



BEN DÊIVIDE DE OLIVEIRA BATISTA

**TESTES DE COMPARAÇÕES MÚLTIPLAS BASEADOS NA
DISTRIBUIÇÃO DA *MIDRANGE* ESTUDENTIZADA
EXTERNAMENTE**

LAVRAS - MG

2016

BEN DÊIVIDE DE OLIVEIRA BATISTA

**TESTES DE COMPARAÇÕES MÚLTIPLAS BASEADOS NA DISTRIBUIÇÃO
DA *MIDRANGE* ESTUDENTIZADA EXTERNAMENTE**

Tese 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 Doutor.

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).**

Batista, Ben Deivide de Oliveira.

Testes de comparações múltiplas baseados na distribuição da
midrange estudentizada externamente / Ben Deivide de Oliveira Batista.
– Lavras : UFLA, 2016.

194 p.: il.

Tese (Doutorado) – Universidade Federal de Lavras, 2016.

Orientador: Daniel Furtado Ferreira.

Bibliografia.

1. Inferência simultânea. 2. Simulação. 3. Poder. 4. Erro tipo I por
experimento. 5. R. I. Universidade Federal de Lavras. II. Título.

BEN DÊIVIDE DE OLIVEIRA BATISTA

**TESTES DE COMPARAÇÕES MÚLTIPLAS BASEADOS NA DISTRIBUIÇÃO
DA *MIDRANGE* ESTUDENTIZADA EXTERNAMENTE**

Tese 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 Doutor.

APROVADA em 15 de julho de 2016.

Dr. Daniel Furtado Ferreira	UFLA
Dr. Augusto Ramalho de Moraes	UFLA
Dr. José Airton Rodrigues Nunes	UFLA
Dr. Eric Batista Ferreira	UNIFAL
Dr. Júlio Silvio de Sousa Bueno Filho	UFLA

Dr. Daniel Furtado Ferreira
Orientador

LAVRAS - MG
2016

*A certeza de que nada sei
é incumbida pela busca do conhecimento.*

Ben Dêvide

*A minha fortaleza, Falcão (Pai), Chica (Mãe),
Colibri (Irmão), Alanaluna (Esposa),
pela paciência e compreensão.*

DEDICO

AGRADECIMENTOS

Ao meu orientador e amigo Daniel Furtado Ferreira, agradeço pela paciência e dedicação que teve na orientação desse trabalho. Tenho uma grande respeito pelo Prof. Daniel, uma vez que consegui dar passos largos sobre a Estatística durante a sua orientação. Em sua defesa de Memorial, o Professor Décio Badin na banca teceu o seguinte comentário: "... conversando com o Professor Conagin, ele me informou que o Professor Daniel é o Pimentel Gomes de Lavras". Isso reflete sua total dedicação no seu trabalho de pesquisa e docência. Assim, tenho o Prof. Daniel como espelho profissional, do qual tenho total admiração.

Ao meu círculo sagrado, Chica e Léo Batista (pais), Álefe Batista (irmão) e Allanna Lopes (Esposa). Queria pedir inicialmente as minhas desculpas pela ausência durante esse tempo. Não é fácil se ausentar das pessoas que amamos. O meu muito obrigado pela paciência que tiveram durante essa jornada. Chica e Léo Batista meus pilares do equilíbrio. Meu irmão Álefe determinado e o meu obrigado por sempre acreditar em mim. E a minha flor e companheira Allanna, em que posso elencar adjetivos mais belos, mas simplifico no amor e dedicação pelo nosso lar.

A minha família, representados por Alcides Batista e Josélia (avós) por parte de pai, e a Raimunda leite e Antônio Bento (*In memoriam*) representando meus familiares por parte de mãe. E representando todos os meus tios, quero deixar a minha homenagem ao velho "Bisaca de Bala", Alcimar Batista (*In memoriam*), obrigado por ter compartilhado com você bons momentos.

Aos amigos da UFLA/DEX, quero agradecer pelos bons momentos de estudo e descontração.

A Capes por ter me proporcionado todo o aparato financeiro para que eu pudesse prosseguir nesse estudo.

RESUMO

Este trabalho teve o objetivo de desenvolver quatro tipos de testes de comparações múltiplas, avaliar o desempenho dos testes propostos, por meio do erro tipo I por experimento e o poder, em diversos cenários de simulação e compará-los com o dos testes presentes na literatura, principalmente com os testes de Tukey, SNK e de Scott-Knott, e finalmente, desenvolver um pacote R, chamado *midrangeMCP*, para auxiliar na aplicação dos testes propostos. Os testes propostos são: teste de Tukey *Midrange* (TM), teste de Student-Newman-Keuls *Midrange* (SNKM), teste de agrupamento de médias *Midrange* (MGM) e teste de agrupamento de médias *Range* (MGR). Os testes de MGM, de MGR e de TM apresentaram desempenho superior aos testes originais de Tukey e de Skott-Knott, na maioria das avaliações feitas, exceto para o erro tipo I por experimento, sob H_0 parcial. O teste de SNKM apresentou desempenho inferior quando comparado ao teste de SNK original em quase todos os cenários simulados. O teste de MGM obteve maior destaque dentre os testes propostos na avaliação de desempenho, agregando a vantagem de não apresentar ambiguidade em seus resultados.

Palavras-chave: Inferência simultânea. Simulação. Poder. Erro tipo I por experimento. R.

ABSTRACT

This paper proposes four types of multiple comparison tests, to evaluate the performance of proposed tests in large experimental situations via Monte Carlo simulation, to compare the experimentwise error rate and power of the proposed tests with the tests in the literature, mainly with the of Tukey, SNK and Scott-Knott tests, and finally, a R package was developed, `midrangeMCP`, to facilitate the users in applying the tests. The proposed tests are: Midrange Tukey test (TM), Midrange Student-Newman-Keuls test (SNKM), Mean Grouping based on the Midrange test (MGM) and Mean Grouping based on the Range test (MGR). The MGM, MGR and TM tests showed superior performance to the original Tukey and Skott-Knott tests, except when the tests were conducted under partial H_0 . The SNKM test was the only one that has a low performance when compared to the original test (SNK test) for all the evaluated situations. The MGM test obtained greater prominence among the proposed tests when evaluated its performance, adding the advantage of not introduce ambiguity in the results.

Keywords: Simultaneous Inference. Simulation. Power. Experimentwise error rate. R.

LISTA DE FIGURAS

Figura 1	Representação da taxa de erro, em inferências simultâneas, em relação ao número de testes.	33
Figura 2	Etapas do desenvolvimento e avaliação dos testes propostos.	69
Figura 3	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de MGM, de TM, de SNKM, em função do número de repetições, hipótese H_0 completa, $n = 10$, $\alpha = 5\%$	89
Figura 4	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de MGM, de TM, de SNKM, em função do número de tratamentos, hipótese H_0 completa, $r = 20$, $\alpha = 5\%$	90
Figura 5	Ordem crescente do rigor do erro tipo I dos procedimentos de comparações múltiplas estudados.	91
Figura 6	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função da diferença de médias consecutivas (δ), sob hipótese H_0 parcial, a) $n = 5$, b) $n = 20$ e c) $n = 100$, para $\alpha = 5\%$ e $r = 10$	95
Figura 7	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de repetições, sob hipótese H_0 parcial, para $n = 10$, $\alpha = 5\%$	96
Figura 8	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de MGM, de TM e de SNKM, em função do número de tratamentos, sob hipótese H_0 parcial, para a) $\delta = 2$, b) $\delta = 4$ e c) $\delta = 16$, com $r = 10$, $\alpha = 5\%$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	98
Figura 9	Poder dos testes de SNK, de SNK modificado, de SNKM, de TM e de Tukey, em porcentagem, sob H_1 completa, para detectar uma diferença entre médias de $2\sigma_{\bar{Y}}$, com $r = 4$ repetições, em função do número de tratamentos, para um $\alpha = 0,05$	99
Figura 10	Poder dos testes de Scott-Knott, de MGR e de MGM, em porcentagem, sob H_1 completa, para detectar uma diferença entre médias de $2\sigma_{\bar{Y}}$, com $r = 4$ repetições, em função do número de tratamentos, para um $\alpha = 0,05$	101

Figura 11	Poder dos testes de Duncan, de Scott-Knott, de MGM, de MGR, de SNK, de SNK modificado, de SNKM, t, t-bayesiana, de TM e de Tukey, em porcentagem, sob H_1 completa, para detectar uma diferença entre médias de $4\sigma_{\bar{Y}}$, com $r = 4$ repetições, em função do número de tratamentos, para um $\alpha = 0,05$	101
Figura 12	Poder dos testes de Scott-Knott, de MGM, de MGR, de SNK, de SNKM, de TM e de Tukey, em porcentagem, sob H_1 completa, para detectar uma diferença entre médias de $2\sigma_{\bar{Y}}$, com (a) $n = 5$, (b) $n = 20$ e (c) $n = 100$, em função do número de repetições, para um $\alpha = 0,05$	103
Figura 13	Poder dos testes de MGM, de MGR, de SNK, de SNKM, de TM e de Tukey, em porcentagem, sob H_1 completa, para detectar diferenças entre médias de 1 a $32\sigma_{\bar{Y}}$, considerando os tratamentos (a) $n = 5$, (b) $n = 20$ e (c) $n = 100$, 4 repetições, e um nível de significância de 5% de probabilidade.	105
Figura 14	Poder dos testes de MGM, de MGR, de SNK, de SNKM, de TM, de Tukey, de Duncan, de SNK modificado e do t-bayesiano, em porcentagem, sob H_1 completa, para detectar diferenças entre médias de 2 a $32\sigma_{\bar{Y}}$, considerando o número de tratamentos (a) $n = 5$, (b) $n = 20$ e (c) $n = 100$, 4 repetições, e um nível de significância de 5% de probabilidade.	107
Figura 15	Poder dos testes de MGM, de MGR, de SNK, de SNKM, de TM, e de Tukey, em porcentagem, sob H_1 completa, para detectar diferenças entre médias de 4 a $32\sigma_{\bar{Y}}$, considerando o número de tratamentos (a) $n = 5$, (b) $n = 20$ e (c) $n = 100$, 4 repetições, e um nível de significância de 5% de probabilidade.	108
Figura 16	Poder dos testes de MGM, de MGR, de SNK, de SNKM, de TM e de Tukey, em porcentagem, sob H_1 completa, com relação aos valores iniciais das diferenças reais entre médias para $4\sigma_{\bar{Y}}$, para $n = 5$, $r = 4$ e $\alpha = 0,05$	109
Figura 17	Poder dos testes de MGM, de MGR, de SNK, de SNKM, de TM, e de Tukey, em porcentagem, sob H_0 parcial, em função do número de tratamentos, para detectar diferenças entre médias de (a) $1\sigma_{\bar{Y}}$, (b) $2\sigma_{\bar{Y}}$, (c) $4\sigma_{\bar{Y}}$ e (d) $8\sigma_{\bar{Y}}$, com 4 repetições, e um nível de significância de 5% de probabilidade.	110

Figura 18 Como acessar as opções de um objeto que recebe a função `MRtest`. . . 138

LISTA DE TABELAS

Tabela 1	Tipos de erros e decisões corretas, com suas respectivas probabilidades em um teste de hipótese.	31
Tabela 2	Dados de teor de nitrogênio das plantas de trevo vermelho relacionados as bactérias em estudo.	84
Tabela 3	Seleção de vinte e cinco genótipos de sorgo com base no período de floração, avaliados pelo teste de Skott-Knott.	85
Tabela 4	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 completa, ao nível de significância $\alpha = 1\%$ de probabilidade, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	87
Tabela 5	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 completa, ao nível de significância $\alpha = 5\%$ de probabilidade, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	88
Tabela 6	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 5\%$ de probabilidade e $\delta = 1$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	92
Tabela 7	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 5\%$ de probabilidade e $\delta = 2$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	93

Tabela 8	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 5\%$ de probabilidade e $\delta = 4\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança 99% de probabilidade.	94
Tabela 9	Resultados da análise do Exemplo 3.1 dos testes de TM e de SNKM, de Tukey, de SNK, de MGM, de MGR e de Scott-Knott.	144
Tabela 10	Resultados dos testes de MGM, de MGR e de Scott-Knott avaliando os 25 genótipos de sorgo apresentado no Exemplo 3.2.	148
Tabela 11	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 5\%$ de probabilidade e $\delta = 8\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	160
Tabela 12	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 5\%$ de probabilidade e $\delta = 16\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	160
Tabela 13	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 5\%$ de probabilidade e $\delta = 32\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	161
Tabela 14	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 1\%$ de probabilidade e $\delta = 1\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	161

Tabela 15	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 1\%$ de probabilidade e $\delta = 2\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	162
Tabela 16	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 1\%$ de probabilidade e $\delta = 4\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	162
Tabela 17	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 1\%$ de probabilidade e $\delta = 8\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	163
Tabela 18	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 1\%$ de probabilidade e $\delta = 16\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	163
Tabela 19	Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 1\%$ de probabilidade e $\delta = 32\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.	164

Tabela 20	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com erro padrão ($1\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 1% de probabilidade.	165
Tabela 21	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com erro padrão ($1\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 5% de probabilidade.	166
Tabela 22	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com erro padrão ($1\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 1% de probabilidade.	167
Tabela 23	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com erro padrão ($1\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 5% de probabilidade.	168
Tabela 24	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com erro padrão ($1\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 1% de probabilidade.	169
Tabela 25	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com erro padrão ($1\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 5% de probabilidade.	170

- Tabela 26 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com dois erros padrões ($2\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 1% de probabilidade. 171
- Tabela 27 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com dois erros padrões ($2\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 5% de probabilidade. 172
- Tabela 28 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com dois erros padrões ($2\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 1% de probabilidade. 173
- Tabela 29 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com dois erros padrões ($2\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 5% de probabilidade. 174
- Tabela 30 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com dois erros padrões ($2\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 1% de probabilidade. 175
- Tabela 31 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com dois erros padrões ($2\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 5% de probabilidade. 176

Tabela 32	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com quatro erros padrões ($4\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 1% de probabilidade.	177
Tabela 33	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com quatro erros padrões ($4\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 5% de probabilidade.	178
Tabela 34	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com quatro erros padrões ($4\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 1% de probabilidade.	179
Tabela 35	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com quatro erros padrões ($4\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 5% de probabilidade.	180
Tabela 36	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com quatro erros padrões ($4\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 1% de probabilidade.	181
Tabela 37	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com quatro erros padrões ($4\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 5% de probabilidade.	182

Tabela 38	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com oito erros padrões ($8\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 1% de probabilidade.	183
Tabela 39	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com oito erros padrões ($8\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 5% de probabilidade.	184
Tabela 40	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com oito erros padrões ($8\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 1% de probabilidade.	185
Tabela 41	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com oito erros padrões ($8\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 5% de probabilidade.	186
Tabela 42	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com oito erros padrões ($8\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 1% de probabilidade.	187
Tabela 43	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com oito erros padrões ($8\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 5% de probabilidade.	188

Tabela 44	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de um erro padrão ($1\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 1% de probabilidade.	189
Tabela 45	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de um erro padrão ($1\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 5% de probabilidade.	189
Tabela 46	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de dois erros padrões ($2\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 1% de probabilidade.	190
Tabela 47	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de dois erros padrões ($2\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 5% de probabilidade.	190
Tabela 48	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de quatro erros padrões ($4\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 1% de probabilidade.	191
Tabela 49	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de quatro erros padrões ($4\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 5% de probabilidade.	191

Tabela 50	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de oito erros padrões ($8\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 1% de probabilidade.	192
Tabela 51	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de oito erros padrões ($8\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_1 parcial, ao nível de significância de 5% de probabilidade.	192
Tabela 52	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de dezesseis erros padrões ($16\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_1 parcial, ao nível de significância de 1% de probabilidade.	193
Tabela 53	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de dezesseis erros padrões ($16\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 5% de probabilidade.	193
Tabela 54	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de trinta e dois erros padrões ($32\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 1% de probabilidade.	194
Tabela 55	Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de trinta e dois erros padrões ($32\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 5% de probabilidade.	194

SUMÁRIO

1	INTRODUÇÃO	23
2	REFERENCIAL TEÓRICO	26
2.1	Procedimento de Comparações Múltiplas	26
2.1.1	A filosofia das comparações múltiplas	29
2.1.2	PCMs baseados na amplitude estudentizada externamente	35
2.1.3	Teste de Scott-Knott	42
2.2	Distribuição da <i>Midrange</i> estudentizada externamente	43
3	MÉTODOS	69
3.1	Filosofia dos testes propostos	70
3.1.1	Teste de Tukey baseado na <i>midrange</i> (TM)	72
3.1.2	Teste de SNK baseado na <i>midrange</i> (SNKM)	74
3.1.3	Teste de agrupamento de médias baseado na <i>midrange</i> (MGM)	74
3.1.4	Teste de agrupamento de médias baseado na amplitude estudentizada (MGR)	75
3.2	Avaliação do desempenho	76
3.3	Pacote <i>midrangeMCP</i>	80
3.3.1	Instalação do pacote <i>midrangeMCP</i>	80
3.3.2	Função <i>MRtest</i>	81
3.3.3	Função <i>MRbarplot</i>	82
3.3.4	Função <i>MRwrite</i>	82
3.3.5	Aplicações do Pacote <i>MidrangeMCP</i>	83
4	RESULTADOS	86
4.1	Erro tipo I	86
4.1.1	Erro tipo I por experimento sob H_0 completa	86
4.1.2	Erro tipo I por experimento sob H_0 parcial	91
4.2	Poder	97
4.2.1	Poder sob H_1 completa	97
4.2.2	Poder sob H_0 parcial	109
4.3	Aplicações dos testes propostos	111
4.3.1	Teste de agrupamento de médias <i>Midrange</i> (MGM) - Solução analítica	111
4.3.2	Teste de agrupamento de médias <i>Range</i> (MGR) - Solução analítica	114

4.3.3	Teste de Student-Newman-Keuls <i>Midrange</i> (SNKM) - Solução analítica .	116
4.3.4	Teste de Tukey <i>Midrange</i> (TM) - Solução analítica	121
4.3.5	Usando o pacote midrangeMCP	125
4.3.6	Comparando os resultados dos testes	144
5	CONCLUSÃO	151
	REFERÊNCIAS	152
	APÊNDICE A	160
	APÊNDICE B	165
	APÊNDICE C	189

1 INTRODUÇÃO

Ao se testar uma hipótese nula de igualdade de duas ou mais médias, duas decisões podem ser tomadas: (i) não rejeitar a hipótese ou (ii) rejeitá-la. Considerando a segunda situação, fica uma outra indagação: quais médias diferem entre si? Vários pesquisadores têm dado atenção a essa questão, propondo uma variedade de procedimentos de comparações múltiplas (PCMs), que identificam as diferenças existentes entre as médias.

O desenvolvimento de PCMs é baseado nos tipos de erros envolvidos nos testes da hipótese estudada. O primeiro é o erro tipo I, cometido quando se rejeita a hipótese nula dado que ela é verdadeira. O segundo é o erro tipo II, em que não se rejeita a hipótese nula dado que ela é falsa. A probabilidade do erro tipo I está relacionada ao poder, sendo este último o complemento da probabilidade de se incorrer no erro tipo II. Assim, o que se deseja no procedimento do teste é que a decisão não contenha esses erros. Na prática isso é impossível, pois até hoje, não há um teste de comparação múltipla na literatura que esteja dentro desse padrão. Na realidade, o que se busca é a minimização de uma combinação linear entre esses dois tipos de erros.

Considerando os PCMs para todos os pares de médias duas a duas, diversos testes foram propostos na literatura. Os mais conhecidos nessa classe foram os testes baseados na distribuição da amplitude estudentizada externamente. O termo “estudentizar” foi primeiro utilizado por William Gosset (*Student*), que o sugeriu a E. S. Pearson para facilitar a utilização da amplitude em problemas práticos. Essa escolha baseou-se na ideia de dividir a amplitude pelo desvio padrão amostral ao invés do desvio padrão populacional. Esse procedimento ficou conhecido como “estudentização”. Assim, ao dividir qualquer estatística por um desvio padrão amostral correspondente, diz-se que esta estatística foi estudentizada.

A distribuição da amplitude estudentizada externamente leva em consideração a máxima diferença entre as observações numa determinada amostra aleatória. Pensando nos procedimentos de comparações múltiplas, as observações são as médias dos efeitos dos tratamentos e, por consequência, o PCM baseado nessa distribuição também levará em conta essa máxima diferença entre médias. Assim, a filosofia para os testes baseados nessa estatística seria da seguinte forma: se a diferença entre duas médias estudentizada em um dado experimento superar o valor do quantil superior $100\alpha\%$ da distribuição da amplitude estudentizada, há evidências para afirmar que essas médias diferem entre si.

Essa é uma proposta de teste interessante. Os testes mais conhecidos baseados nessa distribuição são: teste de Tukey, teste de SNK e teste de Duncan. Apesar desses testes usarem a mesma distribuição, são bem diferentes em sua filosofia. Eles ganharam tanta popularidade que até hoje são utilizados.

Da mesma forma que a amplitude estudentizada externamente, uma outra estatística também baseada em estatísticas de ordem é a *midrange* estudentizada externamente, dada pela razão entre a *midrange* e o desvio padrão amostral, sendo a *midrange* a média entre a menor e a maior estatística de ordem de uma amostra aleatória.

Entretanto, apesar da popularidade dos testes baseados na amplitude estudentizada externamente, existem alguns problemas. Por exemplo, a probabilidade do erro tipo I por experimento para o teste de Duncan são superiores ao nível nominal, o que o torna um teste liberal. Por outro lado, o teste de Tukey tem poder bem inferior aos demais PCMs presentes na literatura à medida que se aumenta o número de médias envolvidas. Além disso, ele apresenta taxas de erros tipo I menores que o nível nominal de significância, o que o caracteriza como um teste conservativo.

Tentando obter testes mais poderosos preservando o nível de significância global, surgiu a proposta de desenvolver PCMs usando a distribuição da *midrange* estudentizada externamente, de modo análogo aos testes de Tukey e de SNK, com o intuito de tentar melhorar o desempenho desses testes. Uma outra proposta foi desenvolver duas versões similares ao teste de Scott-Knott, isto é, testes baseados em agrupamento de médias. Uma baseada na distribuição da amplitude estudentizada externamente e a outra na distribuição da *midrange* estudentizada externamente.

Dessa forma, os objetivos desse trabalho são:

- **Objetivo geral:** desenvolver testes de comparações múltiplas baseado na distribuição da *midrange* estudentizada externamente similares aos testes de Tukey, de SNK e de Scott-Knott. Desenvolver um teste de agrupamento de médias baseado na amplitude estudentizada externamente similar ao teste de Scott-Knott;
- **Objetivos específicos:** Avaliar o desempenho dos testes propostos e compará-los com os dos testes presentes na literatura em diversos cenários experimentais, principalmente com o desempenho dos testes de Tukey, de SNK e de Scott-Knott; Construir um pacote R, chamado *midrangeMCP*, para executar todas as rotinas que envolvem os testes propostos.

O trabalho foi organizado em cinco seções incluindo a Introdução.

A seção 2 apresenta as principais referências utilizadas para o trabalho, dando

enfoque a filosofia dos PCMs e sua história. Será feito uma revisão sobre os testes de comparações múltiplas que servirão como referência para a proposta dos testes, e uma revisão sobre a distribuição da amplitude estudentizada e a distribuição da *midrange* estudentizada externamente.

Na seção 3, é abordada a metodologia do trabalho, apresentando a filosofia dos testes propostos: de Tukey *midrange* (TM), de Student-Newman-Keuls *midrange* (SNKM), de Agrupamento de Médias *midrange* (MGM) e de Agrupamento de Médias *Range* (MGR). Serão abordadas também, a estrutura e as funções do pacote *midrangeMCP*.

Na seção 4, é apresentada a avaliação de desempenho (erro tipo I e poder) dos testes propostos, confrontando-os com os testes originais, por meio de simulação Monte Carlo e resultados encontrados na literatura. Também, serão apresentadas análises de experimentos usando o pacote *midrangeMCP* para ilustrar sua aplicação.

Na seção 5, é mostrada a conclusão do trabalho, enfatizando os resultados mais relevantes.

2 REFERENCIAL TEÓRICO

O pesquisador quando usa a análise de variância para avaliar um experimento, por meio do teste F, deseja obter o máximo de informações sobre a relação entre as verdadeiras médias populacionais ou entre as médias dos tratamentos. Um resultado verificado pela significância do teste F, o deixa completamente insatisfeito (TUKEY, 1949, p. 99), pois o teste F representa um teste médio de diversas comparações independentes. Banzatto e Kronka (2006, p. 62) complementam que, se apenas uma das comparações envolver uma diferença marcante e as outras não, um teste F pode falhar para evidenciar a diferença existente. Como alternativa para identificar essas diferenças específicas, surgiram os procedimentos de comparações múltiplas.

Contudo, os procedimentos de comparações múltiplas (PCM), após o teste F, são a metodologia mais utilizada. Entretanto, se os PCMs são o segundo em uso certamente são o primeiro em mau uso (HSU, 1996, p. vi), uma vez que nem todos os PCMs precisam suceder o teste F, fato bem documentado em Bernhardson (1975) e Wilcox (1987).

A seguir, será apresentado o desenvolvimento histórico das comparações múltiplas, bem como definições e teoremas apresentados para embasar o presente estudo.

2.1 Procedimento de Comparações Múltiplas

Sob uma interpretação ampla, o trabalho que pode ser considerado como um início rudimentar sobre as comparações múltiplas é a obra de Irwin (1925). Esse autor estudou um critério de rejeição de “outliers” inspirado nos trabalhos de Galton (1902) e Pearson (1902). Esses desenvolveram a distribuição da diferença média entre duas observações subsequentes numa amostra aleatória de tamanho n arranjada em ordem de magnitude.

Inspirado nos trabalhos de Tippett (1925) e Pearson (1926) sobre a amplitude de uma população normal, Student (1927) propôs a amplitude como um critério de verificar erros em análises de rotinas em banco de dados. Tempo depois, como inovação não relacionada a esse trabalho, Working e Hotelling (1929) usaram a teoria de erros para interpretar a tendência no modelo de regressão linear simples.

Entretanto, Fisher (1935) acaba sendo um dos primeiros a propor testes de comparações múltiplas. Após a significância do teste F, Fisher propôs o teste t protegido de Fisher. Desenvolveu também como alternativa, o teste t baseado na desigualdade de

Bonferroni (MILLER, 1981, p. 26).

