amostral, à variabilidade, ao nível de significância do teste e aos valores alternativos da média.

REFERÊNCIAS

BARROS, E. A. C.; MAZUCHELI J. Um estudo sobre o tamanho e poder dos testes *t-Student* e Wilcoxon. **Acta Sci. Technol.** Maringá, v. 27, n. 1, p. 23-32, Jan./Jun., 2005.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística Básica**. São Paulo: Saraiva, 2011.

CARRASCO, C. G.; LEMES, T. S. Uma Avaliação do Erro Tipo II no Uso do Teste *t-student*. **C.Q.D. – Revista Eletrônica Paulista de Matemática**, v. 3, n. 1, p. 07-16, 2014.

CARRASCO, C. G.; SILVA, L. A. Avaliação do erro tipo I na aplicação de um teste de hipóteses para a média. **Revista Mosaicum**, Bahia, Nº. 9, p. 63-68, 2009.

CARRASCO, C. G.; SILVA, L. A. Um estudo do erro tipo II em um teste de hipóteses para a média. **Revista Nucleus**, v. 10, n. 2, p. 7-12, 2013.

CASSELLA, G.; BERGER, R. L. **Inferência estatística**. São Paulo: Cengage Learning, 2010.

COSTA NETO, P. L. O. **Estatística**. São Paulo: Edgard Blücher, 1977.

MAGALHÃES, M. N.; LIMA, A. C. P. **Noções de probabilidade e estatística**. São Paulo: Edusp, 2010.

MUN, J. **Applying Monte Carlo simulation, real options analysis, forecasting, and optimization techniques**. New Jersey: John Wiley & Sons, 2006.

PETERNELLI, L. A.; MELLO, M. P. **Conhecendo o R - uma visão estatística**. Viçosa: UFV, 2007.

VENABLES, W. N.; SMITH, D. M. **An Introduction to R: notes on r: a programming environment for data analysis and graphics, Version 3.6.1 (2019-07-05)**. Disponível em: <<https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>>. Acesso em: 10 set. 2019.