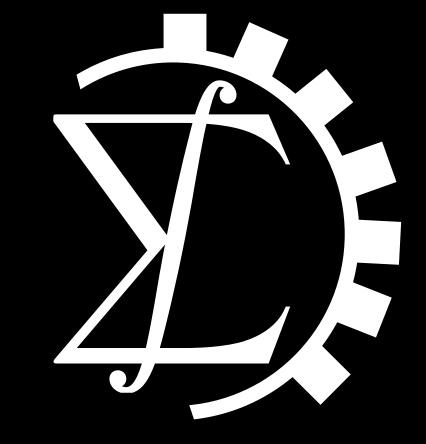


# Estudo de simulação Monte Carlo para testes post hoc

André Felipe B. Menezes andrefelipemaringa@gmail.com

Vinícius Basseto Félix felix\_prot@hotmail.com



# Introdução

Por meio do método de simulação Monte Carlo avaliamos neste trabalho a taxa do erro Tipo I por experimento e o poder de dez testes de comparação múltipla de médias. O estudo foi conduzido no ambiente estatístico R com três grupos balanceados, variando seus tamanhos  $\{2,3,5,10\ e\ 20\}$  e variâncias  $\{1,4\ e\ 9\}$ . Sem perdas de generalidades fixamos o parâmetro de locação  $\mu=0$  e para cada cenário foram geradas M=5000 amostras pseudoaleatórias da distribuição normal. Os testes avaliados foram: LSD, t-Bonferroni, Tukey, SNK, Duncan, Scheffé, Nemenyi, Dunn, Conover e vanWaerden.

## Erro Tipo I por experimento

O erro tipo I por experimento, ou em inglês family-wise Type I error, representa a probabilidade de cometer pelo menos um erros do Tipo I em um conjunto (família) de hipóteses. Para estimar ele geramos amostras independentes sob a hipótese nula e calculamos a proporção de vezes que  $H_0$  foi rejeitada erradamente. Formalmente temos:

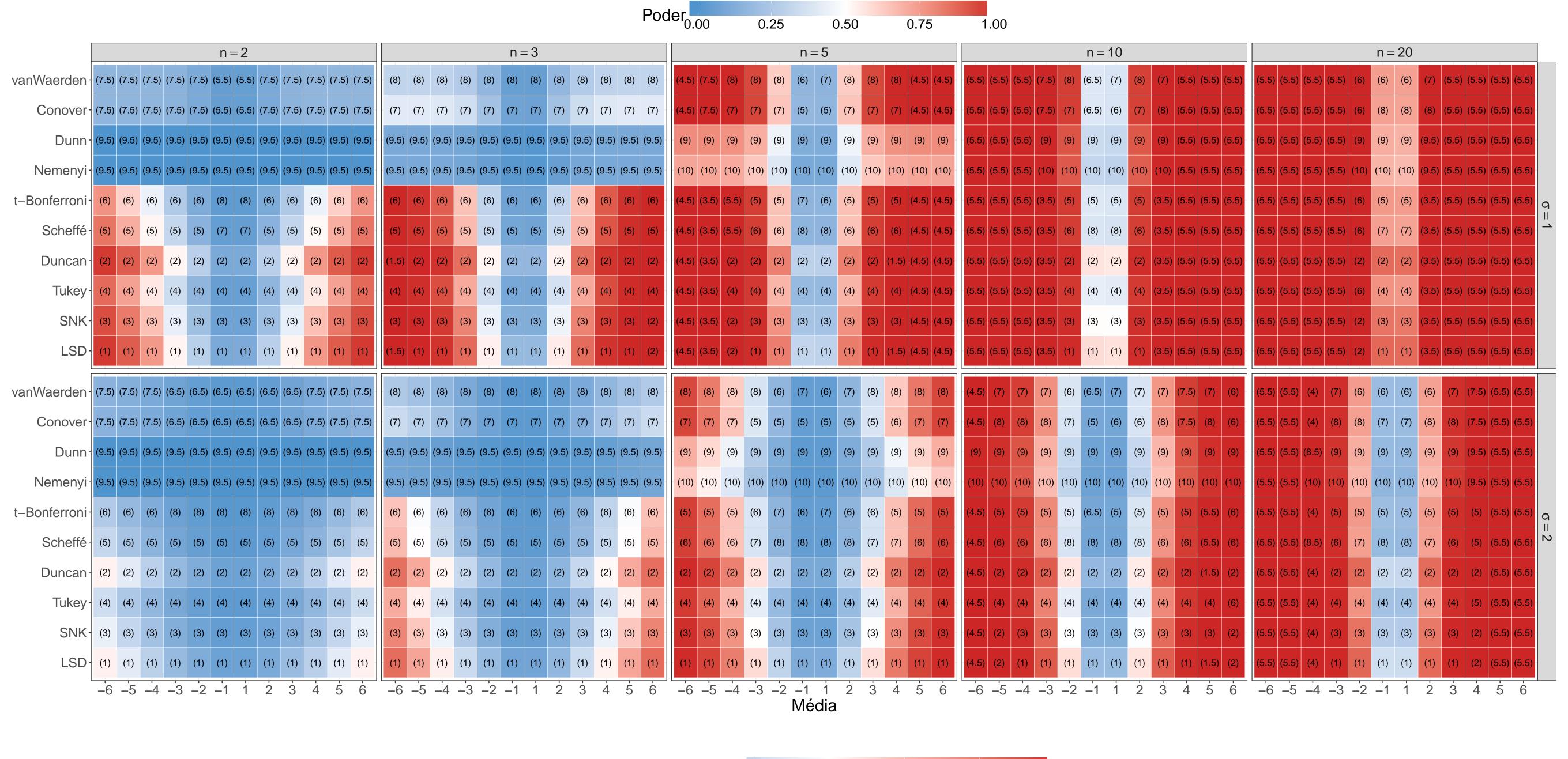
Número de vezes em que  $H_0$  foi rejeitada em pelo menos uma hipótese  $\mid H_0$  é verdadeira