Sugerido por Student (PEARSON, 1939, p.246), Newman (1939) propôs o primeiro PCM com base na distribuição da amplitude estudentizada externamente, dada por $Q = W/S$, em que W é a amplitude de uma amostra de tamanho n de uma população normal e S é o estimador de σ com ν graus de liberdade, sendo S independente de W . Entretanto, um grande problema que existia nessa distribuição era quando $\nu \rightarrow \infty$. Assim, como solução, Newman (1939) usou W/σ ao invés de W/S . Isso facilitou muito os testes baseados nessa distribuição, já que logo após isso, Newman tabulou os quantis de Q . Anos mais tarde, obtendo resultados com maior precisão, Pearson e Hartley (1942, p. 95-99) construíram novas tabelas dos quantis da distribuição de Q . Keuls (1952) desenvolveu uma modificação do teste de Newman que, posteriormente, passou a ser chamado teste de SNK (Student-Newman-Keuls), sendo mais a frente detalhado. A novidade nesse teste é que a diferença mínima significativa (DMS) se modifica à medida que se altera o número de médias envolvidas no contraste entre médias ordenadas. O teste se popularizou tanto que até hoje é usado.

Segundo Miller (1981, p. 2), após a segunda guerra mundial, em especial nos Estados Unidos, as pesquisas em comparações múltiplas passaram a ser foco na área da estatística. É tanto que, durante o período entre 1947 e 1955, as contribuições notórias nessa área foram realizadas pelos principais estatísticos que trabalhavam nos EUA, tais como: David B. Duncan, Charles W. Dunnett, H. O. Hartley, Henry Scheffé e John Tukey.

Tukey (1949) desenvolveu um de seus primeiros testes nos procedimentos de comparações múltiplas, que foi o teste “gap-straggler-variance”. Esse trabalho, segundo Miller (1981, p. 2), talvez tenha sido o principal responsável pela incorporação das técnicas de comparações múltiplas na estatística aplicada. Contudo, esse teste não teve tanta repercussão, talvez devido ao crescente avanço do estudo das comparações múltiplas nessa época. Tukey (1952) estudou vários tipos de taxas de erro e propôs o seu teste baseado na amplitude estudentizada externamente. Ao contrário do teste de SNK, fixou o valor de n como o número máximo de médias de estudo para calcular a DMS. Tukey (1953) escreveu um memorando intitulado “O problema das comparações múltiplas”, estendendo o assunto com o incremento das taxas de erro nos testes de comparações múltiplas. O teste de Tukey é um procedimento aplicado às médias dos tratamentos com mesmo número de repetições. Assim, Kramer (1956) expandiu o teste para comparação de tratamentos com o número de repetições diferentes.

Em sua tese de doutorado, Duncan (1947) elaborou um PCM similar ao teste t.

Duncan (1951, 1952, 1955) aprimorou esse teste passando a usar a distribuição de Q , similar ao teste de Tukey, tendo uma pequena diferença, usava uma proteção no nível de significância, sendo $\alpha_p = (1 - \alpha)^{p-1}$, em que $p = 2, 3, \dots, n$ é o número de comparações envolvidas. Duncan (1965) e Waller e Duncan (1969) também agregaram a inferência bayesiana aos procedimentos de comparações múltiplas.

Scheffé (1953) estabeleceu a ligação entre as comparações múltiplas e os testes de hipóteses dando uma interpretação aos intervalos de confiança simultâneos, método habitual na análise de variância. Dunnett (1955) apresentou um procedimento de comparação múltipla comparando todas as médias com um controle (uma média padrão). Hartley (1955) desenvolveu um PCM de quadrados médios baseado no teste F sequencial.

Miller (1966) em o seu livro “Simultaneous Statistical Inference” popularizou o uso dos PCMs, dando impulso para novas pesquisas nessa área, talvez sendo um dos primeiros livros sobre o assunto. Esse livro foi reeditado em 1981 pelo mesmo autor. Duas outras obras complementaram o assunto, o livro intitulado “Multiple comparison procedures” de Hochberg e Tamhane (1987) e “Multiple comparisons - theory and methods” de Hsu (1996).

Quase todos os PCMs têm um problema de interpretação prática: a ambiguidade em seus resultados, isto é, quando dois níveis do fator, tidos como diferentes entre si, não diferem de um terceiro (BORGES; FERREIRA, 2003, p. 68). Também conhecido como falta de transitividade da não significância do testes. O pesquisador, de uma forma geral, tem muita dificuldade na interpretação dos resultados e na recomendação do melhor tratamento. Como solução, Scott e Knott (1974) propuseram um PCM baseado num teste de agrupamento, usando soma de quadrados para a determinação dos grupos. Esse teste visa a separação de médias de tratamentos em grupos distintos, por meio da minimização da variação dentro e maximização da variação entre grupos. Os resultados são facilmente interpretados devido à ausência de ambiguidade. Dessa forma esse procedimento resulta em maior objetividade e clareza (BORGES; FERREIRA, 2003, p. 68). O problema desse teste, é a alta taxa de erro tipo I sob H_0 parcial.

Para mais detalhes sobre artigos publicados entre 1966 e 1976, ver Miller (1977). Alguns trabalhos no período de 1990 a 2004 podem ser pesquisados em Benjamini, Bretz e Sarkar (2004). Entretanto, em Rao e Swarupchand (2009), há uma extensa nota bibliográfica com 573 obras citadas para consultas sobre os PCMs. Trabalhos recentes sobre os procedimentos de comparações múltiplas utilizando estatística computacional podem

ser encontrados nos livros de Bretz, Hothorn e Westfall (2011), Zieffler, Harring e Long (2011) e Dickhaus (2014).

Apesar de muitos trabalhos no campo das comparações múltiplas, essa área da Estatística ainda está em aberto para a pesquisa, pois ainda não há um teste ideal para os PCMs (GELMAN; HILL; YAJIMA, 2012, p. 209).

Após essa breve descrição histórica, a seguir será dado um enfoque sobre a filosofia dos PCMs.

2.1.1 A filosofia das comparações múltiplas

Antes de mostrar os testes de comparações múltiplas usados como referências para a proposta dos novos testes, é interessante entender como esses métodos são desenvolvidos. O título dessa subseção se baseou em Tukey (1991) quando tentou explicar como desenvolver um PCM e quais os problemas envolvidos.

Um dos interesse dos PCMs está em verificar um conjunto de parâmetros originados da combinação linear de médias tomadas duas a duas, expressos por $\theta_l = \mu_i - \mu_{i'}$, $i \neq i' \in \{1, 2, \dots, n\}$, para $l = 1, 2, \dots, N$. A coleção desses parâmetros é chamado de família (HOCHBERG; TAMHANE, 1987, p. 5), sendo o número de parâmetros dado por

$$N = \binom{n}{2} = \frac{n(n-1)}{2}, \quad (1)$$

considerando as diferenças das n médias duas a duas.

Cada parâmetro é avaliado por meio de um teste de hipóteses. E cada teste está associado erros de decisão: erro tipo I e erro tipo II. Assim, o desenvolvimento de um PCM é realizado avaliando o seu desempenho por meio do erro tipo I e o poder, sendo este último o complemento do erro tipo II. Essa avaliação se baseia em um processo de simulação, em que nesse estudo será com base no método Monte Carlo. Para a avaliação do erro tipo I, a simulação realizada supõe hipótese nula global (H_0), isto é, a simulação realizada gera amostras de n tratamentos em um experimento da mesma população, portanto, com mesma média populacional. Define-se a hipótese H_0 da seguinte forma,

Definição 2.1 — Hipótese nula global H_0 . Seja um fator, de efeito fixo e n níveis, então a hipótese nula global é definida por

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_n, \quad (2)$$

sendo $\mu_1, \mu_2, \dots, \mu_n$ as n médias das populações. ■

A avaliação do poder do teste é baseada na hipótese H_1 completa, definida por:

Definição 2.2 — Hipótese H_1 completa. Seja um fator, de efeito fixo e n níveis, então a hipótese H_1 completa é definida por

$$H_1 : \mu_1 \neq \mu_2 \neq \dots \neq \mu_n, \quad (3)$$

sendo $\mu_1, \mu_2, \dots, \mu_n$ as n médias de n populações. ■

A simulação realizada gera amostras de n tratamentos em um experimento de n populações com médias populacionais diferentes. Contudo, na prática essas n médias dos tratamentos não pertencem necessariamente a n populações. O que há na realidade são g grupos de médias pertencentes a g populações, com $g \leq n$. Nesse caso, a hipótese é expressa numa configuração de médias populacionais em que algumas, mas não todas as médias sejam iguais. Essa hipótese é chamada de hipótese nula parcial. Uma definição formal é apresentada a seguir.

Definição 2.3 — Hipótese nula parcial H_{0_p} . Seja um fator qualquer, de efeito fixo, com n médias e g grupos, em que $g \leq n$, e as p_k médias iguais, para $k = 1, 2, \dots, g$, estão dispostas no j -ésimo grupo, com $j = 1, 2, \dots, p_k$, então a hipótese nula parcial é definida por

$$H_{0_p} : \mu_{11} = \mu_{12} \dots = \mu_{1p_1} \neq \dots \neq \mu_{21} = \mu_{22} \dots = \mu_{2p_2} \neq \dots \neq \mu_{g1} = \mu_{g2} \dots = \mu_{gp_g}, \quad (4)$$

sendo $\sum_{k=1}^g p_k = n$. ■

Observe que essa definição acaba sendo uma hipótese geral em relação às anteriores, pois se for considerado $g = 1$, nada mais é do que a hipótese nula global. Se $g = n$, tem-se a hipótese H_1 completa. Assim, sob H_{0_p} parcial, dentro do j -ésimo grupo, pode

ser avaliado o erro tipo I, e entre grupos, o poder do teste.

De um modo formal, define-se o erro tipo I,

Definição 2.4 — Erro tipo I. Seja H_0 uma hipótese verdadeira. A decisão equivocada de rejeitar H_0 , sendo esta verdadeira, é chamada de erro tipo I. O tamanho do erro tipo I é definido pela probabilidade $\alpha = P(\text{Rejeitar } H_0 | H_0 \text{ verdadeira})$. ■

Diz-se que quando, em uma inferência, se rejeita uma hipótese, dado que ela é verdadeira, esta é chamada de inferência falso positiva. O ideal para um teste é manter o tamanho do erro tipo I no nível de significância adotado pelo pesquisador, chamado de nível de significância global. Por outro lado, não rejeitar uma hipótese que realmente deveria ser rejeitada, comete-se o erro tipo II.

Definição 2.5 — Erro tipo II. Seja H_0 uma hipótese falsa. A decisão equivocada de não rejeitar H_0 , sendo esta falsa, é chamada de erro tipo II. O tamanho do erro tipo II é definido pela probabilidade $\beta = P(\text{Não Rejeitar } H_0 | H_0 \text{ falsa})$. ■

Se por um lado, a probabilidade de se incorrer no erro tipo I é conhecida e fixada pelo pesquisador, por outro, a probabilidade do erro tipo II não é conhecida e nem pode ser especificada *a priori*. O que se sabe é que a probabilidade α é inversamente proporcional à probabilidade β . Relacionado ao erro tipo II, uma outra probabilidade que se torna mais fácil de ser entendida é o poder do teste, sendo apresentado a seguir.

Definição 2.6 — Poder do teste. A probabilidade de rejeitar uma hipótese, dado que ela é falsa, denotada por $1 - \beta$, é chamada de poder do teste. ■

A Tabela 1 mostra o resumo dos tipos de erros envolvidos numa tomada de decisão sob a hipótese nula H_0 e nas respectivas probabilidades.

Tabela 1 Tipos de erros e decisões corretas, com suas respectivas probabilidades em um teste de hipótese.

Decisão	Realidade	
	H_0 verdadeira	H_0 falsa
Não Rejeitar H_0	Decisão correta $1 - \alpha$	Erro Tipo II β
Rejeitar H_0	Erro tipo I α	Poder $1 - \beta$

Diante dessas considerações, segundo Rao e Swarupchand (2009, p. 66) o termo

“comparações múltiplas” refere-se à aplicação de vários testes de significância estatística das diferenças entre as médias (ou proporções ou variâncias, etc) dentro de um grupo. Esses autores ainda definem os procedimentos de comparações múltiplas, apresentados a seguir.

Definição 2.7 — Procedimentos de comparações múltiplas. Os procedimentos estatísticos designados ao controle adequado de efeitos de multiplicidade são chamados de procedimentos de comparações múltiplas. ■

O problema é que se inúmeras hipóteses, e em especial no tipo $H_{0_l} : \mu_i - \mu_{i'}$, para $i \neq i' \in \{1, 2, \dots, n\}$ e $l = 1, 2, \dots, N$, são testadas simultaneamente, nos testes tradicionais ocorre um crescimento no erro tipo I. Este é o efeito de multiplicidade, isto é, o crescimento do erro tipo I com o aumento do número de testes. A forma de medir esse crescimento é realizada pela taxa de erro, definida a seguir.

Definição 2.8 — Taxa de erro em inferências simultâneas. Seja um conjunto de hipóteses H_{0_l} verdadeiras e independentes, com $l = 1, 2, \dots, N$. A probabilidade de pelo menos uma hipótese H_{0_l} ser rejeitada é chamada de taxa de erro, definida por

$$\text{Taxa de erro} = 1 - \prod_{l=1}^N [1 - P(\text{Rejeitar } H_{0_l} | H_{0_l} \text{ verdadeira})]. \quad (5)$$

Um exemplo para ilustrar essa situação é:

■ **Exemplo 2.1 — Efeito de Multiplicidade de PCMs.** Considere um conjunto de 10 testes de hipóteses verdadeiras e independentes, com probabilidade de erro tipo I igual a $P(\text{erro tipo I}) = \alpha$, sendo $\alpha = 0,05$. A taxa de erro é dada por:

$$\begin{aligned} \text{Taxa de erro} &= 1 - \prod_{l=1}^N (1 - \alpha) \\ &= 1 - (1 - \alpha)^N \\ &= 1 - (1 - 0,05)^{10} \\ &= 0,4013. \end{aligned} \quad (6)$$

Considerando 30 testes de hipóteses verdadeiras, a taxa de erro é 0,7854. Se forem

considerados 60 testes, a taxa de erro será 0,9539. Na Figura 1, observa-se assim, que à medida que o número de testes aumenta, a taxa de erro aumenta. Essa taxa de erro nada mais é do que o erro tipo I conjunto de todos os testes.

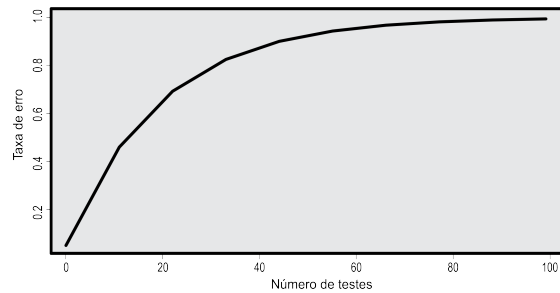


Figura 1 Representação da taxa de erro, em inferências simultâneas, em relação ao número de testes.

Percebe-se que na realidade que as N comparações não são independentes, de modo que essa não é a probabilidade exata, mas uma aproximação conservadora (HOCKING, 2003, p. 655). Diz-se então que esta é a probabilidade máxima ou um limite superior de cometer pelo menos um erro tipo I.

Assim, é com base nas taxas de erro tipo I que os procedimentos de comparações múltiplas são desenvolvidos, com o objetivo de controlar adequadamente a probabilidade de pelo menos uma hipótese verdadeira ser rejeitada, isto é, mesmo que com o aumento dos testes de hipóteses não haja também o aumento na taxa de erro.

Outro ponto importante para a avaliação dos PCMs ser baseada no erro tipo I, é que o erro tipo II não pode ser controlado diretamente pelo pesquisador. Segundo Hochberg e Tamhane (1987, p. 10), a medida de probabilidade do erro tipo I é frequentemente escolhida devido à facilidade de ser analisada e controlada.

Nos PCMs existem duas formas de avaliar o erro tipo I. A primeira avaliação é feita verificando-se a probabilidade de rejeitar uma hipótese verdadeira em todas as possíveis combinações dos níveis dos tratamentos tomados dois a dois, sendo conhecida por taxa de erro tipo I por comparação (TEC), do inglês, “comparisonwise error rate” ou “per-comparison error rate” (RAMALHO; FERREIRA; OLIVEIRA, 2005, p. 86). Segundo O’Neill e Wetherill (1971, p. 221) esse tipo de erro é calculado por:

$$\text{TEC} = \frac{\text{N}^\circ \text{ de comparações rejeitadas indevidamente}}{\text{Número total de comparações}}$$

A segunda forma é avaliada pela taxa de erro tipo I por experimento (TEE), do inglês “experimentwise error rate” (FERREIRA, 2009, p. 215). Segundo O’Neill e Wetherill (1971, p. 221) esse tipo de erro é calculado por:

$$TEE : \frac{\text{N}^\circ \text{ exper. com pelo menos uma hipótese rejeitada indevidamente}}{\text{Número total de experimentos}}.$$

A relação entre esses dois tipos de taxas de erro é relatado por Hochberg e Tamhane (1987, p. 8), dada por

$$TEE = 1 - (1 - TEC)^N, \quad (7)$$

sendo N o número de combinações a serem feitas, conforme expressão (1).

A questão sobre qual das taxas de erros deve ser controlada foi motivo de muitas discussões. Mas, com base em Tukey (1953, cap. 8), Miller (1981, p. 6) e Hochberg e Tamhane (1987, p. 3) a TEE deve ser padrão, pois além de proteger o nível de significância global α , determinado pelo pesquisador, permite calibrar uniformemente procedimentos diferentes para um referencial comum, e assim, comparar as suas características operacionais de maneira mais justa.

Fica evidente pela relação entre essas duas medidas, expressão (7), que o controle da taxa de erro tipo I por experimento controla a TEC de acordo com nível de significância global, porém o contrário não é verdade. Por exemplo, considerando a $TEE = 0,05$, considerando o nível de significância global de $\alpha = 0,05$ para $N = 10$, a taxa de erro por comparação pode ser dada por:

$$TEC = 1 - (1 - TEE)^{1/N} = 1 - (1 - 0,05)^{1/10} = 0,00512,$$

sendo $TEC \leq 0,05$. Já se for considerado $TEC = 0,05$ para $N = 10$, a taxa de erro por experimento é expressa em (6). Assim, é com base na TEE que os testes atualmente são propostos. Pelo fato dessas medidas não terem sido criadas na mesma época, alguns testes foram baseados na TEC, caso clássico como o teste t de *Student* e o teste de Duncan, que controlam a TEC, mas não controlam a TEE.

Com base nessas afirmações, Hsu (1996, p. 27) define a taxa de erro por experimento da seguinte forma:

Definição 2.9 — Taxa de erro por experimento (TEE). A taxa de erro por experimento de um método de comparação múltipla é o supremo, sob H_0 , da probabilidade de se ter ao menos uma afirmativa incorreta, isto é

$$\text{TEE} = \sup_{H_0} P_{H_0}(\text{ao menos uma rejeição incorreta}), \quad (8)$$

sendo H_0 expressa na Definição 2.1. ■

Essa taxa de erro por experimento é definida independentemente se uma hipótese nula completa ou parcial está sendo considerada. Portanto, será a TEE o método adotado para a avaliação dos testes propostos nesse trabalho.

2.1.2 PCMs baseados na amplitude estudentizada externamente

O estudo da amplitude foi iniciado por Karl Pearson (PEARSON, 1902), como um caso especial do problema da diferença de Galton (GALTON, 1902). Mas foi Tippet (1925) um dos primeiros a tratar a amplitude amostral como medida de dispersão, dando uma aplicação a esta medida. Este autor tabulou os quantis da distribuição da amplitude para amostras de tamanho n de uma população com distribuição normal expressa em termos do desvio padrão populacional. Posteriormente, Pearson (1926, 1932), filho de Karl Pearson, ampliou esse trabalho acrescentando outras configurações para computar os quantis dessa distribuição, e para $n = 3$ McKay e Pearson (1933) encontraram a distribuição exata da amplitude para uma população normal. Entretanto, tabelas mais completas sobre a distribuição da amplitude expressa em termos do desvio padrão populacional foram apresentadas por Hartley (1942) e Pearson e Hartley (1942).

Student (1927) ajustou curvas do primeiro e quarto momento teórico da distribuição da amplitude para vários valores de amostra de tamanho n . Esse artigo foi um ponto de partida para uma mudança drástica na aplicação da amplitude utilizada nas comparações múltiplas. Até então, as aplicações utilizadas eram feitas em controle de qualidade. Ao invés do desvio padrão, a amplitude passou a ser utilizada, por ser uma medida de dispersão mais simples, no controle de qualidade aplicado às mudanças na variação da qualidade na indústria (PEARSON; HAINES, 1935, p. 83).

Entretanto, em 1932, *Student* enviou uma carta para E. S. Pearson, publicada mais tarde por Pearson (1939, p. 246), com a ideia de usar a razão entre a amplitude e

o desvio padrão amostral. Essa técnica ficou conhecida por “estudentização”. Isso foi sugerido, pois usar a amplitude expressa em termos do desvio padrão populacional como usada até então, era inviável nas aplicações práticas, pois dificilmente o desvio padrão populacional é conhecido. Assim, o tipo de problema que *Student* tinha em mente já se baseava em aplicações na experimentação, dado um número de médias de tratamentos e um estimador do desvio padrão populacional, o desvio padrão amostral, que é independente da amplitude com ν graus de liberdade. Dessa forma, logo se pensou em um método rápido de verificar a significância entre quaisquer diferenças entre duas médias de tratamentos. Bastava comparar essa diferença com um determinado quantil superior $100\alpha\%$ adotado da distribuição da amplitude estudentizada externamente. Caso a diferença entre as duas médias fosse superior a esse quantil, haveria indicativos de que essas duas médias não viriam de uma mesma população. Esse procedimento é similar ao que foi feito por Student (1927), exceto que a amplitude agora está expressa em termos do desvio padrão amostral.

Seguindo a sugestão de *Student*, Newman (1939) calculou os quantis da distribuição da amplitude estudentizada externamente e propôs um teste baseado nesses quantis. As tabelas de Newman foram obtidas pelas aproximações feitas por Pearson (1932), com relação a amplitude expressa em termos do desvio padrão populacional, também chamado de amplitude padronizada. Tabelas de quantis da distribuição da amplitude estudentizada externamente também foram obtidas por Pearson e Hartley (1943).

A partir desses estudos, um crescente avanço sobre os métodos de comparações múltiplas baseados nos quantis da distribuição da amplitude estudentizada externamente foram desenvolvidos nessa época. Na década de 1950, foram propostos os principais PCMs, tais como: o teste de SNK iniciado por Student (1927), Newman (1939), e adaptado por Keuls (1952), teste de Tukey (1953) e o teste de Duncan (1955). Outros testes podem ter sido propostos baseados na amplitude estudentizada externamente, porém com pouco reconhecimento como estes citados.

Após essa revisão histórica sobre a amplitude, serão mostradas as ideias que esses pesquisadores tiveram para o desenvolvimento de PCMs baseados nessa estatística.

Sejam $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$, as estatísticas de ordem de uma amostra aleatória Y_1, Y_2, \dots, Y_n de tamanho n com função de distribuição $F_Y(y)$ e função densidade $f_Y(y)$, então define-se

Definição 2.10 — Amplitude. A amplitude da amostra é definida por:

$$W = Y_{(n)} - Y_{(1)}. \quad (9)$$

A função de distribuição e a função densidade de W são respectivamente,

$$F_W(w) = n \int_{-\infty}^{\infty} f_Y(z) [F_Y(w+z) - F_Y(z)]^{n-1} dz, \quad (10)$$

e

$$f_W(w) = \int_{-\infty}^{\infty} n(n-1) f_Y(z) f_Y(w+z) [F_Y(w+z) - F_Y(z)]^{n-2} dz, \quad (11)$$

como mostrado em Batista (2012, p. 50-52), David e Nagaraja (2003, p. 13) e Gumbel (1947, p. 385). Considere agora, amostras de tamanho n de uma distribuição normal com média $\mu \in \mathbb{R}$ e variância $\sigma^2 > 0$, cuja função de distribuição é:

$$F_Y(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2} \frac{(t-\mu)^2}{\sigma^2}\right\} dt. \quad (12)$$

Então, define-se a amplitude estudentizada externamente¹ da seguinte forma:

Definição 2.11 — Amplitude Studentizada externamente. A amplitude estudentizada externamente é definida pela razão entre W , Definição 2.10, e S , em que S é o estimador do desvio padrão populacional σ associado a ν graus de liberdade e independente de W , isto é,

$$Q = \frac{W}{S}. \quad (13)$$

Redefinindo Q da Definição 2.11, em busca de obter as funções de distribuição e densidade, tem-se:

$$Q = \frac{W}{S} = \frac{W'}{X},$$

¹O termo “externamente” é utilizado para mostrar que as estatísticas W e S são independentes.

em que $W' = W/\sigma$ é a amplitude padronizada, e $X = S/\sigma$. W' e X são variáveis aleatórias independentes. Essa explicação é bem documentada nos trabalhos de Geary (1936, p. 296), Kamat (1953, p. 124) e David, Hartley e Pearson (1954, p. 483). Esses autores corroboram essa ideia da seguinte forma: seja uma amostra aleatória Y_1, Y_2, \dots, Y_n de tamanho n de uma população normal com variância σ^2 . Considere ainda $\bar{Y} = \sum_{i=1}^n Y_i/n$ e $S^2 = \sum_{i=1}^n (Y_i - \bar{Y})^2/(n-1)$, a média e a variância amostral, respectivamente. Eles mostraram que a distribuição conjunta da razão $(Y_i - \bar{Y})/S$, $i = 1, 2, \dots, n$, é independente de S . Assim, qualquer função desse tipo e em particular $(Y_{(1)} - \bar{Y})/S$ e $(Y_{(n)} - \bar{Y})/S$, ou W/S , também serão independentes de S .

Considerando $X = S/\sigma$, obtido em uma amostra da distribuição normal de tamanho $\nu + 1$, então $\nu S^2/\sigma^2 \sim \chi_\nu^2$, ou seja, $\nu S^2/\sigma^2$ tem distribuição de qui-quadrado com ν graus de liberdade (MOOD; GRAYBILL; BOES, 1974, p. 245). Assim, com abuso de notação, como $\nu S^2/\sigma^2 \sim \chi_\nu^2$, então $\nu S^2/\nu\sigma^2 \sim \chi_\nu^2/\nu$. Dessa forma, pode-se afirmar que $\nu S^2/\nu\sigma^2 = S^2/\sigma^2$ e, conseqüentemente, $S^2/\sigma^2 \sim \chi_\nu^2/\nu$.

Teorema 2.1 — Função densidade de X. Se $X = S/\sigma$ é uma variável aleatória de uma amostra de tamanho $\nu + 1$ de uma distribuição normal com média μ e variância σ^2 , então a função densidade é dada por

$$f_X(x; \nu) = \frac{\nu^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2-1}} x^{\nu-1} e^{-\nu x^2/2}, \quad x > 0. \quad (14)$$

■

Demonstração. Sejam $U = \nu S^2/\sigma^2$ e $X = S/\sigma$, variáveis aleatórias obtidas de uma amostra de tamanho $\nu + 1$ de uma população com distribuição normal com parâmetros μ e σ^2 . A variável aleatória U tem distribuição de qui-quadrado com ν graus de liberdade. Então a distribuição de X pode ser obtida da transformação $U = \nu X^2$. O jacobiano dessa transformação é dado por

$$J = \frac{du}{dx} = 2\nu x,$$

para $x > 0$.

A função densidade de X , amostrada de uma normal, é obtida por:

$$\begin{aligned} f_X(x; \nu) &= f_U(u; \nu)|J| = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} u^{\nu/2-1} e^{-u/2} |J| \\ &= \frac{1}{2^{\nu/2}\Gamma(\nu/2)} (\nu x^2)^{\nu/2-1} e^{-\nu x^2/2} 2\nu x \\ &= \frac{2\nu\nu^{\nu/2-1}}{2^{\nu/2}\Gamma(\nu/2)} x(x^2)^{\nu/2-1} e^{-\nu x^2/2}, \end{aligned}$$

resultando em

$$f_X(x; \nu) = \frac{\nu^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2-1}} x^{\nu-1} e^{-\nu x^2/2}, \quad x \geq 0. \quad \blacksquare$$

Batista (2012, p. 54-55) mostrou que a função de distribuição e a função densidade da amplitude estudentizada externamente são dadas, respectivamente, por:

$$F_Q(q; n, \nu) = \int_0^{\infty} \int_{-\infty}^{\infty} n\phi(y)[\Phi(xq + y) - \Phi(y)]^{n-1} f_X(x; \nu) dy dx, \quad (15)$$

e

$$\begin{aligned} f_Q(q; n, \nu) &= \int_0^{\infty} \int_{-\infty}^{\infty} n(n-1)x\phi(y)\phi(xq + y)[\Phi(xq + y) - \Phi(y)]^{n-2} \times \\ &\quad \times f_X(x; \nu) dy dx, \end{aligned} \quad (16)$$

em que $\phi(y)$ e $\Phi(y)$ são a função densidade e função de distribuição de uma variável aleatória normal padrão, com $y \in \mathbb{R}$ e $f_X(x; \nu)$ é a função densidade de X apresentada no Teorema 2.1. Esses resultados são expressos em função da distribuição normal padrão. Isto é, independente dos parâmetros da distribuição inicial normal, a função densidade e a função de distribuição de Q serão expressos sempre em termos da distribuição normal padrão.

Os testes baseados no quantil superior $100\alpha\%$ da distribuição da amplitude estudentizada externamente, denotado por $q_{(\alpha; n, \nu)}$, com n tratamentos e ν graus de liberdade do resíduo serão apresentados na sequência. Os valores de $q_{(\alpha; \nu, n)}$, podem ser obtidos pela função `qtukey()` (R CORE TEAM, 2015), tais que:

$$P [Q \leq q_{(\alpha; \nu, n)}] = 1 - \alpha. \quad (17)$$

Considere a seguinte amostra aleatória $Y_{11}, Y_{12}, \dots, Y_{1r}, Y_{21}, \dots, Y_{2r}, \dots, Y_{i1}, Y_{i2}, \dots, Y_{ij}, \dots, Y_{ir}, \dots, Y_{n1}, Y_{n2}, \dots, Y_{nr}$, sendo Y_{ij} a observação aleatória referente ao i -ésimo tratamento e a sua j -ésima repetição, $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, r$, em que a média do i -ésimo tratamento é dada por:

$$\bar{Y}_{i.} = \frac{\sum_{j=1}^r Y_{ij}}{r} = \frac{Y_{i.}}{r}, \quad (18)$$

sendo a variável aleatória $\bar{Y}_{i.}$ é normalmente distribuída, independente e homocedástica (variância σ^2), então a estatística Q , Definição 2.11, pode ser reescrita como

$$Q = \frac{\bar{Y}_{(n).} - \bar{Y}_{(1).}}{S_{\bar{Y}_{i.}}}, \quad (19)$$

em que $\bar{Y}_{(n).}$ e $\bar{Y}_{(1).}$ são as estatísticas de ordem, referentes a maior e a menor média em n tratamentos, respectivamente; $S_{\bar{Y}_{i.}} = \sqrt{QME/r}$ é o erro padrão para uma média amostral; e QME é o quadrado médio do resíduo.

Os testes de comparações múltiplas baseados na estatística Q comparam todas as médias duas a duas, ou seja, testa a hipótese $H_0 : \mu_i - \mu_{i'} = 0$ para $i \neq i' \in \{1, 2, \dots, n\}$. A decisão do teste é baseada na diferença mínima significativa (Δ) ao nível de significância α , que será definida posteriormente para cada teste apresentado. Assim, todos os contrastes $D = \bar{Y}_{i.} - \bar{Y}_{i'}$, que superam o valor de Δ , serão considerados significativamente diferentes de zero, e a hipótese $H_0 : \mu_i - \mu_{i'} = 0$ deve ser rejeitada naquele nível nominal de significância α estabelecido.

O primeiro teste apresentado é o teste de Tukey. Este teste requer que todos os níveis de tratamentos tenham o mesmo número de repetições, as médias envolvidas sejam independentes, e haja homogeneidade de variâncias entre os tratamentos envolvidos. Segundo Hinkelmann e Kempthorne (2007, p. 226), o teste de Tukey foi desenvolvido para controlar o erro tipo I por experimento. A seguir é apresentado o teste.

Proposição 2.1 — Teste de Tukey (1953). Considere n amostras normalmente distribuídas, independentes e homocedásticas, de tamanho r , e $\bar{Y}_{i.}$, a média da i -ésima amostra (tratamento), $i = 1, 2, \dots, n$. Então a hipótese $H_0 : \mu_i - \mu_{i'} = 0$ para $i \neq i' \in \{1, 2, \dots, n\}$, será testada comparando $|\bar{Y}_{i.} - \bar{Y}_{i'.}|$ com $\Delta = q_{(\alpha; n, \nu)} \sqrt{\frac{QME}{r}}$, em que $q_{(\alpha; n, \nu)}$ é o quantil superior a $100\alpha\%$ da distribuição da amplitude estudentizada

externamente, Definição (2.11). Se

$$|\bar{Y}_i - \bar{Y}_{i'}| \geq q_{(\alpha;n,\nu)} \sqrt{\frac{QME}{r}}, \quad (20)$$

a hipótese H_0 é rejeitada, isto é, μ_i e $\mu_{i'}$ são consideradas significativamente diferentes de zero, no nível de significância α . ■

Equivalentemente, pode-se obter o intervalo de confiança simultâneo $(1-\alpha)100\%$ para todas as comparações $\mu_i - \mu_{i'}$, da seguinte forma

$$(\bar{Y}_i - \bar{Y}_{i'}) \pm q_{(\alpha;n,\nu)} \sqrt{\frac{QME}{r}}. \quad (21)$$

O segundo teste de SNK foi desenvolvido por Student (1927), Newman (1939) e Keuls (1952). Este teste é similar ao teste de Tukey, com uma pequena diferença que leva em consideração o número de médias abrangidas no contraste entre médias ordenadas.

Proposição 2.2 — Teste de Student-Newman-Keuls (SNK). Considere n amostras normalmente distribuídas, independentes e homocedásticas, de tamanho r , e \bar{Y}_i , a média da i -ésima amostra (tratamento), $i = 1, 2, \dots, n$. Então a hipótese $H_0 : \mu_i - \mu_{i'} = 0$ para $i \neq i' \in \{1, 2, \dots, n\}$, será testada comparando $|\bar{Y}_i - \bar{Y}_{i'}|$ com $\Delta = q_{(\alpha;p,\nu)} \sqrt{\frac{QME}{r}}$, em que $q_{(\alpha;p,\nu)}$ é o quantil superior a $100\alpha\%$ da distribuição da amplitude estudentizada externamente, Definição 2.11, sendo $p = 2, 3, \dots, n$. Nesse caso, as médias devem ser ordenadas em primeiro lugar. Se

$$|\bar{Y}_i - \bar{Y}_{i'}| \geq q_{(\alpha;p,\nu)} \sqrt{\frac{QME}{r}}, \quad (22)$$

a hipótese H_0 é rejeitada, isto é, μ_i e $\mu_{i'}$ são consideradas estatisticamente diferentes, no nível nominal de significância α . ■

Apesar de usar o quantil da distribuição da amplitude estudentizada externamente, a diferença mínima significativa varia de acordo com o número p de médias abrangidas, e portanto, um total de $(n - 1)$ valores críticos.

O terceiro teste abordado baseado na amplitude estudentizada externamente, mas que não fará parte dos estudos desse trabalho, apresentado apenas pela sua importância, é o teste de Duncan. Esse teste sofre duras críticas, pois apresenta alta taxa de erro tipo I por experimento (O'NEILL; WETHERILL, 1971).

Proposição 2.3 — Teste de Duncan (1955). Considere n amostras normalmente distribuídas, independentes e homocedásticas, de tamanho r , e \bar{Y}_i , a média da i -ésima amostra (tratamento), $i = 1, 2, \dots, n$. Então a hipótese $H_0 : \mu_i - \mu_{i'} = 0$ para $i \neq i' \in \{1, 2, \dots, n\}$, será testada comparando $|\bar{Y}_i - \bar{Y}_{i'}|$ com $\Delta = q_{(\alpha_p; p, \nu)} \sqrt{\frac{QME}{r}}$, em que $q_{(\alpha_p; p, \nu)}$ é o quantil superior a $100\alpha_p\%$ da distribuição da amplitude estudentizada externamente, Definição 2.11 e expressão (17), e $\alpha_p = 1 - (1 - \alpha)^{p-1}$, sendo $p = 2, 3, \dots, n$. Se

$$|\bar{Y}_i - \bar{Y}_{i'}| \geq q_{(\alpha_p; p, \nu)} \sqrt{\frac{QME}{r}}, \quad (23)$$

a hipótese H_0 é rejeitada, isto é, μ_i e $\mu_{i'}$ são consideradas estatisticamente diferentes, no nível nominal de significância α . ■

O teste de Duncan tem uma pequena modificação quanto ao teste de SNK. No cálculo do quantil da distribuição da amplitude estudentizada externamente, o nível de significância é variável de acordo com o número de médias abrangidas, proporcionando um total de $(n - 1)$ valores críticos. Além disso, o valor do quantil da amplitude estudentizada externamente varia conforme o número de p médias abrangidas pelo contraste.

2.1.3 Teste de Scott-Knott

O teste criado por Scott e Knott (1974), tem como base o uso da razão de verossimilhança para testar a existência de mais de um grupo de médias. Esse teste tem vantagem em relação aos PCMs descritos, a de não apresentar ambiguidade em seus resultados. Assim, seguindo o algoritmo apresentado em Ramalho, Ferreira e Oliveira (2005, p. 93-94) e acrescentando algumas atualizações para a aplicação do teste Scott-Knott, tem-se:

1. as médias dos tratamentos devem ser ordenadas: $\bar{Y}_{(1).}, \bar{Y}_{(2).}, \dots, \bar{Y}_{(n).}$. Nessa situação, o número de partições é $n - 1$ (FISHER, 1958, p. 796). Inicialmente o número de tratamentos é $g = n$;
2. determinar a partição entre dois grupos que maximize a soma de quadrados entre grupos, como sugerido por Edwards e Cavalli-Sforza (1965, p. 364). Essa soma de quadrados será definida por β_0 . Sejam T_1 e T_2 os totais dos dois grupos com k_1 e

k_2 tratamentos em cada um. Assim, β_0 é estimado da seguinte forma:

$$\beta_0 = \frac{T_1^2}{k_1} + \frac{T_2^2}{k_2} - \frac{(T_1 + T_2)^2}{k_1 + k_2},$$

sendo

$$T_1 = \sum_{i=1}^{k_1} \bar{Y}_{(i)} \quad \text{e} \quad T_2 = \sum_{i=k_1+1}^g \bar{Y}_{(i)},$$

em que $\bar{Y}_{(i)}$ é a média do tratamento da posição ordenada i . Os dois grupos deverão ser identificados por meio da inspeção das somas de quadrados das $g - 1$ partições possíveis, sendo g o número de tratamentos envolvidos no grupo de médias considerado;

3. determinar o valor da estatística λ da seguinte forma:

$$\lambda = \frac{\pi\beta_0}{2\hat{\sigma}_0^2(\pi - 2)},$$

em que $\hat{\sigma}_0^2$ é o estimador de máxima verossimilhança de σ_Y^2 . Seja $S_Y^2 = QME/r$ o estimador não viesado de σ_Y^2 e ν graus de liberdade associados a este estimador. Então,

$$\hat{\sigma}_0^2 = \frac{1}{g + \nu} \left[\sum_{i=1}^g (\bar{Y}_{(i)} - \bar{Y})^2 + \nu S_Y^2 \right].$$

4. Se $\lambda \geq \chi_{(\alpha; g/(\pi-2))}^2$, rejeita-se a hipótese de que os dois grupos são idênticos em favor da hipótese alternativa de que os dois grupos diferem;
5. No caso de rejeitar essa hipótese, os dois subgrupos formados serão independentemente submetidos aos passos (1) a (3), fazendo respectivamente $g = k_1$ e $g = k_2$. O processo em cada subgrupo se encerra ao não rejeitar H_0 no passo (4) ou se cada subgrupo contiver apenas uma média.

2.2 Distribuição da *Midrange* estudentizada externamente

Uma estatística também muito interessante é a *midrange*, pois segundo Rider (1957, p. 541), é mais eficiente (estimador com menor variância para a média populacional) que a média aritmética em algumas populações, tais como população cosseno, população parabólica, população retangular e população parabólica invertida. Com base

nessa informação, percebeu-se que essa estatística poderia trazer bons resultados para métodos de comparações múltiplas.

Na literatura, alguns trabalhos sobre a *midrange* foram publicados. Gumbel (1958), David e Nagaraja (2003), entre outros, obtiveram a função de distribuição e função densidade da *midrange* para o caso de uma população com distribuição normal. Batista (2012) e Ferreira (2013) obtiveram estas funções para o caso da *midrange* estudentizada externamente. Como consequência desses trabalhos, contruiu-se um pacote R, denominado SMR (BATISTA; FERREIRA, 2014a), com a implementação dos algoritmos do pacote publicados em Batista e Ferreira (2014b), que calcula a função de distribuição, função densidade, e retorna os valores dos quantis para a distribuição dessa estatística.

Como a distribuição da *midrange* será importante para os estudos desse trabalho, serão mostrados os principais resultados de Batista (2012). Sejam $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$, as estatísticas de ordem de uma amostra aleatória Y_1, Y_2, \dots, Y_n de tamanho n com função de distribuição $F_Y(y)$ e função densidade $f_Y(y)$. Então define-se

Definição 2.12 — Midrange. A *midrange* é definida pela média entre a menor ($Y_{(1)}$) e a maior ($Y_{(n)}$) estatística de ordem em uma amostra aleatória de tamanho n , dada por

$$\bar{R} = \frac{Y_{(1)} + Y_{(n)}}{2}. \quad (24)$$

A função de distribuição da *midrange* é apresentada no seguinte Teorema,

Teorema 2.2 — Distribuição da *midrange*. A função densidade e a função de distribuição de \bar{R} de uma amostra aleatória Y_1, Y_2, \dots, Y_n de tamanho n , em que Y_i tem função de distribuição $F_Y(y)$ e função densidade $f_Y(y)$ para todo $i, i = 1, 2, \dots, n$ são

$$f_{\bar{R}}(\bar{r}) = \int_{-\infty}^{\bar{r}} 2n(n-1)f_Y(z)f_Y(2\bar{r}-z)[F_Y(2\bar{r}-z) - F_Y(z)]^{n-2}dz, \quad (25)$$

e

$$F_{\bar{R}}(\bar{r}) = n \int_{-\infty}^{\bar{r}} f_Y(z)[F_Y(2\bar{r}-z) - F_Y(z)]^{n-1}dz, \quad (26)$$

respectivamente. ■

Demonstração. Seja a distribuição conjunta de $Y_{(1)}$ e $Y_{(n)}$, segundo David e Nagaraja (2003, p. 12), dada por

$$f_{Y_{(1)}, Y_{(n)}}(u, x) = n(n-1)f_Y(u)f_Y(x)[F_Y(x) - F_Y(u)]^{n-2}, \quad u < x. \quad (27)$$

Então, para obter a distribuição de \bar{R} são consideradas as seguintes transformações $Z = Y_{(1)} = g_1^{-1}(\bar{R}, Z)$ e $\bar{R} = Y_{(n)} = g_2^{-1}(\bar{R}, Z)$. As funções inversas dessas transformações são $Y_{(1)} = g_1^{-1}(\bar{R}, Z) = Z$ e $Y_{(n)} = g_2^{-1}(\bar{R}, Z) = 2\bar{R} - Z$. O Jacobiano dessas transformações é

$$J = \begin{vmatrix} \frac{\partial y_{(1)}}{\partial z} & \frac{\partial y_{(1)}}{\partial \bar{r}} \\ \frac{\partial y_{(n)}}{\partial z} & \frac{\partial y_{(n)}}{\partial \bar{r}} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix} = 2.$$

Portanto, a função densidade conjunta de \bar{R} e Z , é dada por

$$\begin{aligned} f_{\bar{R}, Z}(\bar{r}, z) &= |J|f_{Y_{(1)}, Y_{(n)}}(g_1^{-1}(\bar{r}, z), g_2^{-1}(\bar{r}, z)), \\ &= 2n(n-1)f_Y(z)f_Y(2\bar{r} - z)[F_Y(2\bar{r} - z) - F_Y(z)]^{n-2}. \end{aligned}$$

A função densidade desejada $f_{\bar{R}}(\bar{r})$ pode ser obtida pela integração da densidade conjunta em relação a z , resultando em

$$f_{\bar{R}}(\bar{r}) = \int_{-\infty}^{\bar{r}} 2n(n-1)f_Y(z)f_Y(2\bar{r} - z)[F_Y(2\bar{r} - z) - F_Y(z)]^{n-2} dz.$$

A função de distribuição de \bar{R} pode ser obtida por

$$\begin{aligned} F_{\bar{R}}(\bar{r}) &= \int_{-\infty}^{\bar{r}} \int_{-\infty}^{\bar{r}} 2n(n-1)f_Y(z)f_Y(2t - z) \\ &\quad [F_Y(2t - z) - F_Y(z)]^{n-2} dz dt. \end{aligned} \quad (28)$$

A simplificação da equação (28) pode ser desenvolvida pela inversão da ordem das integrais. Como há uma dependência entre \bar{R} e Z , fixou-se z , e t variará no intervalo $[z, \bar{r}]$. Portanto, os resultados da mudança da ordem de integração é

$$F_{\bar{R}}(\bar{r}) = \int_{-\infty}^{\bar{r}} \left[\int_z^{\bar{r}} 2n(n-1)f_Y(z)f_Y(2t - z) \times \right.$$

$$\times [F_Y(2t - z) - F_Y(z)]^{n-2} dt] dz.$$

Note que

$$\frac{d}{dt} [F_Y(2t - z) - F_Y(z)]^{n-1} = 2(n-1)f_Y(2t - z)[F_Y(2t - z) - F_Y(z)]^{n-2}.$$

Assim,

$$F_{\bar{R}}(\bar{r}) = \int_{-\infty}^{\bar{r}} n f_Y(z) \{ [F_Y(2t - z) - F_Y(z)]^{n-1} \}_{t=\bar{r}}^{t=z} dz.$$

Como,

$$\begin{aligned} \{ [F_Y(2t - z) - F_Y(z)]^{n-1} \}_{t=\bar{r}}^{t=z} &= [F_Y(2\bar{r} - z) - F_Y(z)]^{n-1} - [F_Y(2z - z) - F_Y(z)]^{n-1}, \\ &= [F_Y(2\bar{r} - z) - F_Y(z)]^{n-1}, \end{aligned}$$

o resultado da função de distribuição acumulada $F_{\bar{R}}(\bar{r})$ é

$$F_{\bar{R}}(\bar{r}) = n \int_{-\infty}^{\bar{r}} f_Y(z) [F_Y(2\bar{r} - z) - F_Y(z)]^{n-1} dz,$$

como mostrado em Gumbel (1958, p. 108). ■

Um importante Teorema, apresentado a seguir, mostra a equivalência de populações normais. Estas equivalências serão importantes para a determinação da distribuição das estatísticas \bar{R} , \bar{W} e \bar{Q} .

Teorema 2.3 — Identidades da distribuição normal. Considere uma variável aleatória Y de uma população normal com parâmetros μ e σ^2 , em que $\mu \in \mathbb{R}$ e $0 < \sigma^2 < \infty$. A função densidade e a função de distribuição de Y são dadas por:

$$\phi_{\mu, \sigma^2}(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}},$$

e

$$\Phi_{\mu, \sigma^2}(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt,$$

respectivamente. Assim, valem as equivalências:

(a) 1ª equivalência:

$$\phi_{\mu,\sigma^2}(y) = \frac{1}{\sigma} \phi_{\frac{\mu}{\sigma},1}(y/\sigma) \quad (29)$$

e

$$\Phi_{\mu,\sigma^2}(y) = \Phi_{\frac{\mu}{\sigma},1}(y/\sigma); \quad (30)$$

(b) 2ª equivalência:

$$\phi_{\frac{\mu}{\sigma},1}(y) = \phi(y - \mu/\sigma), \quad (31)$$

sendo $\phi(y)$ a função densidade da normal padrão, isto é, $Y \sim N(0,1)$, e

$$\Phi_{\frac{\mu}{\sigma},1}(y) = \Phi(y - \mu/\sigma), \quad (32)$$

sendo $\Phi(y)$ a função distribuição da normal padrão, isto é, $Y \sim N(0,1)$;

(b) 3ª equivalência:

$$\phi_{0,\sigma^2}(y) = \frac{1}{\sigma} \phi(y/\sigma) \quad (33)$$

e

$$\Phi_{0,\sigma^2}(y) = \Phi(y/\sigma). \quad (34)$$

■

Demonstração. A prova será apresentada para cada equivalência.

(a) 1ª equivalência

Basta mostrar que $\frac{1}{\sigma} \phi_{\frac{\mu}{\sigma},1}(y/\sigma) = \phi_{\mu,\sigma^2}(y)$. Logo,

$$\frac{1}{\sigma} \phi_{\frac{\mu}{\sigma},1}(y/\sigma) = \frac{1}{\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y/\sigma - \mu/\sigma)^2}{2}}$$

$$= \underbrace{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}}_{\phi_{\mu,\sigma^2}(y)}.$$

Seja a função de distribuição

$$\begin{aligned}\Phi_{\mu,\sigma^2}(y) &= \int_{-\infty}^y \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \\ &= \int_{-\infty}^y \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t/\sigma-\mu/\sigma)^2}{2}} dt\end{aligned}$$

então fazendo $x = t/\sigma$, tem-se que $dx = dt/\sigma$, logo,

$$\begin{aligned}\Phi_{\mu,\sigma^2}(y) &= \int_{-\infty}^{y/\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu/\sigma)^2}{2}} dx \\ &= \Phi_{\frac{\mu}{\sigma},1}(y/\sigma).\end{aligned}$$

(b) 2ª equivalência

Observe que,

$$\begin{aligned}\phi(y - \mu/\sigma) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-\mu/\sigma)^2}{2}} \\ &= \phi_{\frac{\mu}{\sigma},1}(y).\end{aligned}$$

Para a equivalência entre as funções distribuição, tem-se que

$$\Phi_{\frac{\mu}{\sigma},1}(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-\mu/\sigma)^2}{2}} dt.$$

Fazendo $x = t - \mu/\sigma$, tem-se que $dx = dt$. Então,

$$\begin{aligned}\Phi_{\frac{\mu}{\sigma},1}(y) &= \int_{-\infty}^{y-\mu/\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x+\mu/\sigma-\mu/\sigma)^2}{2}} dx \\ &= \int_{-\infty}^{y-\mu/\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \Phi(y - \mu/\sigma).\end{aligned}$$

(c) 3ª equivalência

Basta mostrar que $\frac{1}{\sigma}\phi(y/\sigma) = \phi_{0,\sigma^2}(y)$. Logo,

$$\begin{aligned}\frac{1}{\sigma}\phi(z/\sigma) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(z/\sigma)^2}{2}\right\} \\ &= \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{z^2}{2\sigma^2}\right\}}_{\phi_{0,\sigma^2}}.\end{aligned}$$

Nessa mesma linha de raciocínio, a relação entre as funções de distribuição $\Phi_{0,\sigma}(y)$ e $\Phi(y/\sigma)$ é verificada a seguir. Sendo

$$\Phi(z/\sigma) = \int_{-\infty}^{z/\sigma} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2}\right\} dy.$$

Fazendo $y = t/\sigma \Rightarrow dy = dt/\sigma$, então

$$\begin{aligned}\Phi(z/\sigma) &= \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2\sigma^2}\right\} \frac{dt}{\sigma} \\ &= \int_{-\infty}^z \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{t^2}{2\sigma^2}\right\} dt \\ \Phi(z/\sigma) &= \Phi_{0,\sigma^2}(z).\end{aligned}$$

■

Corolário 2.1 No Teorema 2.2, se $Y_i \sim N(\mu, \sigma^2)$, para $i = 1, 2, \dots, n$, então a função densidade e a função de distribuição de \bar{R} são

$$\begin{aligned}f_{\bar{R}}(\bar{r}; \mu, \sigma^2) &= \int_{-\infty}^{\bar{r}} 2n(n-1)\phi_{\mu, \sigma^2}(z)\phi_{\mu, \sigma^2}(2\bar{r}-z) \\ &\quad [\Phi_{\mu, \sigma^2}(2\bar{r}-z) - \Phi_{\mu, \sigma^2}(z)]^{n-2} dz, \quad (35)\end{aligned}$$

e

$$F_{\bar{R}}(\bar{r}; \mu, \sigma^2) = n \int_{-\infty}^{\bar{r}} \phi_{\mu, \sigma^2}(z) [\Phi_{\mu, \sigma^2}(2\bar{r}-z) - \Phi_{\mu, \sigma^2}(z)]^{n-1} dz, \quad (36)$$

respectivamente. ■

Demonstração. Considerando $Y_i \sim N(\mu, \sigma^2)$ para $i = 1, 2, \dots, n$, então

$$\phi_{\mu, \sigma^2}(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y - \mu)^2}{2\sigma^2}\right\}, \quad -\infty \leq y \leq \infty,$$

e

$$\Phi_{\mu, \sigma^2}(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(t - \mu)^2}{2\sigma^2}\right\} dt, \quad -\infty \leq y \leq \infty,$$

são respectivamente, a função densidade e função de distribuição de uma normal com média μ e variância σ^2 . Assim, de modo trivial, as expressões (25) e (26), pelo Teorema 2.2, podem ser reescritas como

$$f_{\bar{R}}(\bar{r}; \mu, \sigma^2) = \int_{-\infty}^{\bar{r}} 2n(n-1)\phi_{\mu, \sigma^2}(z)\phi_{\mu, \sigma^2}(2\bar{r} - z) [\Phi_{\mu, \sigma^2}(2\bar{r} - z) - \Phi_{\mu, \sigma^2}(z)]^{n-2} dz,$$

e

$$F_{\bar{R}}(\bar{r}; \mu, \sigma^2) = n \int_{-\infty}^{\bar{r}} \phi_{\mu, \sigma^2}(z) [\Phi_{\mu, \sigma^2}(2\bar{r} - z) - \Phi_{\mu, \sigma^2}(z)]^{n-1} dz,$$

respectivamente. ■

Outra importante estatística é a *midrange* padronizada, \bar{W} . Esta estatística é a padronização da *midrange*, sendo definida a seguir.