### Poder do Teste

Um dos principais problemas em estudo do poder do teste é o número ilimitado de hipóteses alternativas que podem ser formuladas. Neste trabalho foi considerado um tratamento com média diferente dos demais. Assim sendo, o poder empírico do teste foi obtido por:

Número de vezes em que  $H_0$  é rejeitada nas hipóteses específicas  $\parallel$ 

#### Resultados



				Taxa de erro	Гіро I <sub>0.00</sub>	0.05	0.10	0.15	0.20		
	$\sigma = 1$						$\sigma = 2$				
vanWaerden-	0.067 <sup>(5.5)</sup>	0.046 <sup>(3)</sup>	0.043 <sup>(4)</sup>	0.044 <sup>(3)</sup>	0.046 <sup>(3.5)</sup>		0.067 <sup>(5.5)</sup>	0.042 <sup>(4)</sup>	0.048 <sup>(1)</sup>	0.047 <sup>(3)</sup>	0.047 <sup>(3.5)</sup>
Conover-	0.067 <sup>(5.5)</sup>	0.054 <sup>(2)</sup>	0.047 <sup>(2)</sup>	0.046 <sup>(2)</sup>	0.047 <sup>(2)</sup>		0.067 <sup>(5.5)</sup>	0.049 <sup>(1)</sup>	0.054 <sup>(3)</sup>	0.051 <sup>(1)</sup>	0.048 <sup>(1)</sup>
Dunn-	0 <sup>(9.5)</sup>	0.009 <sup>(7.5)</sup>	0.037 <sup>(7)</sup>	0.038 <sup>(7)</sup>	0.045 <sup>(5)</sup>		0 <sup>(9.5)</sup>	0.013 <sup>(7.5)</sup>	0.042 <sup>(6)</sup>	0.044 <sup>(5.5)</sup>	0.045 <sup>(5)</sup>
Nemenyi-	0 <sup>(9.5)</sup>	0.009 <sup>(7.5)</sup>	0.029 <sup>(8)</sup>	0.033 <sup>(8)</sup>	0.037 <sup>(8)</sup>		0 <sup>(9.5)</sup>	0.013 <sup>(7.5)</sup>	0.033 <sup>(8)</sup>	0.038 <sup>(8)</sup>	0.039 <sup>(8)</sup>
Ocheffé Scheffé	0.034 <sup>(4)</sup>	0.04 <sup>(5)</sup>	0.038 <sup>(5.5)</sup>	0.043 <sup>(4)</sup>	0.046 <sup>(3.5)</sup>		0.035 <sup>(4)</sup>	0.041 <sup>(5)</sup>	0.045 <sup>(4)</sup>	0.048 <sup>(2)</sup>	0.047 <sup>(3.5)</sup>
Scheffé-	0.044 <sup>(2)</sup>	0.045 <sup>(4)</sup>	0.038 <sup>(5.5)</sup>	0.039 <sup>(5)</sup>	0.039 <sup>(7)</sup>		0.047 <sup>(2)</sup>	0.046 <sup>(3)</sup>	0.044 <sup>(5)</sup>	0.044 <sup>(5.5)</sup>	0.042 <sup>(6)</sup>
Duncan-	0.095 <sup>(7)</sup>	0.104 <sup>(9)</sup>	0.098 <sup>(9)</sup>	0.098 <sup>(9)</sup>	0.1 <sup>(9)</sup>		0.098 <sup>(7)</sup>	0.105 <sup>(9)</sup>	0.101 <sup>(9)</sup>	0.104 <sup>(9)</sup>	0.105 <sup>(9)</sup>
Tukey-	0.05 <sup>(1)</sup>	0.053 <sup>(1)</sup>	0.048 <sup>(1)</sup>	0.053 <sup>(1)</sup>	0.053 <sup>(1)</sup>		0.052 <sup>(1)</sup>	0.053 <sup>(2)</sup>	0.053 <sup>(2)</sup>	0.056 <sup>(4)</sup>	0.053 <sup>(2)</sup>
SNK-	0.059 <sup>(3)</sup>	0.061 <sup>(6)</sup>	0.057 <sup>(3)</sup>	0.062 <sup>(6)</sup>	0.061 <sup>(6)</sup>		0.059 <sup>(3)</sup>	0.061 <sup>(6)</sup>	0.063 <sup>(7)</sup>	0.062 <sup>(7)</sup>	0.061 <sup>(7)</sup>
LSD-	0.096 <sup>(8)</sup>	0.117 <sup>(10)</sup>	0.117 <sup>(10)</sup>	0.117 <sup>(10)</sup>	0.124 <sup>(10)</sup>		0.099 <sup>(8)</sup>	0.114 <sup>(10)</sup>	0.121 <sup>(10)</sup>	0.127 <sup>(10)</sup>	0.126 <sup>(10)</sup>
	2	3	5	10	Z <sub>0</sub> Tamanho a	most	ral do grupo	3	5	10	20

#### Conclusão

- A medida que o tamanho da amostra aumenta os testes ficam mais poderosos. No entanto, o aumento da variabilidade implica em diminuição de poder dos testes;
- Os testes mais poderosos foram LSD, Duncan, SNK e Tukey. Os testes não paramétricos mostraram-se menos poderosos;
- Em relação a taxa do erro Tipo I por experimento os testes de **Tukey** e **Conover** foram os melhores;
- Ao se considerar ambos os critérios o teste de Tukey obteve a melhor performance.