Definição 2.13 — Midrange padronizada. A *midrange* padronizada é definida por

$$\bar{W} = \frac{\bar{R}}{\sigma}, \quad (37)$$

em que \bar{R} é a *midrange* da Definição 2.12 e σ é o desvio padrão populacional. ■

No Teorema a seguir serão apresentadas as funções densidade e funções de distribuição de \bar{W} .

Teorema 2.4 — Distribuição da *midrange* padronizada \bar{W} . Seja uma amostra aleatória tal que $Y_i \sim N(\mu, \sigma^2)$, para $i = 1, 2, \dots, n$, então as funções densidade e funções de distribuição equivalentes de \bar{W} , Definição 2.13, são

i) expressas em termos de $\phi_{\mu,\sigma^2}(y)$ e $\Phi_{\mu,\sigma^2}(y)$:

$$f_{\bar{W}}(\bar{w}; \mu, \sigma^2) = \sigma^2 \int_{-\infty}^{\bar{w}} 2n(n-1) \phi_{\mu,\sigma^2}(y\sigma) \phi_{\mu,\sigma^2}(2\bar{w}\sigma - y\sigma) \times \\ \times [\Phi_{\mu,\sigma^2}(2\bar{w}\sigma - y\sigma) - \Phi_{\mu,\sigma^2}(y\sigma)]^{n-2} dy, \quad (38)$$

e

$$F_{\bar{W}}(\bar{w}; \mu, \sigma^2) = \sigma \int_{-\infty}^{\bar{w}} n \phi_{\mu,\sigma^2}(y\sigma) [\Phi_{\mu,\sigma^2}(2t\sigma - y\sigma) - \Phi_{\mu,\sigma^2}(y\sigma)]^{n-1} dy; \quad (39)$$

ii) expressas em termos de $\phi_{\frac{\mu}{\sigma},1}(y)$ e $\Phi_{\frac{\mu}{\sigma},1}(y)$:

$$f_{\bar{W}}(\bar{w}; \mu, \sigma^2) = \int_{-\infty}^{\bar{w}} 2n(n-1) \phi_{\frac{\mu}{\sigma},1}(y) \phi_{\frac{\mu}{\sigma},1}(2\bar{w} - y) \times \\ \times [\Phi_{\frac{\mu}{\sigma},1}(2\bar{w} - y) - \Phi_{\frac{\mu}{\sigma},1}(y)]^{n-2} dy, \quad (40)$$

e

$$F_{\bar{W}}(\bar{w}; \mu, \sigma^2) = \int_{-\infty}^{\bar{w}} n \phi_{\frac{\mu}{\sigma},1}(y) [\Phi_{\frac{\mu}{\sigma},1}(2\bar{w} - y) - \Phi_{\frac{\mu}{\sigma},1}(y)]^{n-1} dy; \quad (41)$$

iii) expressas em termos de $\phi(y)$ e $\Phi(y)$:

$$f_{\bar{W}}(\bar{w}; \mu, \sigma^2) = \int_{-\infty}^{\bar{w}} 2n(n-1) \phi(y - \mu/\sigma) \phi(2\bar{w} - y - \mu/\sigma) \times \\ \times [\Phi(2\bar{w} - y - \mu/\sigma) - \Phi(y - \mu/\sigma)]^{n-2} dy, \quad (42)$$

e

$$F_{\bar{W}}(\bar{w}; \mu, \sigma^2) = \int_{-\infty}^{\bar{w}} n \phi(y - \mu/\sigma) [\Phi(2\bar{w} - y - \mu/\sigma) - \Phi(y - \mu/\sigma)]^{n-1} dy. \quad (43)$$

■

Demonstração. Para a prova de (i), se $\bar{W} = \bar{R}/\sigma$, então

$$f_{\bar{W}}(\bar{w}; \mu, \sigma^2) = f_{\bar{R}}(\sigma\bar{w}; \mu, \sigma^2) \left| \frac{d\bar{r}}{d\bar{w}} \right|$$

$$\begin{aligned}
&= \sigma \int_{-\infty}^{\bar{w}\sigma} 2n(n-1)\phi_{\mu,\sigma^2}(z)\phi_{\mu,\sigma^2}(2\bar{w}\sigma - z) \times \\
&\quad \times [\Phi_{\mu,\sigma^2}(2\bar{w}\sigma - z) - \Phi_{\mu,\sigma^2}(z)]^{n-2} dz.
\end{aligned}$$

Fazendo $Y = Z/\sigma$, tem-se que $dz = \sigma dy$. Assim,

$$\begin{aligned}
f_{\bar{W}}(\bar{w}; \mu, \sigma^2) &= \sigma^2 \int_{-\infty}^{\bar{w}} 2n(n-1)\phi_{\mu,\sigma^2}(y\sigma)\phi_{\mu,\sigma^2}(2\bar{w}\sigma - y\sigma) \times \\
&\quad \times [\Phi_{\mu,\sigma^2}(2\bar{w}\sigma - y\sigma) - \Phi_{\mu,\sigma^2}(y\sigma)]^{n-2} dy.
\end{aligned}$$

Tem-se, então, a prova para a função densidade da parte (i). A função de distribuição é

$$\begin{aligned}
F_{\bar{W}}(\bar{w}; \mu, \sigma^2) &= \int_{-\infty}^{\bar{w}} f_{\bar{W}}(t; \mu, \sigma^2) dt \\
&= \sigma^2 \int_{-\infty}^{\bar{w}} \int_{-\infty}^{\bar{w}} 2n(n-1)\phi_{\mu,\sigma^2}(y\sigma)\phi_{\mu,\sigma^2}(2t\sigma - y\sigma) \times \\
&\quad \times [\Phi_{\mu,\sigma^2}(2t\sigma - y\sigma) - \Phi_{\mu,\sigma^2}(y\sigma)]^{n-2} dy dt.
\end{aligned}$$

Usando a inversão de integrais, observa-se que \bar{W} e Y são estatísticas padronizadas, e portanto estão em mesma escala de unidades. Percebe-se ainda que estas estatísticas são dependentes. Assim, com a inversão, \bar{W} variará no intervalo $[y, \bar{w}]$. Dessa forma, segue que

$$\begin{aligned}
F_{\bar{W}}(\bar{w}; \mu, \sigma^2) &= \sigma^2 \int_{-\infty}^{\bar{w}} \int_y^{\bar{w}} 2n(n-1)\phi_{\mu,\sigma^2}(y\sigma)\phi_{\mu,\sigma^2}(2t\sigma - y\sigma) \times \\
&\quad \times [\Phi_{\mu,\sigma^2}(2t\sigma - y\sigma) - \Phi_{\mu,\sigma^2}(y\sigma)]^{n-2} dt dy.
\end{aligned}$$

Sabendo que

$$\begin{aligned}
\frac{d}{dt} [\Phi_{\mu,\sigma^2}(2t\sigma - y\sigma) - \Phi_{\mu,\sigma^2}(y\sigma)]^{n-1} &= 2\sigma(n-1)\phi_{\mu,\sigma^2}(2t\sigma - y\sigma) \times \\
&\quad \times [\Phi_{\mu,\sigma^2}(2t\sigma - y\sigma) - \Phi_{\mu,\sigma^2}(y\sigma)]^{n-2},
\end{aligned}$$

logo,

$$F_{\bar{W}}(\bar{w}; \mu, \sigma^2) = \sigma \int_{-\infty}^{\bar{w}} n\phi_{\mu,\sigma^2}(y\sigma)[\Phi_{\mu,\sigma^2}(2\bar{w}\sigma - y\sigma) - \Phi_{\mu,\sigma^2}(y\sigma)]^{n-1} dy.$$

Assim, conclui-se o item (i). A prova de (ii) é consequência de (i), usando as equivalências do Teorema 2.3. Usando a 1ª equivalência deste Teorema, expressões (29) e (30), pode-se obter a função densidade e função de distribuição, respectivamente, das expressões (40) e (41). Assim, aplicando a equivalência da expressão (29) em (38), logo

$$\begin{aligned} f_{\bar{W}}(\bar{w}; \mu, \sigma^2) &= \sigma^2 \int_{-\infty}^{\bar{w}} 2n(n-1) \frac{\phi_{\frac{\mu}{\sigma}, 1}(y)}{\sigma} \frac{\phi_{\frac{\mu}{\sigma}, 1}(2\bar{w}-y)}{\sigma} \times \\ &\quad \times [\Phi_{\frac{\mu}{\sigma}, 1}(2\bar{w}-y) - \Phi_{\frac{\mu}{\sigma}, 1}(y)]^{n-2} dy \\ &= \int_{-\infty}^{\bar{w}} 2n(n-1) \phi_{\frac{\mu}{\sigma}, 1}(y) \phi_{\frac{\mu}{\sigma}, 1}(2\bar{w}-y) \times \\ &\quad \times [\Phi_{\frac{\mu}{\sigma}, 1}(2\bar{w}-y) - \Phi_{\frac{\mu}{\sigma}, 1}(y)]^{n-2} dy \end{aligned}$$

e a função densidade de (ii) é obtida. A função de distribuição de (ii) é desenvolvida seguindo a mesma linha de raciocínio da que foi obtida em (i). Assim,

$$F_{\bar{W}}(\bar{w}; \mu, \sigma^2) = \int_{-\infty}^{\bar{w}} n \phi_{\frac{\mu}{\sigma}, 1}(y) [\Phi_{\frac{\mu}{\sigma}, 1}(2\bar{w}-y) - \Phi_{\frac{\mu}{\sigma}, 1}(y)]^{n-1} dy.$$

Por fim, as funções densidade e de distribuição de (iii) podem ser obtidas, respectivamente, usando os resultados de (40) e (41) com a 2ª equivalência do Teorema 2.3, expressões (31) e (32). Assim, aplicando (31) em (40), logo

$$\begin{aligned} f_{\bar{W}}(\bar{w}; \mu, \sigma^2) &= \int_{-\infty}^{\bar{w}} 2n(n-1) \phi(y - \mu/\sigma) \phi(2\bar{w} - y - \mu/\sigma) \times \\ &\quad \times [\Phi(2\bar{w} - y - \mu/\sigma) - \Phi(y - \mu/\sigma)]^{n-2} dy \end{aligned}$$

e a função densidade de (iii) foi provada. Para obter a função de distribuição de (iii), basta usar a mesma lógica do que foi feito para obter a função de distribuição de (i). Logo,

$$F_{\bar{W}}(\bar{w}; \mu, \sigma^2) = \int_{-\infty}^{\bar{w}} n \phi(y - \mu/\sigma) [\Phi(2\bar{w} - y - \mu/\sigma) - \Phi(y - \mu/\sigma)]^{n-1} dy.$$

■

Observe que as funções densidade e funções de distribuição apresentadas no Teorema 2.4 são equivalentes. Por exemplo, se $Y_i \sim N(0, \sigma^2)$, estas resultam na mesma

função densidade e na mesma função de distribuição, isto é,

$$f_{\bar{W}}(\bar{w}) = \int_{-\infty}^{\bar{w}} 2n(n-1)\phi(y)\phi(2\bar{w}-y)[\Phi(2\bar{w}-y) - \Phi(y)]^{n-2} dy,$$

e

$$F_{\bar{W}}(\bar{w}) = \int_{-\infty}^{\bar{w}} n\phi(y)[\Phi(2\bar{w}-y) - \Phi(y)]^{n-1} dy.$$

A única situação em que a distribuição de \bar{W} passa a não depender dos parâmetros da distribuição inicial, sendo esta uma distribuição normal, é quando $\mu = 0$ independente do valor de σ^2 .

Com o objetivo de obter a distribuição de \bar{Q} , define-se

Definição 2.14 — *Midrange* estudentizada externamente. A *midrange* estudentizada externamente é definida por

$$\bar{Q} = \frac{\bar{R}}{S},$$

em que S é um estimador do desvio padrão populacional σ com ν graus de liberdade obtido independentemente de \bar{R} (Definição 2.12). ■

Observe que a *midrange* estudentizada externamente \bar{Q} (Definição 2.14), pode ser redefinida por

$$\bar{Q} = \frac{\bar{R}}{S} = \frac{\bar{R}/\sigma}{S/\sigma} = \frac{\bar{W}}{X}, \quad (44)$$

em que \bar{W} está expressa na Definição 2.13, X é definido no Teorema 2.1, e S é o estimador do desvio padrão populacional, σ . \bar{W} e S com ν graus de liberdade, são distribuídos independentemente. Isso ocorre quando \bar{W} é relacionado às médias de um fator com Isso ocorre quando, por exemplo, \bar{W} é proveniente de uma amostra aleatória normal de tamanho n , e S , o desvio padrão, obtido em outra amostra aleatória normal de tamanho $\nu + 1$. O n níveis e $S = \sqrt{QME/r}$, sendo QME o quadrado médio do erro experimental com ν graus de liberdade e r é o número de repetições, fato bem documentado na literatura (SEARLE, 1987).

Usar a relação expressa em (44) foi o meio para obter a função densidade e função distribuição de \bar{Q} , já que as funções densidade de \bar{W} e X foram demonstradas nos Teoremas 2.4 e 2.1, respectivamente. Dessa forma, a distribuição de \bar{Q} é apresentada a seguir.

Teorema 2.5 — Distribuição da *midrange* estudentizada externamente \bar{Q} . As funções densidade e funções de distribuição de \bar{Q} , da Definição 2.14, considerando uma amostra aleatória $Y_i \sim N(\mu, \sigma^2)$, para $i = 1, 2, \dots, n$, são

i) expressas em termos de $\phi_{\mu, \sigma^2}(y)$ e $\Phi_{\mu, \sigma^2}(y)$:

$$f_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) = \sigma^2 \int_0^\infty \int_{-\infty}^{x\bar{q}} 2n(n-1)x\phi_{\mu, \sigma^2}(y\sigma)\phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) \times \\ \times [\Phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)]^{n-2} f_X(x; \nu) dy dx, \quad (45)$$

e

$$F_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) = \sigma \int_0^\infty \int_{-\infty}^{x\bar{q}} n\phi_{\mu, \sigma^2}(y\sigma) [\Phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) - \\ - \Phi_{\mu, \sigma^2}(y\sigma)]^{n-1} f_X(x; \nu) dy dx; \quad (46)$$

ii) expressas em termos de $\phi_{\frac{\mu}{\sigma}, 1}(y)$ e $\Phi_{\frac{\mu}{\sigma}, 1}(y)$:

$$f_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) = \int_0^\infty \int_{-\infty}^{x\bar{q}} 2n(n-1)x\phi_{\frac{\mu}{\sigma}, 1}(y)\phi_{\frac{\mu}{\sigma}, 1}(2x\bar{q} - y) \times \\ \times [\Phi_{\frac{\mu}{\sigma}, 1}(2x\bar{q} - y) - \Phi_{\frac{\mu}{\sigma}, 1}(y)]^{n-2} f_X(x; \nu) dy dx, \quad (47)$$

e

$$F_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) = \int_0^\infty \int_{-\infty}^{x\bar{q}} n\phi_{\frac{\mu}{\sigma}, 1}(y) [\Phi_{\frac{\mu}{\sigma}, 1}(2x\bar{q} - y) - \Phi_{\frac{\mu}{\sigma}, 1}(y)]^{n-1} \times \\ \times f_X(x; \nu) dy dx; \quad (48)$$

iii) expressas em termos de $\phi(y)$ e $\Phi(y)$:

$$f_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) = \int_0^\infty \int_{-\infty}^{x\bar{q}} 2n(n-1)x\phi(y - \mu/\sigma)\phi(2x\bar{q} - y - \mu/\sigma) \times \\ \times [\Phi(2x\bar{q} - y - \mu/\sigma) - \Phi(y - \mu/\sigma)]^{n-2} f_X(x; \nu) dy dx, \quad (49)$$

e

$$F_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) = \int_0^\infty \int_{-\infty}^{x\bar{q}} n\phi(y - \mu/\sigma)[\Phi(2x\bar{q} - y - \mu/\sigma) - \Phi(y - \mu/\sigma)]^{n-1} f_X(x; \nu) dy dx. \quad (50)$$

■

Demonstração. Para a prova (i), observe que a distribuição de $Q = \bar{R}/S = \bar{W}/X$ pode ser obtida por meio da distribuição conjunta de $\bar{W} = \bar{R}/\sigma$ e $X = S/\sigma$, realizando algumas transformações de variáveis convenientemente. Como \bar{W} e X são independentemente distribuídas, tem-se que a densidade conjunta é o produto das densidades marginais e é dada por

$$f_{\bar{W},X}(\bar{w}, x; \mu, \sigma^2, n, \nu) = \sigma^2 \int_{-\infty}^{\bar{w}} 2n(n-1)\phi_{\mu, \sigma^2}(y)\phi_{\mu, \sigma^2}(2\bar{w}\sigma - y\sigma) \times [\Phi_{\mu, \sigma^2}(2\bar{w}\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)]^{n-2} f_X(x; \nu) dy, \quad (51)$$

em que $f_X(x; \nu)$ é dada na expressão (14).

Utilizando as transformações de variáveis $\bar{Q} = \bar{W}/X$ e $X = X$, tem-se que o jacobiano da transformação é

$$J = \begin{vmatrix} \frac{\partial \bar{w}}{\partial \bar{q}} & \frac{\partial \bar{w}}{\partial x} \\ \frac{\partial x}{\partial \bar{q}} & \frac{\partial x}{\partial x} \end{vmatrix} = \begin{vmatrix} x & \bar{q} \\ 0 & 1 \end{vmatrix} = x, \quad x > 0.$$

Como $\bar{W} = X\bar{Q}$ e observando a distribuição conjunta de \bar{W} e X , dada em (51), tem-se

$$\begin{aligned} f_{\bar{Q},X}(\bar{q}, x; \mu, \sigma^2, n, \nu) &= f_{\bar{W},X}(x\bar{q}, x; \mu, \sigma^2, n, \nu) |J|, \\ &= \sigma^2 \int_{-\infty}^{x\bar{q}} 2n(n-1)\phi_{\mu, \sigma^2}(y\sigma)\phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) \times \\ &\quad \times [\Phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)]^{n-2} f_X(x; \nu) |x| dy, \end{aligned}$$

que se for integrada em relação a x , definida no domínio de 0 a ∞ , fornece a função densidade marginal de \bar{Q} . Portanto,

$$f_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) = \sigma^2 \int_0^\infty \int_{-\infty}^{x\bar{q}} 2n(n-1)x\phi_{\mu, \sigma^2}(y\sigma)\phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) \times [\Phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)]^{n-2} f_X(x; \nu) dy dx. \quad (52)$$

Para encontrar a função de distribuição de \bar{Q} , basta integrar a expressão (52) ao domínio de \bar{Q} , definida no intervalo $(-\infty, \bar{q}]$. Realizando ainda uma troca de variável para t , como artifício de integração, tem-se que a função de distribuição de \bar{Q} é dada por

$$F_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) = \sigma^2 \int_{-\infty}^{\bar{q}} \int_0^{\infty} \int_{-\infty}^{x\bar{q}} 2n(n-1)x\phi_{\mu, \sigma^2}(y\sigma)\phi_{\mu, \sigma^2}(2xt\sigma - y\sigma) \times \\ \times [\Phi_{\mu, \sigma^2}(2xt\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)]^{n-2} f_X(x; \nu) dy dx dt.$$

Fazendo a troca da ordem das integrais e fixando a menor estatística de ordem estudentizada, o limite inferior com relação a \bar{Q} passa a ser y/x . Assim,

$$F_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) = \sigma^2 \int_0^{\infty} \int_{-\infty}^{x\bar{q}} n\phi_{\mu, \sigma^2}(y\sigma) \left\{ \int_{y/x}^{\bar{q}} 2x(n-1)\phi_{\mu, \sigma^2}(2xt\sigma - y\sigma) \times \right. \\ \left. \times [\Phi_{\mu, \sigma^2}(2xt\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)]^{n-2} dt \right\} f_X(x; \nu) dy dx.$$

Sabendo que,

$$\frac{d}{dt} [\Phi_{\mu, \sigma^2}(2xt\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)]^{n-1} = 2x\sigma(n-1)\phi_{\mu, \sigma^2}(2xt\sigma - y\sigma) \\ [\Phi_{\mu, \sigma^2}(2xt\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)]^{n-2},$$

então,

$$F_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) = \sigma \int_0^{\infty} \int_{-\infty}^{x\bar{q}} n\phi_{\mu, \sigma^2}(y\sigma) [\{\Phi_{\mu, \sigma^2}(2xt\sigma - y\sigma) - \\ - \Phi_{\mu, \sigma^2}(y\sigma)\}^{n-1}]_{t=y/x}^{t=\bar{q}} \times f(x; \nu) dy dx.$$

Resolvendo,

$$[\{\Phi_{\mu, \sigma^2}(2xt\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)\}^{n-1}]_{t=y/x}^{t=\bar{q}} = \{\Phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)\}^{n-1} - \\ - \{\Phi_{\mu, \sigma^2}(2x \cdot \frac{y}{x}\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)\}^{n-1}, \\ = \{\Phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)\}^{n-1},$$

a função de distribuição de \bar{Q} é dada por

$$F_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) = \sigma \int_0^{\infty} \int_{-\infty}^{x\bar{q}} n\phi_{\mu, \sigma^2}(y\sigma) [\Phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)]^{n-1} \times$$

$$\times f_X(x; \nu) dy dx$$

Então, a primeira parte fica concluída. Para a prova (ii), basta usar as equivalências das expressões (29) e (30) do Teorema 2.3, e aplicá-las nas expressões (45) e (46). Assim, para determinar a função densidade do item (ii), tem-se que

$$\begin{aligned} f_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) &= \sigma^2 \int_0^\infty \int_{-\infty}^{x\bar{q}} 2n(n-1)x\phi_{\mu, \sigma^2}(y\sigma)\phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) \times \\ &\quad \times [\Phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)]^{n-2} f_X(x; \nu) dy dx \\ &= \sigma^2 \int_0^\infty \int_{-\infty}^{x\bar{q}} 2n(n-1)x \frac{\phi_{\frac{\mu}{\sigma}, 1}(y)}{\sigma} \frac{\phi_{\frac{\mu}{\sigma}, 1}(2x\bar{q} - y)}{\sigma} \times \\ &\quad \times [\Phi_{\frac{\mu}{\sigma}, 1}(2x\bar{q} - y) - \Phi_{\frac{\mu}{\sigma}, 1}(y)]^{n-2} f_X(x; \nu) dy dx \\ &= \int_0^\infty \int_{-\infty}^{x\bar{q}} 2n(n-1)x\phi_{\frac{\mu}{\sigma}, 1}(y)\phi_{\frac{\mu}{\sigma}, 1}(2x\bar{q} - y) \times \\ &\quad \times [\Phi_{\frac{\mu}{\sigma}, 1}(2x\bar{q} - y) - \Phi_{\frac{\mu}{\sigma}, 1}(y)]^{n-2} f_X(x; \nu) dy dx \end{aligned}$$

Para determinar a função de distribuição de (ii), tem-se que

$$\begin{aligned} F_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) &= \sigma \int_0^\infty \int_{-\infty}^{x\bar{q}} n\phi_{\mu, \sigma^2}(y\sigma) [\Phi_{\mu, \sigma^2}(2x\bar{q}\sigma - y\sigma) - \Phi_{\mu, \sigma^2}(y\sigma)]^{n-1} \times \\ &\quad \times f_X(x; \nu) dy dx \\ &= \sigma \int_0^\infty \int_{-\infty}^{x\bar{q}} n \frac{\phi_{\frac{\mu}{\sigma}, 1}(y)}{\sigma} [\Phi_{\frac{\mu}{\sigma}, 1}(2x\bar{q} - y) - \Phi_{\frac{\mu}{\sigma}, 1}(y)]^{n-1} \times \\ &\quad \times f_X(x; \nu) dy dx \\ &= \int_0^\infty \int_{-\infty}^{x\bar{q}} n\phi_{\frac{\mu}{\sigma}, 1}(y) [\Phi_{\frac{\mu}{\sigma}, 1}(2x\bar{q} - y) - \Phi_{\frac{\mu}{\sigma}, 1}(y)]^{n-1} f_X(x; \nu) dy dx. \end{aligned}$$

Por fim, para a prova do item (iii) serão aplicadas as equivalências das expressões (31) e (32) nas funções densidade e de distribuição expressas em (47) e (48), respectivamente. Assim, para a função densidade tem-se que

$$\begin{aligned} f_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) &= \int_0^\infty \int_{-\infty}^{x\bar{q}} 2n(n-1)x\phi_{\frac{\mu}{\sigma}, 1}(y)\phi_{\frac{\mu}{\sigma}, 1}(2x\bar{q} - y) \times \\ &\quad \times [\Phi_{\frac{\mu}{\sigma}, 1}(2x\bar{q} - y) - \Phi_{\frac{\mu}{\sigma}, 1}(y)]^{n-2} f(x; \nu) dy dx \end{aligned}$$

$$= \int_0^\infty \int_{-\infty}^{x\bar{q}} 2n(n-1)x\phi(y-\mu/\sigma)\phi(2x\bar{q}-y-\mu/\sigma) \times \\ \times [\Phi(2x\bar{q}-y-\mu/\sigma) - \Phi(y-\mu/\sigma)]^{n-2} f(x;\nu) dy dx.$$

Para a função de distribuição, tem-se que

$$F_{\bar{Q}}(\bar{q}; \mu, \sigma^2, n, \nu) = \int_0^\infty \int_{-\infty}^{x\bar{q}} n\phi_{\frac{\mu}{\sigma}, 1}(y) [\Phi_{\frac{\mu}{\sigma}, 1}(2x\bar{q}-y) - \Phi_{\frac{\mu}{\sigma}, 1}(y)]^{n-1} f(x;\nu) dy dx \\ = \int_0^\infty \int_{-\infty}^{x\bar{q}} n\phi(y-\mu/\sigma) [\Phi(2x\bar{q}-y-\mu/\sigma) - \Phi(y-\mu/\sigma)]^{n-1} \times \\ f(x;\nu) dy dx,$$

concluindo-se, assim, a prova do Teorema. ■

Observe que para uma população inicial com distribuição normal com média $\mu = 0$, independente do valor de σ^2 , as três funções densidade e as três funções de distribuição resultam, respectivamente, em

$$f_{\bar{Q}}(\bar{q}; n, \nu) = \int_0^\infty \int_{-\infty}^{x\bar{q}} 2n(n-1)x\phi(y)\phi(2x\bar{q}-y) \times \\ \times [\Phi(2x\bar{q}-y) - \Phi(y)]^{n-2} f(x;\nu) dy dx. \quad (53)$$

e

$$F_{\bar{Q}}(\bar{q}; n, \nu) = \int_0^\infty \int_{-\infty}^{x\bar{q}} n\phi(y) [\Phi(2x\bar{q}-y) - \Phi(y)]^{n-1} \times \\ f(x;\nu) dy dx. \quad (54)$$

Assim, a distribuição de \bar{Q} não dependerá dos parâmetros da distribuição da população normal.

Para suporte na proposta dos PCMs, é importante entender onde a distribuição da *midrange* está centrada, isto é, sua esperança. Para isso, segue a seguinte Definição,

Definição 2.15 — Distribuição simétrica. Seja Y uma variável aleatória. A distri-

buição de Y é simétrica, centrada em μ , se

$$F_Y(y + \mu) = 1 - F_Y(\mu - y) \quad (55)$$

e

$$f_Y(y + \mu) = f_Y(\mu - y), \quad (56)$$

em que $F_Y(y)$ e $f_Y(y)$ são a função de distribuição e função densidade de Y , respectivamente. ■

Para mostrar que a esperança de Y com distribuição simétrica centrada em μ é $E[Y] = \mu$, será apresentado no seguinte Teorema,

Teorema 2.6 — Esperança de uma distribuição simétrica. Seja Y uma variável aleatória com distribuição simétrica centrada em μ , sendo $\mu \in \mathbb{R}$, então $E[Y] = \mu$. ■

Demonstração. Seja Y uma variável aleatória com função densidade $f_Y(y)$, sendo esta simétrica em torno de μ , isto é, $f_Y(y + \mu) = f_Y(\mu - y)$, Definição 2.15, então

$$\begin{aligned} E[Y] &= \int_{-\infty}^{+\infty} y f_Y(y) dy \\ &= \int_{-\infty}^{+\infty} (y - \mu + \mu) f_Y(y) dy \\ &= \int_{-\infty}^{+\infty} [\mu + (y - \mu)] f_Y(y) dy \\ &= \int_{-\infty}^{+\infty} \mu f_Y(y) dy + \int_{-\infty}^{+\infty} (y - \mu) f_Y(y) dy \\ &= \mu \int_{-\infty}^{+\infty} f_Y(y) dy + \int_{-\infty}^{+\infty} (y - \mu) f_Y(y) dy \\ &= \mu + \int_{-\infty}^{+\infty} (y - \mu) f_Y(y) dy. \end{aligned} \quad (57)$$

Basta agora mostrar que $\int_{-\infty}^{+\infty} (y - \mu) f_Y(y) dy = 0$. Fazendo uma mudança de variável, $t = y - \mu$, tem-se

$$\int_{-\infty}^{+\infty} (y - \mu) f_Y(y) dy = \int_{-\infty}^{+\infty} t f_Y(t + \mu) dt. \quad (58)$$

Fazendo uma partição nos limites de integração da expressão (58), então

$$\int_{-\infty}^{+\infty} t f_Y(t + \mu) dt = \int_{-\infty}^0 t f_Y(t + \mu) dy + \int_0^{\infty} t f_Y(t + \mu) dt. \quad (59)$$

Para a primeira integral da expressão (59), fazendo uma mudança de variável do tipo $z = -t$, tem-se

$$\int_{-\infty}^0 (-1)(-z) f_Y(\mu - z) dz.$$

Usando o fato que $f_Y(\mu - z) = f_Y(z + \mu)$, e fazendo a inversão dos limites de integração, segue que

$$\int_{-\infty}^0 (-1)(-z) f_Y(\mu - z) dz = - \int_0^{\infty} z f_Y(\mu + z) dz.$$

Fazendo $z = t$, tem-se que

$$- \int_0^{\infty} z f_Y(\mu + z) dz = - \int_0^{\infty} t f_Y(\mu + t) dt. \quad (60)$$

Usando os resultados (59) e (60), a esperança de Y é

$$\begin{aligned} E[Y] &= \mu + \int_{-\infty}^{+\infty} (y - \mu) f_Y(y) dy \\ &= \mu + \left(\int_{-\infty}^0 t f_Y(t + \mu) dy + \int_0^{\infty} t f_Y(t + \mu) dt \right) \\ &= \mu + \left(- \int_0^{\infty} t f_Y(\mu + t) dz + \int_0^{\infty} t f_Y(t + \mu) dt \right) \\ &= \mu + 0 \\ &= \mu, \end{aligned}$$

concluindo-se a demonstração. ■

Um importante Teorema provado por Gumbel, Carlson e Mustafi (1965, p. 1052), mostrou que se a distribuição inicial² (ou distribuição parental) é simétrica em torno de

²O termo distribuição inicial ou distribuição parental refere-se à distribuição da população da qual os dados foram amostrados. Por exemplo, a distribuição da *midrange* depende da função densidade e função

μ , a distribuição da *midrange* também é simétrica em torno de μ , o qual será apresentado a seguir, sem prova.

Teorema 2.7 Se a distribuição inicial é ilimitada em seu suporte, diferenciável, simétrica e unimodal, então a distribuição da *midrange* é também ilimitada, diferenciável, simétrica e unimodal. ■

Demonstração. Ver Gumbel, Carlson e Mustafi (1965, p. 1052-1054). ■

Com base nessas informações, pode-se apresentar a esperança da *midrange*.

Teorema 2.8 — Esperança de \bar{R} . Seja uma amostra aleatória Y_1, Y_2, \dots, Y_n de uma população com distribuição simétrica centrada em μ , e *midrange* (Definição 2.12), denotada por \bar{R} , então a esperança de \bar{R} é dada por

$$E[\bar{R}] = \mu. \quad (61)$$

Demonstração. Da Definição 2.12, sabe-se que

$$\bar{R} = \frac{Y_{(n)} + Y_{(1)}}{2}$$

e que as funções densidade do mínimo ($Y_{(1)}$) e do máximo ($Y_{(n)}$) são dadas por

$$f_{Y_{(1)}}(y) = n f_Y(y) [1 - F_Y(y)]^{n-1} \quad (62)$$

e

$$f_{Y_{(n)}}(y) = n f_Y(y) [F_Y(y)]^{n-1}, \quad (63)$$

respectivamente, como mostrado em Batista (2012, p. 47). Observe que $F_Y(y)$ e $f_Y(y)$ são a função distribuição e função densidade, respectivamente, da população parental. A esperança de \bar{R} pode ser expressa por

$$E[\bar{R}] = E[(Y_{(1)} + Y_{(n)})/2]$$

distribuição da população. Então a *midrange* depende da distribuição inicial dos dados.

$$= 1/2(E[Y_{(1)}] + E[Y_{(n)}]). \quad (64)$$

A esperança do mínimo é dada por

$$E[Y_{(1)}] = n \int_{-\infty}^{\infty} y f_Y(y) [1 - F_Y(y)]^{n-1} dy.$$

Fazendo uma mudança de variável $z = y - \mu$, então

$$\begin{aligned} E[Y_{(1)}] &= n \int_{-\infty}^{\infty} (z + \mu) f_Y(z + \mu) [1 - F_Y(z + \mu)]^{n-1} dz \\ &= n \int_{-\infty}^{\infty} z f_Y(z + \mu) [1 - F_Y(z + \mu)]^{n-1} dz + \\ &\quad + \mu \int_{-\infty}^{\infty} n f_Y(z + \mu) [1 - F_Y(z + \mu)]^{n-1} dz \\ &= n \int_{-\infty}^{\infty} z f_Y(z + \mu) [1 - F_Y(z + \mu)]^{n-1} dz + \\ &\quad + (-1)(-1)\mu \int_{-\infty}^{\infty} n f_Y(z + \mu) [1 - F_Y(z + \mu)]^{n-1} dz. \end{aligned}$$

Para a segunda integral, tem-se que

$$\frac{d}{dz} \{ [1 - F_Y(z + \mu)]^n \} = -n f_Y(z + \mu) [1 - F_Y(z + \mu)]^{n-2}.$$

Assim,

$$\begin{aligned} E[Y_{(1)}] &= n \int_{-\infty}^{\infty} z f_Y(z + \mu) [1 - F_Y(z + \mu)]^{n-1} dz - \\ &\quad - \mu \{ [1 - F_Y(z + \mu)]^n \}_{z=-\infty}^{z=\infty} \\ &= n \int_{-\infty}^{\infty} z f_Y(z + \mu) [1 - F_Y(z + \mu)]^{n-1} dz - \\ &\quad - \mu \left\{ \lim_{z \rightarrow +\infty} [1 - F_Y(z + \mu)]^{n-1} - \lim_{z \rightarrow -\infty} [1 - F_Y(z + \mu)]^{n-1} \right\} \\ &= n \int_{-\infty}^{\infty} z f_Y(z + \mu) [1 - F_Y(z + \mu)]^{n-1} dz - \mu(0 - 1) \\ &= n \int_{-\infty}^{\infty} z f_Y(z + \mu) [1 - F_Y(z + \mu)]^{n-1} dz + \mu. \end{aligned} \quad (65)$$

A esperança do máximo é dada por

$$E[Y_{(n)}] = n \int_{-\infty}^{\infty} y f_Y(y) [F_Y(y)]^{n-1} dy.$$

Fazendo uma mudança de variável $z = y - \mu$, então

$$\begin{aligned} E[Y_{(n)}] &= n \int_{-\infty}^{\infty} (z + \mu) f_Y(z + \mu) [F_Y(z + \mu)]^{n-1} dz \\ &= n \int_{-\infty}^{\infty} z f_Y(z + \mu) [F_Y(z + \mu)]^{n-1} dz + \\ &\quad + n \int_{-\infty}^{\infty} \mu f_Y(z + \mu) [F_Y(z + \mu)]^{n-1} dz \\ &= n \int_{-\infty}^{\infty} z f_Y(z + \mu) [F_Y(z + \mu)]^{n-1} dz + \mu \{ [F_Y(z + \mu)]^n \}_{-\infty}^{+\infty} \\ &= n \int_{-\infty}^{\infty} z f_Y(z + \mu) [F_Y(z + \mu)]^{n-1} dz + \\ &\quad + \mu \left\{ \lim_{z \rightarrow +\infty} [F_Y(z + \mu)]^n - \lim_{z \rightarrow -\infty} [F_Y(z + \mu)]^n \right\} \\ &= n \int_{-\infty}^{\infty} z f_Y(z + \mu) [F_Y(z + \mu)]^{n-1} dz + \mu(1 - 0) \\ &= n \int_{-\infty}^{\infty} z f_Y(z + \mu) [F_Y(z + \mu)]^{n-1} dz + \mu. \end{aligned} \tag{66}$$

Agora, usando os resultados de (65) e (66) em (64), encontra-se a esperança da *midrange* da seguinte forma

$$E[\bar{R}] = \frac{1}{2} \left\{ \underbrace{n \int_{-\infty}^{\infty} z f_Y(z + \mu) [1 - F_Y(z + \mu)]^{n-1} dz}_A + \underbrace{n \int_{-\infty}^{\infty} z f_Y(z + \mu) [F_Y(z + \mu)]^{n-1} dz + 2\mu}_B \right\}. \tag{67}$$

Pela simetria da distribuição de Y , a área A pode ser reescrita como

$$A = n \int_{-\infty}^{\infty} z f_Y(-z + \mu) [F_Y(-z + \mu)]^{n-1} dz.$$

Fazendo $y = -z$, tem-se

$$\begin{aligned} A &= n \int_{+\infty}^{-\infty} (-y) f_Y(y + \mu) [F_Y(y + \mu)]^{n-1} (-1) dy \\ &= n \int_{+\infty}^{-\infty} y f_Y(y + \mu) [F_Y(y + \mu)]^{n-1} dy. \end{aligned}$$

Invertendo os limites de integração, logo

$$\begin{aligned} A &= -n \int_{-\infty}^{\infty} y f_Y(y + \mu) [F_Y(y + \mu)]^{n-1} dy \\ &= -B. \end{aligned} \tag{68}$$

Substituindo o resultado (68) em (67), logo

$$\begin{aligned} E[\bar{R}] &= \frac{1}{2} \{-B + B + 2\mu\} \\ &= \mu, \end{aligned}$$

concluindo-se a demonstração. ■

Para o caso de uma população com distribuição normal com média $\mu = 0$, isto é, $Y_i \sim N(0, \sigma^2)$, para $i = 1, 2, \dots, n$, sabe-se que a distribuição de \bar{R} é também simétrica centrada na esperança de Y_i , isto é, $\mu = 0$. Assim, $E[\bar{R}] = 0$.

A esperança de \bar{W} de uma amostra aleatória Y_1, Y_2, \dots, Y_n de uma população com distribuição simétrica centrada em μ é apresentada no Teorema a seguir.

Teorema 2.9 — Esperança da *midrange* padronizada. Seja uma amostra aleatória Y_1, Y_2, \dots, Y_n de uma população com distribuição simétrica centrada em μ , e *midrange* padronizada, Definição 2.13, denotada por \bar{W} , então a esperança de \bar{W} é dada por

$$E[\bar{W}] = \frac{\mu}{\sigma}. \tag{69}$$

Demonstração. A esperança de \bar{W} pode expressa da seguinte forma:

$$E[\bar{W}] = E \left[\frac{\bar{R}}{\sigma} \right]$$

$$= \frac{1}{\sigma} E[\bar{R}].$$

Pelo Teorema 2.8, sabe-se que para qualquer distribuição simétrica centrada em μ a esperança $E[\bar{R}] = \mu$. Portanto,

$$E[\bar{W}] = \frac{\mu}{\sigma}.$$

■

Também será importante saber a esperança de \bar{Q} para uma amostra aleatória de uma população com distribuição normal, o que será apresentada a seguir.

Teorema 2.10 — Esperança da \bar{Q} . Seja uma amostra X_1, X_2, \dots, X_n de uma população normal com média μ e variância σ^2 , isto é $X_i \sim N(\mu, \sigma^2)$, para $i=1, 2, \dots, n$. Então a esperança da *midrange* estudentizada externamente, \bar{Q} , da Definição 2.14, é dada por

$$E[\bar{Q}] = \frac{\mu \left(\frac{\nu}{2}\right)^{1/2} \Gamma\left(\frac{\nu-1}{2}\right)}{\sigma \Gamma(\nu/2)}. \quad (70)$$

■

Demonstração. Pela expressão (44), \bar{Q} pode ser escrita como:

$$\bar{Q} = \frac{\bar{R}/\sigma}{S/\sigma} = \frac{\bar{W}}{X}.$$

Sabe-se que \bar{R} e S são estimados independentemente. Assim, a esperança de \bar{Q} é:

$$E[\bar{Q}] = E\left[\frac{\bar{R}/\sigma}{X}\right].$$

Pela propriedade da independência entre duas variáveis (MOOD; GRAYBILL; BOES, 1974, p. 160), a esperança de \bar{Q} pode ser expressa por

$$E[\bar{Q}] = E[\bar{R}/\sigma] \times E\left[\frac{1}{X}\right].$$

A esperança $E[\bar{R}/\sigma]$ foi obtida no Teorema 2.9. Basta apenas fazer a esperança $E[1/X]$,

que será apresentada a seguir.

$$E[1/X] = \int_0^{\infty} \frac{1}{x} f_X(x; \nu) dx,$$

sendo que a função densidade de X , $f_X(x; \nu)$, foi obtida pelo Teorema 2.1. Assim,

$$\begin{aligned} E[1/X] &= \int_0^{\infty} \frac{1}{x} \frac{\nu^{\nu/2}}{\Gamma(\nu/2) 2^{\nu/2-1}} x^{\nu-1} e^{-\nu x^2/2} dx \\ &= \int_0^{\infty} \frac{1}{x} \frac{\nu^{\nu/2}}{\Gamma(\nu/2) 2^{\nu/2-1}} x^{\nu-1} e^{-\nu x^2/2} \frac{x\nu}{x\nu} dx \\ &= \int_0^{\infty} \frac{\nu^{\nu/2-1}}{\Gamma(\nu/2) 2^{\nu/2}} x^{\nu-3} e^{-\nu x^2/2} 2\nu x dx. \end{aligned}$$

Fazendo a transformação $U = \nu X^2 \Rightarrow du = 2\nu x \Rightarrow x = (u/\nu)^{1/2}$, para $x > 0$. Logo,

$$\begin{aligned} E[1/X] &= \int_0^{\infty} \frac{\nu^{\nu/2-1}}{\Gamma(\nu/2) 2^{\nu/2}} \left(\left(\frac{u}{\nu} \right)^{\frac{1}{2}} \right)^{\nu-3} e^{-u/2} du \\ &= \frac{1}{\Gamma(\nu/2)} \int_0^{\infty} \frac{1}{2^{\nu/2}} \frac{\nu^{\nu/2-1}}{\nu^{\nu/2-3/2}} u^{\nu/2-3/2} e^{-u/2} du \\ &= \frac{1}{\Gamma(\nu/2)} \int_0^{\infty} \frac{1}{2^{\nu/2}} \frac{\nu^{\nu/2-1}}{\nu^{\nu/2-3/2}} u^{(\nu-1)/2-1} e^{-u/2} du \\ &= \frac{\nu^{1/2}}{\Gamma(\nu/2) 2^{1/2}} \int_0^{\infty} \frac{1}{2^{(\nu-1)/2}} u^{(\nu-1)/2-1} e^{-u/2} du \\ &= \frac{(\nu/2)^{1/2}}{\Gamma(\nu/2)} \int_0^{\infty} \frac{1}{2^{(\nu-1)/2}} u^{(\nu-1)/2-1} e^{-u/2} du \\ &= \frac{(\frac{\nu}{2})^{1/2} \Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})} \underbrace{\int_0^{\infty} \frac{1}{\Gamma(\frac{\nu-1}{2}) 2^{\frac{\nu-1}{2}}} u^{\frac{\nu-1}{2}-1} e^{-u/2} du}_{=1} \\ &= \frac{(\frac{\nu}{2})^{1/2} \Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})}. \end{aligned}$$

Portanto, a esperança de \bar{Q} é

$$E[\bar{Q}] = \frac{\mu}{\sigma} \frac{(\frac{\nu}{2})^{1/2} \Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)}.$$

■

Caso a distribuição normal esteja centrada em $\mu = 0$, $E[\bar{Q}] = 0$ também, isto é, a distribuição de \bar{Q} está centrada em 0 para essa situação.

Com base nessas definições e Teoremas apresentados, o próximo passo é apresentar a metodologia para a propostas dos testes de Tukey Midrange, de SNK Midrange, e os testes de agrupamento de médias baseados na *midrange* estudentizada externamente e na amplitude estudentizada externamente.

3 MÉTODOS

Quatro métodos de comparações múltiplas foram propostos nesse trabalho, três deles baseados na distribuição da *midrange* estudentizada externamente e outro baseado na amplitude estudentizada. A descrição dos novos testes será feita nas seções subsequentes. Uma segunda parte do trabalho constitui da validação dos testes por simulação Monte Carlo. Nesse caso, duas situações foram consideradas. Na primeira, avaliou-se o erro tipo I por experimento dos testes. Na segunda situação, foram avaliados os desempenhos dos testes em relação ao poder.

Para o estudo do poder, considerou-se duas etapas. Na primeira, um modelo sob H_1 completa e na segunda, um modelo sob H_0 parcial. Sob H_0 parcial, além do poder, avaliou-se o erro tipo I por experimento, utilizando-se os contrastes intra grupos, cujas médias de tratamentos, por construção, foram iguais. Em ambas as etapas, os testes de Tukey e SNK baseados na amplitude estudentizada externamente também foram simulados para fins de comparação. Resultados de outros testes encontrados na literatura também serviram de comparação para os testes propostos.

Finalmente, apresentou-se um exemplo real, no qual ilustrou-se a aplicação dos testes propostos. Os testes originais de Tukey e de SNK foram também aplicados a esse conjunto de dados. Abaixo, na Figura 2, segue um fluxograma que auxilia a entender os passos para obtenção e validação dos métodos.

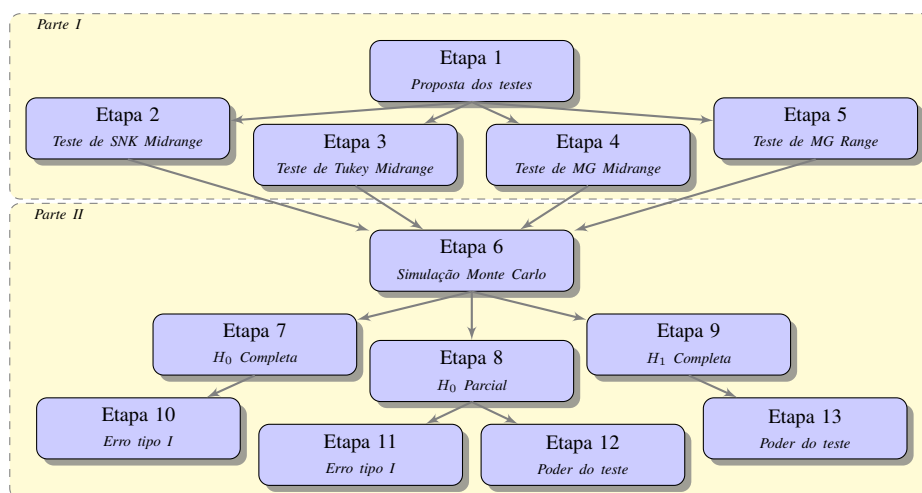


Figura 2 Etapas do desenvolvimento e avaliação dos testes propostos.

3.1 Filosofia dos testes propostos

Para a proposição dos testes a seguinte amostra aleatória foi considerada: $Y_{11}, Y_{12}, \dots, Y_{1r}, Y_{21}, \dots, Y_{2r}, \dots, Y_{i1}, Y_{i2}, \dots, Y_{ij}, \dots, Y_{ir}, \dots, Y_{n1}, Y_{n2}, \dots, Y_{nr}$, sendo Y_{ij} a observação aleatória referente ao i -ésimo tratamento e a sua j -ésima repetição, $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, r$. A média do i -ésimo tratamento é:

$$\bar{Y}_i = \frac{\sum_{j=1}^r Y_{ij}}{r} = \frac{Y_i}{r}. \quad (71)$$

Essa amostra foi submetida a uma análise de variância, adotando-se o seguinte modelo:

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij} = \mu_i + \epsilon_{ij},$$

em que $\epsilon_{ij} \sim N(0, \sigma^2)$ e $\mu_i = \mu + \tau_i$ é a média do i -ésimo tratamento. Assim, o quadrado médio do erro (QME) foi estimado por:

$$QME = \frac{\sum_{i=1}^n \sum_{j=1}^r (Y_{ij} - \bar{Y}_i)^2}{n(r-1)}. \quad (72)$$

Verifica-se que \bar{Y}_i e QME são independentemente distribuídos e que $\hat{V}(\bar{Y}_i) = QME/r$, ver Graybill (1961, p. 206) e Searle (1987, p. 99).

Sob a hipótese nula $H_0 : \mu_1 = \mu_2 = \dots = \mu_n = \mu$, os n tratamentos possuem média comum μ . Nesse caso particular, as estatísticas de ordem $\bar{Y}_{(1)}, \bar{Y}_{(2)}, \dots, \bar{Y}_{(n)}$ são centradas em μ . Dessa forma, a *midrange* estudentizada externamente, definida por

$$\bar{Q} = \frac{\sqrt{r} [(\bar{Y}_{(1)} + \bar{Y}_{(n)})/2]}{\sqrt{QME}}, \quad (73)$$

possui função de distribuição dependente de μ (sob H_0), como apresentado no Teorema 2.5.

Entretanto, μ é desconhecido e dificilmente igual a zero nas situações reais. Dessa forma, utilizar a distribuição de \bar{Q} com μ desconhecido e $\mu \neq 0$ ficaria impossível na proposição dos testes.

Assim, optou-se por utilizar a distribuição da *midrange* no caso específico em que $\mu = 0$, expressão (54). Observa-se que a distribuição dos dados está centrada em μ , que é desconhecido. Logo, para usar a distribuição centrada em 0, houve uma correção na estatística. Como $\bar{R} = (\bar{Y}_{(1)} + \bar{Y}_{(n)})/2$ tem distribuição centrada em μ , a estatística corrigida foi $\bar{R}_n = \bar{R} - \bar{Y}_*^*$, em que \bar{Y}_*^* é um estimador de μ .

Inicialmente, pensou-se na média geral como um estimador de \bar{Y}_*^* , isto é, $\bar{Y}_* = \sum_{i=1}^n \bar{Y}_i./n$. Entretanto, quando os dados experimentais estavam sob H_1 , essa média estimava a média geral dos parâmetros $\mu = \sum_{i=1}^n \mu_i/n$, em que $\mu_i = E(\bar{Y}_i.)$. Assim, sob H_1 , observou-se por meio da simulação que a quantidade $(\mu_1 + \mu_n)/2$ aproximava de μ e portanto, $\bar{R}_n = (\mu_1 + \mu_n)/2 - \mu \approx 0$. Dessa forma o desempenho dos testes para o poder era muito baixo.

Sob H_0 , é claro que essa esperança é nula, o que a princípio respaldaria o uso da distribuição da *midrange* no teste. No entanto, o que se observou, numa avaliação preliminar via simulação, foi que os erros tipo I por experimento foram muito elevados. Inicialmente, especulou-se que isso decorreu do fato de a estatística ser função de $\bar{Y}_*.$, que possui erro amostral associado. Assim, a DMS (diferença mínima significativa) que Representasse o erro padrão de $\bar{Y}_*.$ foi construída. Entretanto, o teste passou a controlar adequadamente o erro tipo I, mas apresentou baixo poder.

Isso ocorreu porque $E(\bar{R}_n)$, embora diferente de zero, sob H_1 , apresentava valores em magnitude não tão diferentes de zero. Assim, buscou-se um estimador de μ que maximizasse $E(\bar{R}_n)$ sob H_1 e que $E(\bar{R}_n) = 0$ sob H_0 . Nesse caso, usou-se \bar{Y}_*^* que corresponderia a média de um dos dois potenciais grupos a serem obtidos no teste. Essa partição se daria entre duas médias ordenadas de máxima amplitude.

Para se obter um estimador de menor erro padrão, usou-se a média do grupo de maior número de médias envolvidas entre os dois grupos considerados. Esse estimador foi determinado baseado em critério empírico e de validação por simulação Monte Carlo. Assim, considerando as partições $\bar{Y}_{(1)}, \bar{Y}_{(2)}, \dots, \bar{Y}_{(k)}$ e $\bar{Y}_{(k+1)}, \bar{Y}_{(k+2)}, \dots, \bar{Y}_{(n)}$, cujo ponto k foi definido para o valor j , em que

$$\max_j (\bar{Y}_{(j+1)} - \bar{Y}_{(j)})$$

ocorre, para $j = 1, 2, \dots, n - 1$. Se houver empates com dois ou mais valores diferentes de k , digamos k_1, k_2, \dots , então forma-se a partição em que $k = \max\{\min(k_1, m - k_1), \min(k_2, m - k_2)\}$.

Dessa forma, tomando-se

$$\bar{Y}_1^* = \frac{\sum_{j=1}^k \bar{Y}_{(j)}}{k}$$

e

$$\bar{Y}_2^* = \frac{\sum_{j=k+1}^n \bar{Y}_{(j)}}{n-k},$$

o valor de \bar{Y}_1^* irá corresponder a \bar{Y}_1^* se $k \geq n - k$ ou igual a \bar{Y}_2^* , caso contrário.

Assim, a estatística ficou da seguinte forma:

$$\bar{R}_n = \frac{\bar{Y}_{(1)} + \bar{Y}_{(n)}}{2} - \bar{Y}_1^* \quad (74)$$

e inicialmente, a DMS para a rejeição ou não da hipótese, foi

$$\Delta_n = \bar{q}_{(\alpha/2; n, \nu)} \sqrt{\frac{QME}{r}} + \frac{1}{\sqrt{n}} \sqrt{\frac{QME}{r}}. \quad (75)$$

Em resultados de simulação Monte Carlo preliminares, observou-se que o erro tipo I por experimento estava muito menor que os níveis nominais de significância e que o poder era baixo. Assim, a contribuição de \bar{Y}_1^* para a DMS, $(1/\sqrt{n}) \times \sqrt{QME/r}$ deveria ser reduzida por um fator entre 0 e 1. Por tentativa e erro em um processo de simulação Monte Carlo, encontrou-se um valor que convergia para $\sqrt{2}/2$. Assim, a DMS final é

$$\Delta_n = \bar{q}_{(\alpha/2; n, \nu)} \sqrt{\frac{QME}{r}} + \frac{1}{\sqrt{2n}} \sqrt{\frac{QME}{r}}. \quad (76)$$

3.1.1 Teste de Tukey baseado na *midrange* (TM)

Considerando as justificativas apresentadas na subseção 3.1, os passos para a aplicação do teste são:

1. As médias dos tratamentos devem ser ordenadas: $\bar{Y}_{(1)}, \bar{Y}_{(2)}, \dots, \bar{Y}_{(n)}$;

2. A diferença mínima significativa é obtida por:

$$\Delta_n = \bar{q}_{(\alpha/2;n,\nu)} \sqrt{\frac{QME}{r}} + \underbrace{\frac{1}{\sqrt{2n}} \sqrt{\frac{QME}{r}}}_{\text{Representa a variação de } \bar{Y}^*};$$

3. Calcula-se o valor da estatística, determinando-se k e \bar{Y}^* como descrito anteriormente, por:

$$\bar{r}_n = \frac{\bar{Y}_{(1).} + \bar{Y}_{(n).}}{2} - \bar{Y}^*;$$

4. Se $|\bar{r}_n| \leq \Delta_n$, então as n médias serão consideradas não diferentes. Caso contrário, vai para o passo 5, considerando $m = n$ médias;
5. Faça $m = m - 1$;
6. Considerando os grupos de médias: $\bar{Y}_{(1).}, \bar{Y}_{(2).}, \dots, \bar{Y}_{(m).}; \bar{Y}_{(2).}, \bar{Y}_{(3).}, \dots, \bar{Y}_{(m+1).}; \dots; \bar{Y}_{(n-m+1).}, \bar{Y}_{(n-m+2).}, \dots, \bar{Y}_{(n).}$; o número de grupos é dado por $l = n - m + 1$;
7. Para cada um dos grupos com m médias, é obtido

$$\bar{r}_m = \frac{\min_j \{\bar{Y}_{(j).}\} + \max_j \{\bar{Y}_{(j).}\}}{2} - \bar{Y}^*,$$

em que \bar{Y}^* é obtida nos mesmos moldes, conforme descrito para o conjunto de todas as n médias, sujeito apenas ao fato que se tem $m < n$ médias, nesse caso.

8. Para cada grupo obtido e marcado como divisível, considere m o número de médias do grupo relacionado, devendo-se usar como diferença mínima significativa dada por:

$$\Delta_n^* = \bar{q}_{(\alpha/2;n,\nu)} \sqrt{\frac{QME}{r}}; \quad (77)$$

9. Se $|\bar{r}_m| \leq \Delta_n^*$, então as m médias não diferem estatisticamente; caso contrário elas são diferentes nesse estágio;
10. Repete-se o processo para todos os l grupos de m médias, refazendo-se os passos 7, 8 e 9. Terminado a comparação de todos os l grupos de m médias, retorna-se ao passo 5, atualizando-se o valor de m . Isso deve ser repetido enquanto $m \geq 2$.

3.1.2 Teste de SNK baseado na *midrange* (SNKM)

O algoritmo para o teste de SNKM é o mesmo utilizado para o teste de TM. O que diferencia é o passo 8, que segue:

8 O valor da diferença mínima significativa é dada por

$$\Delta_m = \bar{q}_{(\alpha/2; m, \nu)} \sqrt{\frac{QME}{r}}. \quad (78)$$

Assim, Δ_m é alterado à medida que o número de médias m se modifica nos $l = n - m + 1$ grupos, modificando-se o argumento n por m , que é um dos parâmetros da distribuição de \bar{Q} .

3.1.3 Teste de agrupamento de médias baseado na *midrange* (MGM)

A sigla MGM, vem do inglês: “*Mean Grouping based on the Midrange*”. Da mesma forma como foi proposto para o teste de TM (Subseção 3.1.1), usou-se o critério de formar uma partição de m médias ordenadas na posição k , em que ocorre

$$\max_j \{\bar{Y}_{(j+1)} - \bar{Y}_{(j)}\} = \bar{Y}_{(k+1)} - \bar{Y}_k.$$

para $j = 1, 2, \dots, m - 1$. Se houver empates com dois ou mais valores diferentes de k , digamos k_1, k_2, \dots , então forma-se a partição em que $k = \max\{\min(k_1, m - k_1), \min(k_2, m - k_2)\}$. Assim, definiu-se

$$\bar{Y}_1^* = \frac{\sum_{j=1}^k \bar{Y}_{(j)}}{k}$$

e

$$\bar{Y}_2^* = \frac{\sum_{j=k+1}^m \bar{Y}_{(j)}}{m - k}.$$

Logo $\bar{Y}_m^* = \bar{Y}_1^*$ se $k \geq m - k$ ou $\bar{Y}_m^* = \bar{Y}_2^*$, caso contrário. Os passos para a aplicação do teste são:

1. Faça $m = n$ e tome as médias ordenadas dos tratamentos por: $\bar{Y}_{(1)}, \bar{Y}_{(2)}, \dots, \bar{Y}_{(m)}$;
2. Determina-se k e \bar{Y}_m^* como discutido anteriormente;
3. Determina-se o valor da estatística por:

$$\bar{r}_m = \frac{\bar{Y}_{(1)} + \bar{Y}_{(m)}}{2} - \bar{Y}_m^*;$$

4. A diferença mínima significativa é

$$\Delta_m = \bar{q}_{(\alpha/2; m, \nu)} \sqrt{\frac{QME}{r}} + \underbrace{\frac{1}{\sqrt{2m}} \sqrt{\frac{QME}{r}}}_{\text{Representa a variação de } \bar{Y}_m^*};$$

5. Se $|\bar{r}_m| \leq \Delta_m$, então as m médias são consideradas não diferentes e marca-se o grupo como não divisível e o processo se encerra. Caso contrário, considere que as médias do grupo $\bar{Y}_{(1)}, \bar{Y}_{(2)}, \dots, \bar{Y}_{(k)}$ como diferentes das médias do grupo $\bar{Y}_{(k+1)}, \bar{Y}_{(k+2)}, \dots, \bar{Y}_{(m)}$. e vá para o passo 6;
6. Para cada grupo obtido e marcado como divisível, considere m o número de médias do grupo relacionado. Repita os passos de 2 a 5, com uma ressalva, no passo 4, deve-se usar como diferença mínima significativa:

$$\Delta_m = \bar{q}_{(\alpha/2; m, \nu)} \sqrt{\frac{QME}{r}}.$$

Esse processo é realizado para todos os grupos até que nenhum outro grupo consiga ser dividido em dois novos ou até que todos os grupos contenham apenas 1 média.

3.1.4 Teste de agrupamento de médias baseado na amplitude estudentizada (MGR)

A sigla MGR, vem do inglês: “*Mean Grouping based on the Range*”. Uma versão similar do teste de Scott-Knott (SCOTT; KNOTT, 1974), baseado na amplitude estudentizada também foi proposta. A essência do teste é a mesma da proposta apresentada para o teste de MGM (Seção 3.1.3). Para divisão dos grupos, usou-se como potencial ponto de partição a posição da máxima amplitude entre médias ordenadas. Assim, para

$\bar{Y}_{(1).}, \bar{Y}_{(2).}, \dots, \bar{Y}_{(m).}$, a partição deve ser considerada na posição k onde se verifica:

$$\max_j \{\bar{Y}_{(j+1).} - \bar{Y}_{(j).}\} = \bar{Y}_{(k+1).} - \bar{Y}_{k.},$$

para $j = 1, 2, \dots, m - 1$. Deve-se considerar para aplicação do teste o quantil superior $100\alpha\%$, $q_{\alpha; m, \nu}$, da amplitude estudentizada externamente.

Os passos para a aplicação do teste são:

1. Fazer $m = n$ e considerar m médias ordenadas: $\bar{Y}_{(1).}, \bar{Y}_{(2).}, \dots, \bar{Y}_{(m).}$;
2. Determinar k conforme discutido anteriormente;
3. Calcular a estatística do teste por:

$$q_m = \bar{Y}_{(m).} - \bar{Y}_{(1).}; \quad (79)$$

4. A diferença mínima significativa é

$$\Delta_m = q_{(\alpha; m, \nu)} \sqrt{\frac{QME}{r}}; \quad (80)$$

5. Se $q_m \leq \Delta_m$, então as m médias são consideradas não diferentes, marque o grupo como não divisível e vá para o passo 6. Caso contrário, considere as médias do grupo $\bar{Y}_{(1).}, \bar{Y}_{(2).}, \dots, \bar{Y}_{(k).}$ como diferentes das médias do grupo $\bar{Y}_{(k+1).}, \dots, \bar{Y}_{(m).}$ e vá para o passo 6.
6. Para cada grupo obtido e marcado como divisível, considere m o número de médias do grupo relacionado. Repita os passos de 2 a 5. Esse processo é realizado para todos os grupos até que nenhum outro grupo consiga ser dividido em dois novos ou até que todos os grupos contenham apenas 1 média.

3.2 Avaliação do desempenho

Dois estratégias foram consideradas neste trabalho. A primeira foi de avaliar as taxas de erro tipo I por experimento dos testes de comparações múltiplas propostos. A segunda foi de avaliar o poder dos testes. Em ambos os casos, foi usada simulação Monte Carlo. Em cada simulação foram aplicados os testes de comparações múltiplas em um nível nominal pré-estabelecido de significância α , verificando se a hipótese nula foi ou não rejeitada. Este processo, em cada caso, foi repetido $N^* = 5000$ vezes e a proporção de experimentos com pelo menos uma decisão incorreta, no primeiro caso, refere-se a

taxa de erro tipo I por experimento empírico e, no segundo caso, a proporção de decisões (rejeições) corretas refere-se ao poder empírico.

Para avaliar a taxa de erro tipo I por experimento empírico simulada via Monte Carlo, usou-se o teste binomial exato com coeficiente de confiança de 99% de probabilidade para testar as hipóteses $H_0 : \alpha = 5\%$ contra $H_1 : \alpha \neq 5\%$ e $H_0 : \alpha = 1\%$ contra $H_1 : \alpha \neq 1\%$. Se a hipótese nula for rejeitada e a taxa de erro tipo I por experimento empírica for considerada significativamente (valor-p < 0,01) inferior ao nível nominal, o teste será considerado conservativo. Se a taxa de erro tipo I por experimento empírica for considerada significativamente (valor-p < 0,01) superior ao nível nominal, o teste será considerado liberal. Caso o valor observado da taxa de erro tipo I por experimento empírica seja não significativa (valor-p > 0,01), o teste será considerado exato (OLIVEIRA; FERREIRA, 2010).

Considerando y o número de hipóteses nula rejeitadas em $N^* = 5000$ simulações Monte Carlo, para um nível nominal de significância α , a estatística do teste usando a relação entre a distribuição F e a distribuição binomial (LEEMIS; TRIVEDI, 1996), com probabilidade de sucesso $p = \alpha$, é dada por

$$F = \left(\frac{y + 1}{N^* - y} \right) \left(\frac{1 - \alpha}{\alpha} \right), \quad (81)$$

sob H_0 . Essa estatística tem distribuição F com $\nu_1 = 2(N^* - y)$ e $\nu_2 = 2(y + 1)$ graus de liberdade. Se $F < F_{0,005}$ ou $F > F_{0,995}$, a hipótese nula deve ser rejeitada ao nível de significância de 1% de probabilidade, em que $F_{0,005}$ e $F_{0,995}$ são os quantis da distribuição F com ν_1 e ν_2 graus de liberdade.

Em ambas as etapas foram simulados dados de acordo com o seguinte modelo estatístico:

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij}, \quad (82)$$

em que μ é constante geral fixada em 100 em todos os casos, sem perda de generalidade, τ_i é o efeito do i -ésimo tratamento e ϵ_{ij} é o efeito do erro aleatório com distribuição normal e independentemente distribuído com média 0 e variância σ^2 comum, também fixada em 100, sem perda de generalidade, sendo ainda $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, r$, em que r é o número de repetições.

Na primeira etapa de avaliação do erro tipo I por experimento, os efeitos do tratamento τ_i foram considerados iguais a 0 para todo i , $i = 1, 2, \dots, n$. Assim, os

dados foram gerados sob hipótese nula completa, ou seja, com todos os tratamentos tendo as mesmas médias paramétricas. A probabilidade de erro tipo I por experimento ($\hat{\alpha}$) foi estimada pela proporção de experimentos com pelo menos uma diferença detectada incorretamente em relação ao total de N^* experimentos simulados, isto é,

$$\hat{\alpha} = \frac{\sum_{k=1}^{N^*} I(E_k = 1)}{N^*},$$

em que E_k é uma variável binária que assume o valor 1 se ocorreu pelo menos um erro tipo I no k -ésimo experimento, para $k = 1, 2, \dots, N^*$ e $I(E_k = 1)$ é a função indicadora que retorna 1 se a igualdade for verificada e 0, caso contrário.

Na segunda etapa de avaliação do poder, os efeitos de tratamentos foram simulados com duas opções, para gerar uma situação de H_1 (hipótese alternativa) completa e de H_0 (hipótese nula) parcial. Assim, no primeiro caso, o efeito do tratamento 1 foi considerado igual a 0, isto é, $\tau_1 = 0$, e os demais fixados por

$$\tau_i = \tau_{i-1} + \delta \frac{\sigma}{\sqrt{r}},$$

para δ , $\delta = 1, 2, 4, 8, 16$ e 32 , representando o número de erros padrões da diferença entre médias para especificar o efeito dos tratamentos consecutivos, considerando $i = 2, 3, \dots, n$. Assim, o poder foi computado pela proporção de rejeições entre médias envolvendo múltiplos de δ , em relação ao número total de comparações envolvendo esta diferença. Assim, entre tratamentos consecutivos, por exemplo, tem-se $n - 1$ comparações por experimento e $N^*(n - 1)$ comparações no total, que corresponde ao poder de detectar δ erros padrões da diferença entre médias; da mesma forma, para os vizinhos com salto 2 (primeira e terceira médias, segunda e quarta até antepenúltima e última médias ordenadas), tem-se $n - 2$ comparações por experimentos envolvendo 2δ erros padrões da diferença de médias para ser detectada. Esse procedimento é feito para todos os casos, até que sejam comparadas a primeira e última médias, ou seja, $(n - 1)\delta$ erros padrões a serem detectados em apenas 1 comparação por experimento e um total de N^* comparações em relação a todos os experimentos simulados.

A segunda opção para o estudo do poder envolvendo uma simulação sob H_0 parcial envolveu simulação de dois grupos de médias, com $k_1 = \lfloor n/2 \rfloor$ e $k_2 = n - k_1$ médias em cada, em que $\lfloor x \rfloor$ refere-se ao maior inteiro menor ou igual a x . As médias

do primeiro grupo foram todas iguais, para as quais os efeitos foram $\tau_i = 0$, $i = 1, 2, 3, \dots, k_1$, sem perda de generalidade. O segundo grupo, com k_2 médias, teve seus efeitos também iguais a

$$\tau_i = \tau_1 + \delta \frac{\sigma}{\sqrt{r}}, \quad i = k_1 + 1, k_1 + 2, \dots, n,$$

em que diferentes valores de δ foram considerados como $\delta = 1, 2, 4, 8, 16$. Nesse caso, a proporção de rejeições envolvendo comparações dos diferentes grupos no total de $N^*k_1k_2$ comparações envolvendo médias dos dois grupos nos N^* experimentos simulados, forneceu uma estimativa do poder. As comparações intra grupos permitiram também avaliar as taxas de erro tipo I por experimento sob H_0 parcial. A proporção de experimentos com pelo menos uma rejeição da hipótese nula de igualdade entre duas médias intra grupos foi uma estimativa dessa taxa de erro dos testes. Todos os testes foram aplicados a cada um desses cenários simulados, as taxas de erro tipo I por experimento (comparação intra grupo) e poder (comparação inter grupos) foram computados e os resultados comparados entre si.

Foram consideradas algumas configurações em ambas as etapas com diferentes valores de n e de r . Assim, foram considerados os casos com $n = 5, 10, 20, 40$ e 100 , e $r = 4, 10$ e 20 . Também foram considerados os níveis de significância nominais de 1 e 5% . O coeficiente de variação (CV) do experimento adotado foi $CV = 10\%$, pois nos resultados encontrados em Borges e Ferreira (2003, p. 74, Tabela 1) e Silva, Ferreira e Bezzi (1999, p. 691), percebeu-se que os PCMs avaliados não foram influenciados quanto a variação do CV, considerando população normal, para a avaliação de desempenho do erro tipo I por experimento e poder, uma vez que, na simulação, quando as médias se diferenciavam era sempre em termos de erros padrões. Além do mais, foram feitas análises preliminares com os testes propostos e verificou-se esse mesmo comportamento. Por isso, as simulações foram fixadas em um único CV. Isso mostra um importante resultado, pois grande parte dos pesquisadores usam transformações ou outras técnicas para baixar o CV de seus experimentos, com medo da influência em seus resultados. Portanto, uma prática não indicada para esse intuito.

3.3 Pacote `midrangeMCP`

O pacote `midrangeMCP` (BATISTA; FERREIRA, 2016), desenvolvido no software R, versão 3.2.2 (R CORE TEAM, 2015), terá o objetivo de implementar os algoritmos dos testes propostos com o intuito de facilitar a utilização.

Esse pacote tem três funções: `MRtest`, `MRboxplot` e `MRwrite`. A função `MRtest` tem como objetivo aplicar os quatro testes: MGM teste, MGR teste, SNKM teste e TM teste. A grande versatilidade dessa função é que o objeto de entrada tanto pode ser um vetor da variável resposta e um vetor do efeito de tratamento quanto pode ser um objeto que armazenou o modelo ajustado aos seus dados. Nesse pacote a função que armazena o modelo ajustado deve ser `aov()` ou `lm()`. Outro ponto interessante é que os quatro testes podem ser aplicados de uma só vez. A função `MRboxplot` plotará o resultado dos testes por meio de um gráfico de barras identificando os grupos de médias de acordo com o teste desejado. E por último, a função `MRwrite` exportará os principais resultados dos testes utilizados na análise. Os arquivos exportados com os resultados apresentam as seguintes opções de extensão: “arquivo.csv”, “arquivo.txt”, “arquivo.xlsx” e em código latex. Este último será apresentado no próprio console do R.

O pacote `midrangeMCP` teve as seguintes pendências de outros pacotes: `SMR` (BATISTA; FERREIRA, 2014a), `xtable` (DAHL, 2014) e `WriteXLS` (SCHWARTZ, 2015). O primeiro pacote deu suporte para obter os quantis da distribuição da *midrange* estudentizada externamente, sendo usado na função `MRtest`. Os dois outros pacotes foram utilizados na função `MRwrite`, auxiliando na exportação dos arquivos em código latex e com a extensão “.xlsx”, respectivamente.

3.3.1 Instalação do pacote `midrangeMCP`

Para instalar o pacote basta entrar com os seguintes comandos no console do R:

```
install.packages("midrangeMCP") # instalando o pacote
library("midrangeMCP") # carregando o pacote
```

A seguir será dado mais detalhes sobre essas três funções do pacote `midrangeMCP`.

3.3.2 Função MRtest

A função MRtest é expressa da seguinte forma:

```
MRtest(y, trt = NULL, dferror = NULL, mserror = NULL,
       replication = NULL, alpha = 0.05, main = NULL,
       MCP = "all", ismean = FALSE)
```

sendo:

- `y`: vetor com as variáveis respostas, o modelo desejado ou a média dos tratamentos;
- `trt`: vetor com os níveis do fator em estudo. Caso `y` seja um objeto que recebe um modelo das funções `lm()` ou `aov()`, no argumento `trt` deve-se indicar qual o componente do modelo representa o efeito dos tratamentos. Por exemplo, se `y <- aov(vr ~ treat)`, então `trt = "treat"`. O *default* `trt = NULL` é uma proteção na função caso o usuário esqueça de inserir alguma informação para esse argumento;
- `dferror`: grau de liberdade do resíduo. Essa opção será necessária apenas quando o argumento `y` for um vetor com as variáveis respostas ou vetor de médias de tratamentos. Por isso o *default* desse argumento é `dferror = NULL`;
- `msferror`: quadrado médio do resíduo. Essa opção será necessária apenas quando o argumento `y` for um vetor com as variáveis respostas ou vetor de médias dos tratamentos. Por isso o *default* desse argumento é `mserror = NULL`;
- `alpha`: nível de significância adotado nos testes. O *default* é `alpha = 0.05`;
- `replication`: número de repetições dos tratamentos. Caso os dados sejam desbalanceados, deve-se informar a média harmônica dessas repetições. Esse argumento deve ser informado apenas quando o argumento `y` for um vetor de médias ou as repetições dos tratamentos forem desbalanceados. O *default* é `replication = NULL`;
- `main`: título de sua análise;
- `MCP`: permite escolher o teste de comparação múltipla desejado. As opções são: "MGM", "MGR", "SNKM", "TM" e "all". Este último é o *default*, que resulta em calcular os quatro testes de uma só vez. Caso queira alguns testes, mas não todos, esse argumento pode ser concatenado com as opções desejados. Por exemplo, caso queira realizar os testes de MGM e teste de TM, basta usar `MCP = c("MGM", "TM")`

- `ismean`: argumento lógico. Se `ismean = TRUE`, é porque o argumento `y` é um vetor de médias dos tratamentos. Caso contrário, `ismean = FALSE`.

A função `MRtest` independente de ser armazenado em um objeto, imprime os resultados da análise. Contudo, caso deseje obter alguns resultados armazenados no objeto, os resultados são divididos em quatro partes:

- `Summary`: as estatísticas descritivas da variável resposta;
- `Statistics`: as principais estatísticas para a realização do teste;
- `Groups`: resultado do grupo de médias para cada teste;
- `Tests`: armazena o nome dos testes selecionados.

Entretanto, o objetivo de criar um objeto para armazenar esses resultados é a facilidade que o usuário terá para usar as outras duas funções do pacote `MRbarplot` e `MRwrite`.

3.3.3 Função `MRbarplot`

A função `MRbarplot` é expressa da seguinte forma:

```
MRbarplot(x, MCP = "all", col = heat.colors(10),
          horiz = FALSE, ...)
```

sendo:

- `x`: objeto da função `MRtest`;
- `MCP`: permite escolher o teste de comparação múltipla desejado. O *default* é `"all"`. Essa opção realizará todos os testes que foram escolhidos no objeto `MRtest`. Caso queira alguns dos testes utilizados nesse objeto, pode-se concatená-lo. Ver o argumento `MCP` da função `MRtest`, seção 3.3.2;
- `col`: permite escolher a cor das barras do gráfico. O *default* é `heat.colors(10)`. O valor 10 significa o número de cores desejado. Caso queira o gráfico com coloração na escala preto e branco, pode-se escolher a opção `gray.colors(10)`;
- `horiz`: escolhe a direção das barras. O *default* é `horiz = FALSE`, isto é, barras na vertical. Caso escolha `horiz = TRUE`, as barras serão plotadas na horizontal;
- ... a função permite que o usuário possa acrescentar mais argumentos da função `barplot`.

3.3.4 Função `MRwrite`

A função `MRwrite` é expressa da seguinte forma:

```
MRwrite(x, MCP = "all", extension = "csv",
        dataMR = "all")
```

sendo:

- `x`: objeto da função `MRtest`;
- `MCP`: permite escolher o teste de comparação múltipla desejado. O *default* é `MCP = "all"`. Essa opção realizará todos os testes que foram escolhidos no objeto `MRtest`. Caso queira alguns dos testes utilizados nesse objeto, pode-se concatená-lo. Ver o argumento `MCP` da função `MRtest`, seção 3.3.2;
- `extension`: tipo de formato de arquivo. Existem quatro opções: `"csv"`, `"txt"`, `"xlsx"` e `"latex"`. O *default* é `"csv"`.
- `dataMR`: permite escolher os resultados a serem exportados. Há três opções: `"groups"`, `"summary"` or `"all"`. A opção `"groups"` exporta os grupos de médias de tratamentos avaliados pela escolha do teste no argumento `MCP`. A opção `"summary"` exporta as estatísticas descritivas da variável resposta. A opção `"all"` exporta ambos os resultados.

Todos os arquivos são exportados para o diretório de trabalho, exceto a extensão `"latex"` que o código é exportado no próprio console do R.

3.3.5 Aplicações do Pacote `MidrangeMCP`

A aplicação do pacote `MidrangeMCP` será feita utilizando dois experimentos. Inicialmente, será abordado o desenvolvimento analítico dos testes, e posteriormente, detalhes que como proceder utilizando o pacote.

O primeiro experimento foi retirado de Steel e Torrie (1980, p. 180) que representará o Exemplo 3.1.

■ **Exemplo 3.1** O experimento foi realizado no Delineamento Inteiramente Causalizado (DIC) e objetiva avaliar o efeito das bactérias no teor de nitrogênio das plantas de trevo vermelho. Esse experimento apresentou os seguintes dados:

Para esse exemplo, será mostrado duas formas de entrada de dados para a execução da função `MRtest`. A primeira, o argumento `y` da função `MRtest` recebendo o objeto que armazenou o resultado da função `aov()`, e o segundo o argumento `y` recebendo o vetor da variável resposta (teor de nitrogênio). Posteriormente, recursos gráficos serão apresentados, bem como a exportação dos resultados dos testes executados.

Tabela 2 Dados de teor de nitrogênio das plantas de trevo vermelho relacionados as bactérias em estudo.

Tipo de Bactéria	Repetições				
	R_1	R_2	R_3	R_4	R_4
3DOK1	19,4	32,6	27,0	32,1	33,0
3DOK5	17,7	24,8	27,9	25,2	24,3
3DOK4	17,0	19,4	9,1	11,9	15,8
3DOK7	20,7	21,0	20,5	18,8	18,6
3DOK13	14,3	14,4	11,8	11,6	14,2
COMPOS	17,3	19,4	19,1	16,9	20,8

O segundo experimento foi retirado de Figueiredo et al. (2015) que representará o Exemplo 3.2. Estes dados representam o estudo de seleção de 25 genótipos da cultura do sorgo, dos quais uma das avaliações foi baseada no período de floração (essa variável foi estimada como a época em que 50% das plantas na parcela estavam em floração).

■ **Exemplo 3.2** O experimento foi realizado num delineamento em látice triplo avaliando os genótipos em 7 ambientes. Para esse estudo não será levado em consideração a avaliação dos ambientes, pois extrapola o objetivo de estudo desse trabalho. A avaliação do teste de Skott-Knott para o período de floração também foi realizado por Figueiredo et al. (2015, p. 11219), sendo apresentada na Tabela 3. As informações adicionais desse estudo foram: a análise de variância cujo quadrado médio do resíduo foi de 6,3078 para 252 graus de liberdade. O número de repetições com que as médias dos genótipos foram estimados foi igual a 21. ■

Uma entrada de dados não muito comum nas rotinas é a média dos tratamentos, do qual será apresentado nesse exemplo. Nesse exemplo, será mostrado que informando apenas os resultados do quadrado médio do resíduo, o grau de liberdade e o número de repetições, a função MR_{test} poderá realizar o procedimento dos quatro testes propostos.

Tabela 3 Seleção de vinte e cinco genótipos de sorgo com base no período de floração, avaliados pelo teste de Skott-Knott.

Genótipo	Período de floração	Teste de Skott-Knott
CMSXS643	87,51	A
CMSXS630	85,78	B
BR507	85,17	B
BR506	84,33	B
BR508	83,70	C
CMSXS629	83,27	C
BR501	83,25	C
CMSXS635	82,48	C
CMSXS644	82,35	C
BRS511	81,42	D
CMSXS648	81,12	D
CMSXS633	80,91	D
BR505	79,91	E
CMSXS637	79,59	E
XBSW80140	79,35	E
CMSXS646	78,59	E
BRS601	78,33	E
CMSXS639	78,15	E
CMSXS647	77,29	E
SUGARGRAZE	75,45	F
CMSXS636	75,43	F
V82391	75,36	F
XBSW80007	75,15	F
V82393	73,83	G
V82392	72,52	G

4 RESULTADOS

A seguir, será apresentada a avaliação de desempenho dos testes propostos comparando com os resultados dos testes já existentes na literatura, dando ênfase aos testes de Tukey, de SNK e de Scott-Knott, pois foram estes que serviram de base para os testes propostos. A avaliação de desempenho terá por base dois pontos, o erro tipo I e o poder do teste. Diversos arranjos foram escolhidos para a avaliação de desempenho, como já falado na Seção 3.2. Os resultados serão discutidos e apresentados por meio de tabelas e gráficos para facilitar a exposição e interpretação.

4.1 Erro tipo I

A primeira avaliação dos testes foi baseada nas taxas de erro tipo I por experimento, como já justificado na subseção 2.1.1. Foram avaliados dois cenários para computar o erro tipo I por experimento: sob H_0 completa, Definição 2.1, e sob H_{0_p} parcial, Definição 2.3.

4.1.1 Erro tipo I por experimento sob H_0 completa

Nas Tabelas 4 e 5 são apresentados os resultados do erro tipo I por experimento sob H_0 completa.

Para avaliar o desempenho, os testes propostos de TM, de SNKM, de MGM e de MGR foram comparados com os testes de Tukey e de SNK. Para esses dois últimos testes, além da simulação realizada no presente trabalho, foram utilizados resultados de outras avaliações de desempenho encontrados na literatura. Outros PCMs encontrados na literatura também foram comentados nessa discussão.

Observou-se que os testes propostos controlaram a taxa de erro tipo I por experimento, pois nenhum destes tiveram a TEE empírica rejeitada pelo teste binomial exato, tal que $F \geq F_{0,995}$. Entretanto, em alguns casos, os níveis nominais empíricos para os testes de TM, de SNKM e de MGM foram rejeitados pelo teste binomial exato, tal que $F \leq F_{0,995}$, tornando-os conservativos. Confirmando os resultados do presente trabalho (Tabelas 4 e 5), Carmer e Swanson (1973, p. 69) e Borges e Ferreira (2003, p. 71) também mostraram que os testes de Tukey e de SNK apresentam controle do erro tipo I por experimento.

Tabela 4 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 completa, ao nível de significância $\alpha = 1\%$ de probabilidade, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Rep	Trat	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
4	5	1,160	1,160	1,160	0,640 ⁻⁻	0,640 ⁻⁻	0,640 ⁻⁻
	10	1,100	1,100	1,100	0,720	0,720	0,720
	20	1,040	1,040	1,040	0,900	0,900	0,900
	30	0,820	0,820	0,820	0,660 ⁻⁻	0,660 ⁻⁻	0,660 ⁻⁻
	40	0,820	0,820	0,820	0,700	0,700	0,700
	100	1,060	1,060	1,060	0,600 ⁻⁻	0,600 ⁻⁻	0,600 ⁻⁻
10	5	0,980	0,980	0,980	0,460 ⁻⁻	0,460 ⁻⁻	0,460 ⁻⁻
	10	1,000	1,000	1,000	0,740	0,740	0,740
	20	0,900	0,900	0,900	0,780	0,780	0,780
	30	1,240	1,240	1,240	0,560 ⁻⁻	0,560 ⁻⁻	0,560 ⁻⁻
	40	0,880	0,880	0,880	0,860	0,860	0,860
	100	0,940	0,940	0,940	0,580 ⁻⁻	0,580 ⁻⁻	0,580 ⁻⁻
20	5	0,940	0,940	0,940	0,500 ⁻⁻	0,500 ⁻⁻	0,500 ⁻⁻
	10	1,080	1,080	1,080	0,880	0,880	0,880
	20	1,100	1,100	1,100	0,720	0,720	0,720
	30	1,020	1,020	1,020	0,520 ⁻⁻	0,520 ⁻⁻	0,520 ⁻⁻
	40	0,840	0,840	0,840	0,600 ⁻⁻	0,600 ⁻⁻	0,600 ⁻⁻
	100	1,020	1,020	1,020	0,700	0,700	0,700

* O símbolo “- -” indica que a TEE foi rejeitada pelo teste binomial exato, tal que $F \leq F_{0,005}$. O símbolo “++” indica que a TEE foi rejeitada pelo teste binomial exato, tal que $F \geq F_{0,995}$.

Silva, Ferreira e Bearzoti (1999, p. 691) e Borges e Ferreira (2003, p. 70), avaliando o desempenho do teste de Scott-Knott e considerando as mesmas configurações de simulação do presente trabalho, observaram que alguns valores das taxas de erro tipo I por experimento foram superiores aos níveis nominais de significância (α) de 1% e 5%. Os valores das TEE que ultrapassaram o valor de α (testes liberais) foram aqueles em que o número de tratamentos era 5, embora não tivessem se distanciado muito dos valores nominais.

Independentemente do número de repetições, os testes propostos controlaram as taxas de erro tipo I por experimento (Tabelas 4 e 5). Isso também foi verificado por Borges e Ferreira (2003, p. 75) quando avaliaram o desempenho dos testes de Tukey e o teste de SNK. Eles usaram a mesma metodologia de simulação do presente trabalho, em relação ao número de repetições e ao coeficiente de variação (CV).

A razão para isso é a simulação adotada. Os parâmetros dos tratamentos estão vinculados ao número de repetições e a diferença entre médias é sempre preservada em

Tabela 5 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 completa, ao nível de significância $\alpha = 5\%$ de probabilidade, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Rep	Trat	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
4	5	5,240	5,240	5,240	3,680 ⁻⁻	3,680 ⁻⁻	3,680 ⁻⁻
	10	5,660	5,660	5,660	4,720	4,720	4,720
	20	5,080	5,080	5,080	5,060	5,060	5,060
	30	4,960	4,960	4,960	4,340	4,340	4,340
	40	4,980	4,980	4,980	3,980 ⁻⁻	3,980 ⁻⁻	3,980 ⁻⁻
	100	4,680	4,680	4,680	3,340 ⁻⁻	3,340 ⁻⁻	3,340 ⁻⁻
10	5	4,940	4,940	4,940	3,860 ⁻⁻	3,860 ⁻⁻	3,860 ⁻⁻
	10	5,060	5,060	5,060	4,820	4,820	4,820
	20	5,240	5,240	5,240	5,140	5,140	5,140
	30	4,840	4,840	4,840	4,160 ⁻⁻	4,160 ⁻⁻	4,160 ⁻⁻
	40	4,620	4,620	4,620	4,020 ⁻⁻	4,020 ⁻⁻	4,020 ⁻⁻
	100	5,140	5,140	5,140	3,700 ⁻⁻	3,700 ⁻⁻	3,700 ⁻⁻
20	5	4,880	4,880	4,880	2,540 ⁻⁻	2,540 ⁻⁻	2,540 ⁻⁻
	10	5,060	5,060	5,060	4,440	4,440	4,440
	20	4,940	4,940	4,940	4,760	4,760	4,760
	30	4,960	4,960	4,960	4,120 ⁻⁻	4,120 ⁻⁻	4,120 ⁻⁻
	40	5,020	5,020	5,020	4,180 ⁻⁻	4,180 ⁻⁻	4,180 ⁻⁻
	100	4,820	4,820	4,820	3,720 ⁻⁻	3,720 ⁻⁻	3,720 ⁻⁻

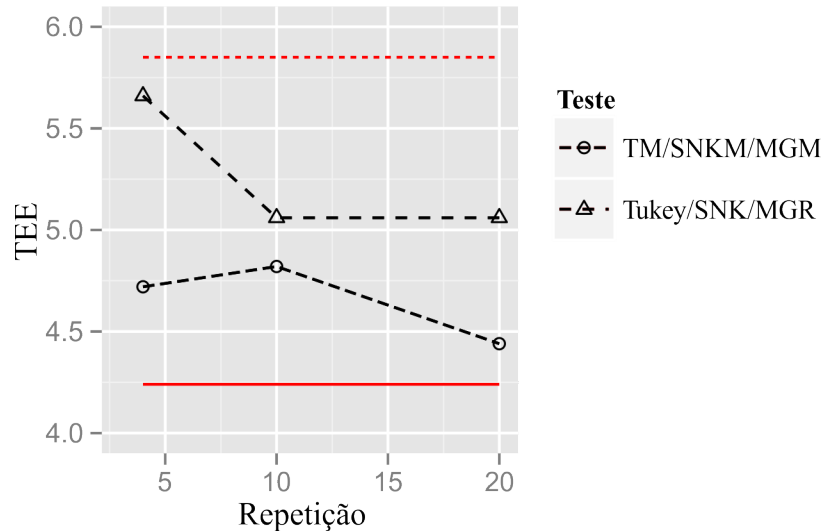
* O símbolo “- -” indica que a TEE foi rejeitada pelo teste binomial exato, tal que $F \leq F_{0,005}$. O símbolo “++” indica que a TEE foi rejeitada pelo teste binomial exato, tal que $F \geq F_{0,995}$.

termos de erro padrão, que por conseguinte, está relacionado com o CV e com o número de repetições. Entretanto, considerando o teste de Scott-Knott, esses mesmos autores observaram que para um número grande de repetições, $r = 20$, apenas quando o número de tratamentos foi pequeno, $n = 5$, o teste tornava-se liberal.

Nas Figuras 3 e 4 que representam a avaliação da taxa de erro tipo I por experimento, pode-se observar duas linhas vermelhas, uma está cheia e a outra tracejada. A primeira delimita a rejeição em que os valores de TEE abaixo dessa linha foram inferiores ao nível nominal global, isto é, a hipótese $H_0 : \alpha = 5\%$ foi rejeitada pelo teste binomial exato, pois $F \leq F_{0,005}$. Assim, trata-se de um teste conservativo. A segunda delimita a região em que os valores de TEE acima dessa linha, foram superiores ao nível nominal global, isto é, a hipótese $H_0 : \alpha = 5\%$ foi rejeitada pelo teste binomial exato, pois $F > F_{0,995}$. Trata-se de um teste liberal.

Na Figura 3, todos os testes propostos e os de Tukey e de SNK controlaram o erro tipo I por experimento, pois nenhum dos testes avaliados pelo teste binomial exato

ultrapassaram as linhas vermelhas identificadoras da rejeição da hipótese H_0 .

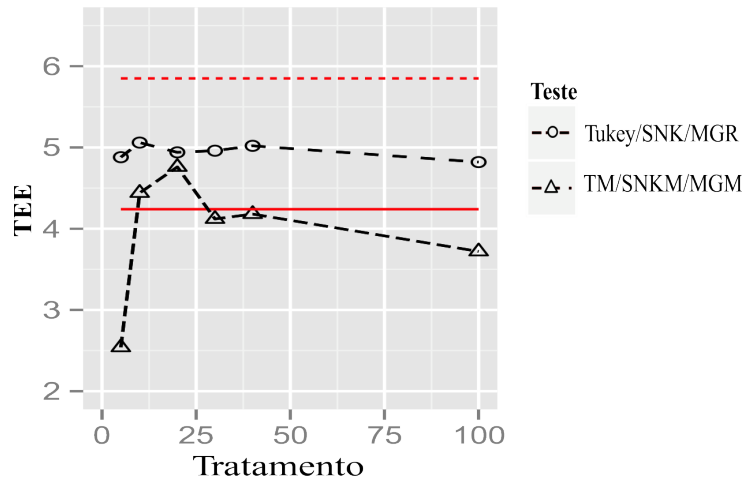


* As linhas de vermelho delimitam a região de rejeição pelo teste binomial exato.

Figura 3 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de MGM, de TM, de SNKM, em função do número de repetições, hipótese H_0 completa, $n = 10$, $\alpha = 5\%$.

Com relação ao número de tratamentos, tendo por base uma representação gráfica para o arranjo $r = 20$ e $\alpha = 5\%$ na Figura 4, verificou-se que os testes de MGR, de TM, de SNKM e de MGM controlam o erro tipo I por experimento. Percebeu-se que os testes de TM, de SNKM e de MGM na medida em que o número de tratamentos aumenta, a taxa de erro por experimento diminui, a ponto de serem conservativos tanto para $\alpha = 1\%$ quanto para $\alpha = 5\%$ (Tabelas 4 e 5). As demais configurações simuladas podem ser observadas nas Tabelas 4 e 5, sendo que os resultados foram similares aos apresentados na Figura 4. Carmer e Swanson (1973, p. 70) e Boardman e Moffitt (1971, p. 743) verificaram esse mesmo comportamento para o teste de Scheffé, considerando 4000 experimentos. Para $n = 20$ tratamentos, a TEE deste teste chegou a praticamente 0% de erro tipo I por experimento, um teste muito conservativo.

Para os testes de Tukey, de SNK e de Scott-Knott, independente do número de tratamentos, considerando uma população normal, Borges e Ferreira (2003, p. 71) mostraram que as taxas de erro tipo I por experimento permanecem iguais ao nível de significância global. Entretanto, quando consideradas populações não-normais, os testes de Tukey e de SNK apresentaram TEE na ordem 55% com 80 tratamentos para a distribui-



* As linhas de vermelho delimitam a região de rejeição pelo teste binomial exato.

Figura 4 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de MGM, de TM, de SNKM, em função do número de tratamentos, hipótese H_0 completa, $r = 20$, $\alpha = 5\%$.

ção log-normal. Esses mesmos autores constataram também que o teste de Scott-Knott apresenta um certo controle do erro tipo I por experimento considerando populações não-normais, já que o viés positivo na taxa de controle do erro era muito pequeno, apresentando uma certa robustez. Com esse resultado, o teste de Scott-Knott pode ser uma alternativa de PCM para ser utilizado quando não for verificada a pressuposição de normalidade dos erros no experimento.

Ao contrário desses PCMs, Bernhardson (1975, p. 231) mostrou que os testes LSD (teste baseado na distribuição t de *Student*) e de Duncan, considerando 10 tratamentos, apresentam altas taxas de erro tipo I, 49,0% e 36,3%, respectivamente. Esse fato também foi confirmado em Boardman e Moffitt (1971, p. 742) e Carmer e Swanson (1973, p. 70).

Percin e Malheiros (1989, p. 51) avaliaram o teste t-bayesiano, proposto por Waller e Duncan (1969). Verificaram-se altas taxas de erro tipo I por experimento para esse teste. Carmer e Swanson (1973, p. 70), estudando também o teste t-bayesiano, observaram que para os números de tratamentos iguais a 5, 10 e 20 e nível de significância $\alpha = 5\%$, os valores das taxas de erros tipo I por experimento foram 15,6%, 18,4% e 18,7%, respectivamente, confirmando tratar-se de um teste liberal.

Fato interessante para os testes propostos de TM, de SNKM e de MGM é que

apresentam TEE idênticas, assim como os testes de Tukey, de SNK e de MGR, independente do número de repetições e de tratamentos, sob H_0 completa. Isso se deve a semelhança no desenvolvimento teórico dos testes. Por exemplo, o teste de Tukey e de SNK para a primeira diferença entre as médias extremas (menor média e a maior média), apresentam a mesma DMS, como observado por Carmer e Swanson (1973, p. 67). Fato também foi observado em Borges e Ferreira (2003, p. 74, Tabela 1), considerando a distribuição normal e $CV = 10\%$ (mesmas condições da simulação desse estudo), em que os testes de Tukey e de SNK apresentam TEE iguais.

Com base nos resultados obtidos por em simulação e os encontrados na literatura, na Figura 5 é apresentado um resumo da avaliação do erro tipo I por experimento, sob H_0 completa, para os testes propostos e os testes citados na literatura para comparação, com base em seu rigor, isto é, qual teste é mais conservativo.

No entanto, o que se observa nos experimentos, é que grupos de médias de tratamentos sejam iguais, e muito raramente, que essas médias sejam todas iguais. Com base nisso, na subseção seguinte, foi considerado o cenário no qual a simulação baseou-se na hipótese nula parcial.

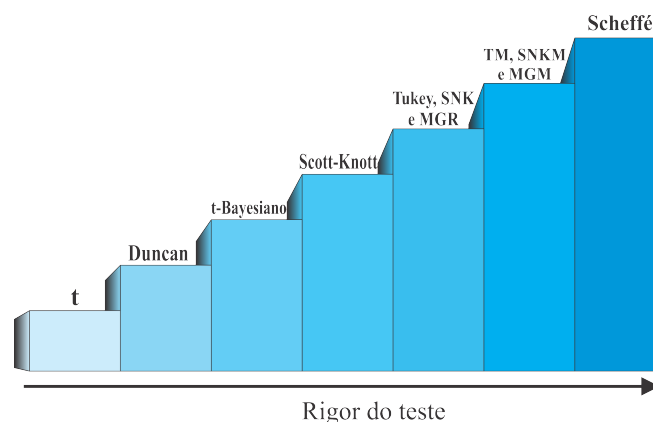


Figura 5 Ordem crescente do rigor do erro tipo I dos procedimentos de comparações múltiplas estudados.

4.1.2 Erro tipo I por experimento sob H_0 parcial

Outra forma de avaliar o erro tipo I é por meio de simulações levando em consideração a hipótese nula parcial (H_{0p}), ver Tabelas 6 a 8 e Apêndice A.

Tabela 6 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 5\%$ de probabilidade e $\delta = 1$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Trat	Rep	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	2,700 ⁻⁻	3,600 ⁻⁻	8,560 ⁺⁺	4,700	4,040 ⁻⁻	5,220
	10	2,820 ⁻⁻	3,740 ⁻⁻	9,960 ⁺⁺	4,200 ⁻⁻	3,760 ⁻⁻	4,760
	20	2,240 ⁻⁻	3,240 ⁻⁻	9,420 ⁺⁺	3,980 ⁻⁻	3,500 ⁻⁻	4,620
10	4	2,980 ⁻⁻	3,560 ⁻⁻	10,820 ⁺⁺	7,480 ⁺⁺	7,040 ⁺⁺	9,400 ⁺⁺
	10	2,880 ⁻⁻	3,560 ⁻⁻	12,360 ⁺⁺	7,220 ⁺⁺	6,840 ⁺⁺	9,080 ⁺⁺
	20	2,460 ⁻⁻	3,000 ⁻⁻	12,900 ⁺⁺	6,280 ⁺⁺	5,780 ⁺⁺	8,080 ⁺⁺
20	4	3,140 ⁻⁻	3,500 ⁻⁻	13,120 ⁺⁺	7,700 ⁺⁺	7,600 ⁺⁺	8,580 ⁺⁺
	10	2,740 ⁻⁻	3,180 ⁻⁻	13,440 ⁺⁺	6,800 ⁺⁺	6,780 ⁺⁺	7,900 ⁺⁺
	20	2,520 ⁻⁻	3,000 ⁻⁻	14,160 ⁺⁺	6,920 ⁺⁺	6,820 ⁺⁺	7,640 ⁺⁺
40	4	3,140 ⁻⁻	3,480 ⁻⁻	15,400 ⁺⁺	5,460	5,460	5,600 ⁺⁺
	10	3,400 ⁻⁻	3,400 ⁻⁻	16,000 ⁺⁺	5,600	5,560	5,820
	20	2,700 ⁻⁻	2,880 ⁻⁻	14,980 ⁺⁺	5,240	5,240	5,440
100	4	3,200 ⁻⁻	3,380 ⁻⁻	17,960 ⁺⁺	4,560	4,560	4,560
	10	3,380 ⁻⁻	3,460 ⁻⁻	19,400 ⁺⁺	4,720	4,720	4,720
	20	2,920 ⁻⁻	3,020 ⁻⁻	18,320 ⁺⁺	4,200 ⁻⁻	4,200 ⁻⁻	4,200 ⁻⁻

* O símbolo “- -” indica que a TEE foi rejeitada pelo teste binomial exato, tal que $F \leq F_{0,005}$. O símbolo “++” indica que a TEE foi rejeitada pelo teste binomial exato, tal que $F \geq F_{0,995}$.

Diferentemente dos trabalhos de Silva, Ferreira e Bearzoti (1999) e Borges e Ferreira (2003), para que houvesse mais informações sobre o erro tipo I por experimento sob H_0 parcial, o número de diferenças entre médias consecutivas foi maior, isto é, $1\sigma_{\bar{Y}}$, $4\sigma_{\bar{Y}}$, $8\sigma_{\bar{Y}}$, $16\sigma_{\bar{Y}}$ e $32\sigma_{\bar{Y}}$ erros padrões.

Na Figura 6 é apresentada a avaliação de desempenho dos testes com relação a diferença de médias consecutivas (δ), fixando o número de tratamentos ($n = 5, 20$ e 100) e o número de repetições ($r = 10$). Observou-se, de um modo geral, que os testes propostos ultrapassam os níveis de significância estabelecidos, principalmente quando a diferença de grupos de médias consecutivas é superior a $2\sigma_{\bar{Y}}$.

Pela simulação realizada no presente estudo, o teste de Tukey é o único teste com TEE igual ao nível de significância global, independentemente da configuração do experimento. Isso também foi verificado no trabalho de Borges e Ferreira (2003, p. 76). Esses autores confirmaram o mesmo comportamento para o teste de SNK. Entretanto, quando a diferença entre grupos de médias consecutivas foi superior a $4\sigma_{\bar{Y}}$, a TEE desse teste superou o nível nominal tanto para $\alpha = 0,01$ quanto para $\alpha = 0,05$, o que o caracteriza como um teste liberal. Esse resultado também foi confirmado no presente

Tabela 7 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 5\%$ de probabilidade e $\delta = 2$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Trat	Rep	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	2,460 ⁻⁻	5,500	18,420 ⁺⁺	9,280 ⁺⁺	6,420 ⁺⁺	10,760 ⁺⁺
	10	2,540 ⁻⁻	5,600	21,700 ⁺⁺	8,720 ⁺⁺	6,140 ⁺⁺	10,140 ⁺⁺
	20	2,240 ⁻⁻	3,240 ⁻⁻	9,420 ⁺⁺	12,150 ⁺⁺	9,820 ⁺⁺	15,020 ⁺⁺
10	4	2,600 ⁻⁻	4,640	31,220 ⁺⁺	13,400 ⁺⁺	11,160 ⁺⁺	22,560 ⁺⁺
	10	3,140 ⁻⁻	5,260	38,220 ⁺⁺	12,320 ⁺⁺	9,920 ⁺⁺	21,620 ⁺⁺
	20	2,460 ⁻⁻	3,000 ⁻⁻	12,900 ⁺⁺	13,460 ⁺⁺	11,640 ⁺⁺	22,140 ⁺⁺
20	4	3,480 ⁻⁻	4,800	42,920 ⁺⁺	14,200 ⁺⁺	12,780 ⁺⁺	22,220 ⁺⁺
	10	3,120 ⁻⁻	4,380	48,640 ⁺⁺	13,240 ⁺⁺	12,120 ⁺⁺	20,900 ⁺⁺
	20	2,520 ⁻⁻	3,000 ⁻⁻	12,900 ⁺⁺	12,920 ⁺⁺	11,700 ⁺⁺	20,640 ⁺⁺
40	4	2,820 ⁻⁻	3,800 ⁻⁻	53,500 ⁺⁺	10,620 ⁺⁺	10,280 ⁺⁺	13,240 ⁺⁺
	10	2,980 ⁻⁻	3,860 ⁻⁻	58,160 ⁺⁺	9,020 ⁺⁺	8,740 ⁺⁺	11,460 ⁺⁺
	20	2,520 ⁻⁻	3,000 ⁻⁻	14,160 ⁺⁺	9,680 ⁺⁺	9,320 ⁺⁺	12,680 ⁺⁺
100	4	3,360 ⁻⁻	3,800 ⁻⁻	67,120 ⁺⁺	5,660	5,660	5,760
	10	2,980 ⁻⁻	3,440 ⁻⁻	69,600 ⁺⁺	5,020	5,000	5,260
	20	2,700 ⁻⁻	2,880 ⁻⁻	14,980 ⁺⁺	5,680	5,660	5,780

* O símbolo “- -” indica que a TEE foi rejeitada pelo teste binomial exato, tal que $F \leq F_{0,005}$. O símbolo “++” indica que a TEE foi rejeitada pelo teste binomial exato, tal que $F \geq F_{0,995}$.

trabalho, ver Tabelas 8 a 13, e no apêndice A, Tabelas 16 a 19.

O teste de MGR apresentou controle no nível de significância global quando a diferença entre grupos de médias consecutivas foi maior ou igual a $8\sigma_{\bar{Y}}$. Os testes de TM, de SNKM e de MGM controlaram o nível de significância global apenas para o número de tratamentos grande ($n = 100$) e $\delta \leq 2$, ver Tabelas 6 e 7, e Apêndice A nas Tabelas 14 e 15.

Carmer e Swanson (1973, p. 70) afirmaram que além do teste de Tukey, outro teste que apresenta TEE de acordo com o nível de significância global é o teste de Scheffé, com valores das TEEs até mais baixos que os apresentados pelo teste de Tukey. Nesse mesmo trabalho, observou-se também que os testes de Duncan e o t-bayesiano, apresentam as maiores taxas de erro tipo I por experimento sob H_0 parcial, sendo o último de maior magnitude, chegando a ordem de 58% de rejeição.

O teste de Scott-Knott, avaliado em Silva, Ferreira e Bearzoti (1999, p. 694), ultrapassou os níveis nominais estabelecidos com viés muito alto. Entretanto, Borges e Ferreira (2003, p. 74) verificaram apenas uma pequena influência no aumento da TEE com relação a variação em δ . O que pode estar provocando a discordância dos resultados

Tabela 8 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 5\%$ de probabilidade e $\delta = 4\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança 99% de probabilidade.

Trat	Rep	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	2,560 ⁻⁻	9,740 ⁺⁺	21,720 ⁺⁺	23,020 ⁺⁺	11,940 ⁺⁺	24,880 ⁺⁺
	10	2,820 ⁻⁻	9,620 ⁺⁺	23,420 ⁺⁺	23,240 ⁺⁺	11,800 ⁺⁺	25,240 ⁺⁺
	20	2,460 ⁻⁻	9,360 ⁺⁺	23,100 ⁺⁺	23,040 ⁺⁺	11,740 ⁺⁺	25,500 ⁺⁺
10	4	2,980 ⁻⁻	8,860 ⁺⁺	37,680 ⁺⁺	31,080 ⁺⁺	18,240 ⁺⁺	51,320 ⁺⁺
	10	2,560 ⁻⁻	9,100 ⁺⁺	40,040 ⁺⁺	33,200 ⁺⁺	19,120 ⁺⁺	53,220 ⁺⁺
	20	2,920 ⁻⁻	9,580 ⁺⁺	40,000 ⁺⁺	33,320 ⁺⁺	19,100 ⁺⁺	53,800 ⁺⁺
20	4	3,000 ⁻⁻	8,540 ⁺⁺	57,660 ⁺⁺	26,780 ⁺⁺	17,820 ⁺⁺	64,980 ⁺⁺
	10	2,660 ⁻⁻	9,100 ⁺⁺	58,240 ⁺⁺	28,280 ⁺⁺	18,080 ⁺⁺	64,460 ⁺⁺
	20	3,140 ⁻⁻	9,500 ⁺⁺	57,480 ⁺⁺	27,860 ⁺⁺	18,300 ⁺⁺	65,380 ⁺⁺
40	4	3,540 ⁻⁻	8,660 ⁺⁺	78,060 ⁺⁺	15,100 ⁺⁺	10,900 ⁺⁺	58,720 ⁺⁺
	10	2,600 ⁻⁻	8,360 ⁺⁺	78,120 ⁺⁺	14,600 ⁺⁺	10,280 ⁺⁺	57,140 ⁺⁺
	20	2,920 ⁻⁻	8,840 ⁺⁺	76,840 ⁺⁺	14,860 ⁺⁺	10,900 ⁺⁺	58,000 ⁺⁺
100	4	3,000 ⁻⁻	7,800 ⁺⁺	97,080 ⁺⁺	6,540 ⁺⁺	5,340	32,720 ⁺⁺
	10	3,280 ⁻⁻	8,260 ⁺⁺	96,980 ⁺⁺	6,760 ⁺⁺	5,220	32,980 ⁺⁺
	20	2,940 ⁻⁻	7,720 ⁺⁺	96,440 ⁺⁺	6,480 ⁺⁺	5,180	32,720 ⁺⁺

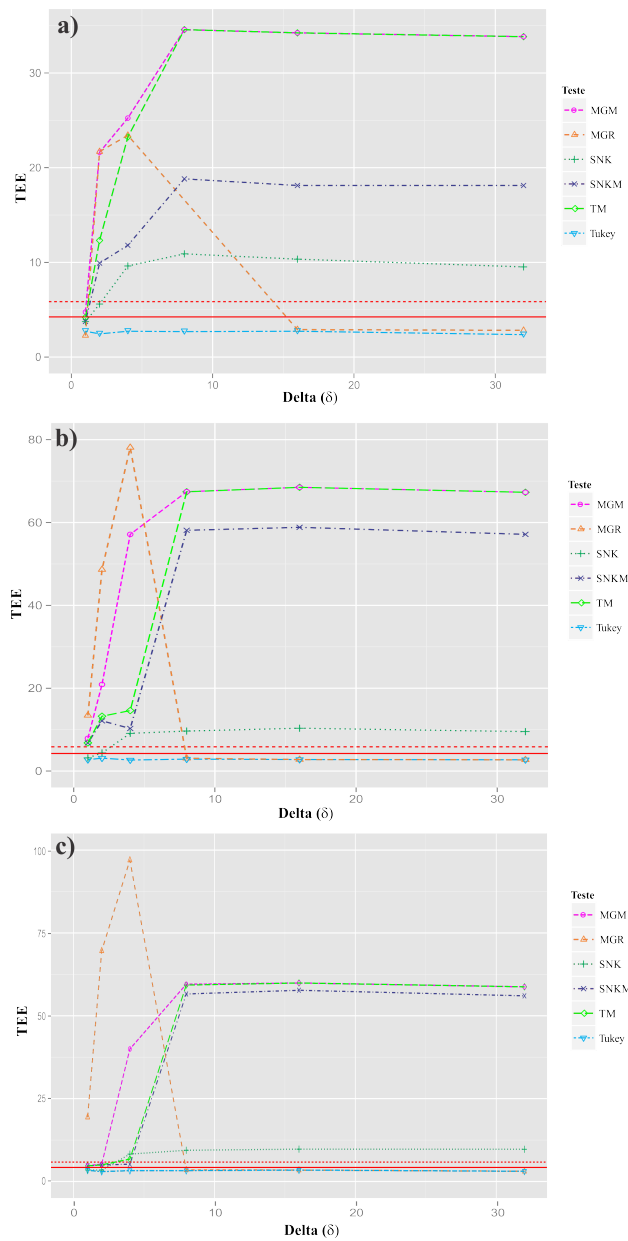
* O símbolo “- -” indica que a TEE foi rejeitada pelo teste binomial exato, tal que $F \leq F_{0,005}$. O símbolo “++” indica que a TEE foi rejeitada pelo teste binomial exato, tal que $F \geq F_{0,995}$.

é o processo de simulação para a avaliação do teste de Scott-knott. O processo de simulação do primeiro estudo assemelha-se aos que foram realizados no presente trabalho.

Apenas o teste de Scott-Knott sofreu uma pequena influência no aumento da TEE com relação ao aumento no número de repetições (BORGES; FERREIRA, 2003, p. 73). Os demais testes não sofreram influências expressivas, como pode ser visto na Figura 7, do qual são apresentados os resultados das TEEs para os testes propostos com relação ao número de repetições, para $n = 10$, $\alpha = 5\%$ e diferença entre médias de grupos consecutivas de $4\sigma_{\bar{Y}}$.

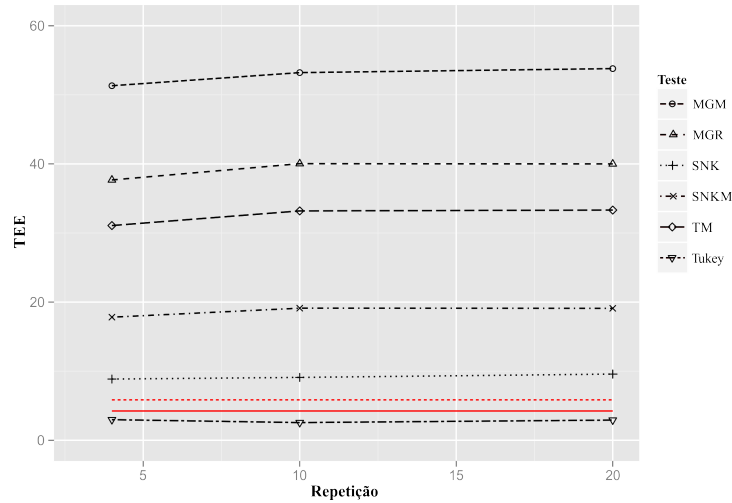
O número de tratamentos de um modo geral não exerceu influência no controle da TEE ao nível de significância global, sob H_0 parcial (Figura 8). Mas os comportamentos dos testes foram diferentes entre si. Inicialmente, o teste de Tukey preservou a TEE no nível de significância quando se variou o número de tratamentos. Comparando com a avaliação de desempenho feita por Borges e Ferreira (2003, p. 76), percebeu-se que os resultados da TEE para o teste de Tukey são semelhantes.

O teste de SNK, para $\delta \leq 2$, apresentou a TEE de acordo com o nível de significância adotado, independente do número de tratamentos, ver Tabelas 6 e 7, e no Apêndice



* As linhas de vermelho delimitam a região de rejeição pelo teste binomial.

Figura 6 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função da diferença de médias consecutivas (δ), sob hipótese H_0 parcial, a) $n = 5$, b) $n = 20$ e c) $n = 100$, para $\alpha = 5\%$ e $r = 10$.



* As linhas de vermelho delimitam a região de rejeição pelo teste binomial exato.

Figura 7 Taxa de erro tipo I por experimento, em porcentagem, dos testes de de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de repetições, sob hipótese H_0 parcial, para $n = 10$, $\alpha = 5\%$.

A nas Tabelas 14 e 15. Entretanto, quando $\delta > 2$ o teste apresentou a $TEE > \alpha$, tornando-se um teste liberal. Comparando essa avaliação de desempenho do teste de SNK com as avaliações realizadas em Borges e Ferreira (2003, p. 76), percebeu-se que apesar de esse teste em simulação ter sido liberal, ele apresenta valores de TEEs discrepantes. Para esses autores, considerando $\delta = 4$, $\alpha = 5\%$, $r = 10$ e $n = 40$, a TEE foi de 41,25%. Nessa mesma configuração, pela simulação realizada no presente estudo (Tabela 8), obteve a TEE de 8,36%.

Observou-se que o teste de MGR, para $\delta < 8$, aumenta consideravelmente os valores da TEE com o aumento do número de tratamentos. Para $\delta \geq 8$, a TEE desse teste estabiliza e passa a ser idêntico ao nível de significância global. Quando $\delta = 16$, as taxas de erro tipo I por experimento são idênticas a TEE do teste de Tukey. Isso é devido a semelhança entre as estruturas dos dois métodos. Assim, tanto faz usar um como o outro, sob H_0 parcial, quando $\delta \geq 8$. O problema é que essa diferença entre médias não é muito comum em situações práticas.

Para os testes de TM, de SNKM e de MGM, quando $\delta \geq 8$, observa-se que aumentando o número de tratamentos, esses testes também aumentam a TEE, tornando-se testes muito liberais, sob H_0 parcial. Quando $\delta < 8$, as TEEs desses testes em função do número de tratamentos (n), apresentam um comportamento de uma parábola (Figura

8), tendo um pico nos valores da TEE quando o número de tratamentos é igual a 20.

A avaliação de desempenho de outros testes encontrados na literatura não são comparáveis com os resultados encontrados no presente estudo, devido a forma de como os testes foram avaliados. Um dos poucos trabalhos que a avaliação de desempenho se assemelha a esse estudo foi encontrado em Carmer e Swanson (1973, p. 70). Apesar da metodologia ter sido um pouco diferente desse estudo, eles verificaram que os testes de Tukey e de Scheffé não ultrapassaram a TEE de 3,1%, em todas as configurações realizadas, considerando um nível de significância de $\alpha = 5\%$.

4.2 Poder

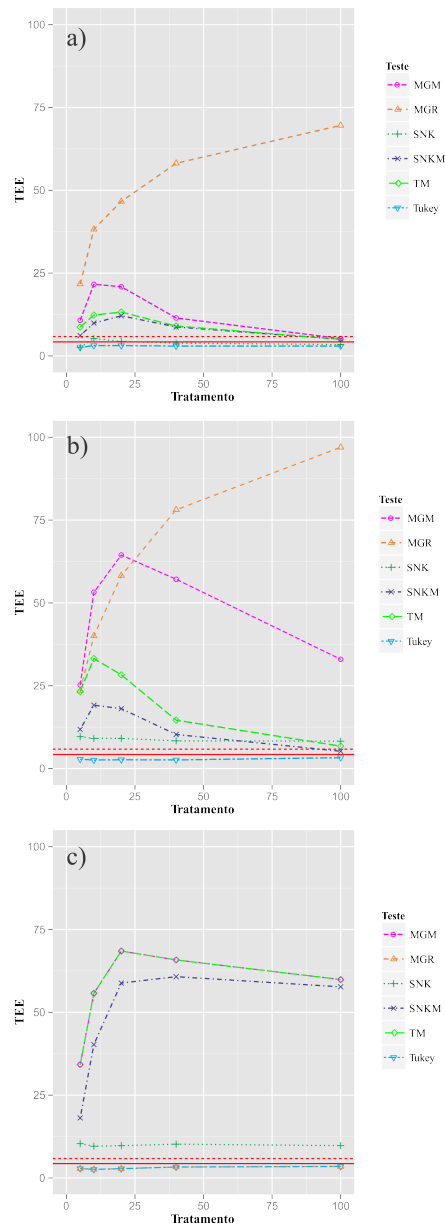
Nessa segunda etapa, comparou-se os testes por meio do estudo do poder (Definição 2.6). Foram consideradas diversas situações: número de repetições, número de tratamentos, diferenças entre médias, nível de significância, e número de populações. Nesta última situação foram simulados dois grupos que possuíam mesmas médias internamente e diferiam entre si por uma quantidade δ erros padrões, isto é, sob H_0 parcial (Definição 2.3). O estudo do poder também foi avaliado sob a hipótese H_1 , em que foi considerado comparações entre grupos de médias diferentes.

4.2.1 Poder sob H_1 completa

No Apêndice B é apresentado as Tabelas 20 a 43 sobre as percentagens de decisões corretas dos testes, sob H_1 completa, com relação ao número de tratamentos, número de repetições, diferença entre médias, o valor inicial da diferença entre médias e nível de significância. Os resultados serão discutidos de acordo com a simulação realizada nesse estudo e também com base em avaliações de desempenho de outros testes de trabalhos na literatura.

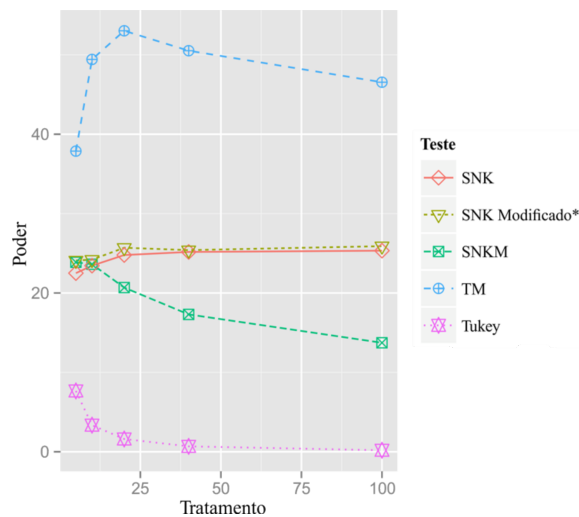
O poder dos testes foi influenciado pelo número de tratamentos. Na Figura 9, pode-se observar a avaliação de desempenho dos testes de Tukey, de Tukey *Midrange*, de SNK, de SNK *Midrange* e de SNK modificado, para diferença entre médias de $2\sigma_{\bar{X}}$, $r = 4$ repetições e um nível de significância de $\alpha = 5\%$ de probabilidade, Tabela 27. Este último teste foi avaliado por Percin e Malheiros (1989) no mesmo cenário de avaliação de desempenho do presente estudo.

Os testes de SNK, de SNK modificado e de TM aumentaram o poder com o



* As linhas de vermelho delimitam a região de rejeição pelo teste binomial exato.

Figura 8 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de MGM, de TM e de SNKM, em função do número de tratamentos, sob hipótese H_0 parcial, para a) $\delta = 2$, b) $\delta = 4$ e c) $\delta = 16$, com $r = 10$, $\alpha = 5\%$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.



* Resultado de Percin e Malheiros (1989).

Figura 9 Poder dos testes de SNK, de SNK modificado, de SNKM, de TM e de Tukey, em porcentagem, sob H_1 completa, para detectar uma diferença entre médias de $2\sigma_{\bar{Y}}$, com $r = 4$ repetições, em função do número de tratamentos, para um $\alpha = 0,05$.

aumento do número de tratamentos (n), sendo que o teste de TM obteve o maior poder entre esses dois testes.

O teste de TM foi o que apresentou maior poder, sendo uma diferença muito grande principalmente em relação ao teste de Tukey, que apresentou o pior desempenho. Outro fato interessante, foi que sob H_0 completa, o teste de TM preserva o erro tipo I, podendo ser um teste alternativo ao teste de Tukey. O valor inicial do poder para o teste de TM é de 37,87%, atingindo um valor máximo de 53,03% quando $n = 20$. Os testes de SNK e de SNK modificado apresentaram desempenho semelhante, sendo o poder desses testes inferior ao do teste de TM. Para n variando de 5 a 100, o poder dos testes de SNK e de SNK modificado foi entre 22,50% e 25,9%, respectivamente. Houve um pequeno aumento do poder com a variação do número de tratamentos.

Entretanto, os testes de Tukey e de SNKM diminuem o poder com o aumento do número de tratamentos. O teste de Tukey tem praticamente 0% de poder quando o número de tratamentos é igual a 100. Esse fato também foi observado por Percin e Malheiros (1989, p. 48). Esses autores mostraram que o teste de Tukey apresenta o menor poder dentre os testes avaliados. Com 100 médias, o poder de detectar diferença

de até 6 erros padrões da média foi inferior a 50%. Esse fato mostra que o teste de Tukey não seja recomendável para comparação de médias duas a duas com um grande número de tratamentos. O teste de SNKM também diminui o poder com o aumento do número de tratamentos. Contudo, o poder desse teste é bem superior ao do teste de Tukey.

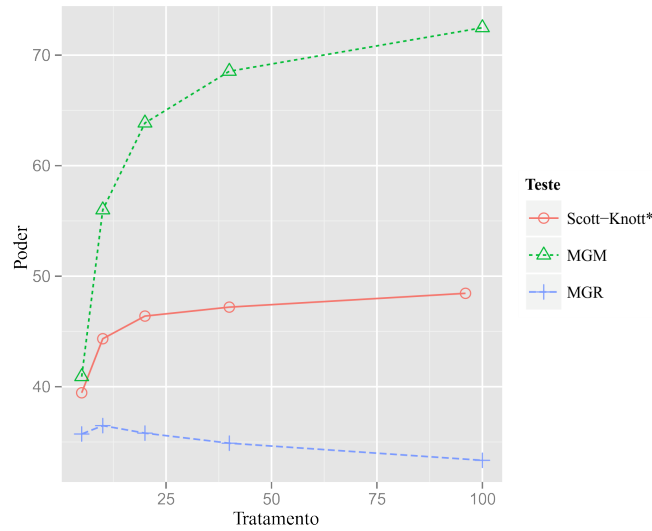
Outra avaliação de desempenho do poder foi realizada para os testes de MGR, de MGM e de Scott-Knott, Figura 10. Essa avaliação foi realizada com base no número de tratamentos, fixando a diferença entre médias de $2\sigma_{\bar{X}}$, $r = 4$ repetições e $\alpha = 5\%$ de probabilidade. O poder do teste de Scott-Knott foi retirado do trabalho de Silva, Ferreira e Bearzoti (1999). Esses autores avaliaram o teste de Scott-Knott no mesmo cenário experimental do presente trabalho.

O teste de MGM teve um aumento no poder com o aumento do número de tratamentos. Esse comportamento também foi verificado para o teste de Scott-Knott. Entretanto, o teste de MGR praticamente não alterou o poder com a variação do número de tratamentos. O que ocorreu foi uma pequena diminuição do poder com o aumento de n . Para $n = 5$, o poder dos testes de MGM, de Scott-Knott e de MGR foi 42,99%, 39,45% e 37,59%, respectivamente. Ao passo que para $n = 100$, o poder dos testes foi 72,64%, 48,45% e 33,51%, respectivamente. O teste de MGM apresentou o maior poder do que os testes de Scott-Knott e de MGR. Este último apresentou o pior desempenho entre os três testes.

As avaliações de desempenho dos testes foram realizadas separadamente com o objetivo de comparar os testes propostos com os teste originais. Para uma avaliação mais ampla, foram comparados todos os testes apresentados até o momento com as avaliações de desempenho de outros testes feitas por Perecin e Malheiros (1989, p. 48). Esses autores mostraram que o teste que apresentou maior poder foi o teste t-bayesiano, seguido do teste t, teste de Duncan, teste de Newman-Keuls modificado e o teste de Newman-Keuls. Todos apresentaram poder acima de 60% para $4\sigma_{\bar{X}}$ erros padrões, sendo esse efeito mais expressivo com o teste t-bayesiano, poder acima de 78%.

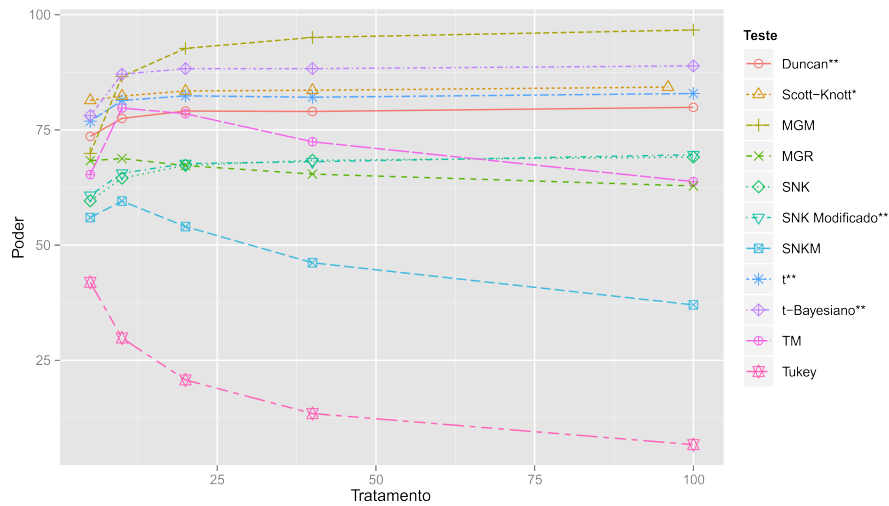
Na Figura 11 é apresentado o poder dos testes no cenário $\delta = 4$, $r = 4$ repetições e $\alpha = 5\%$ de probabilidade. Esse cenário serviu de base para apresentar as demais situações, pois os resultados foram equivalentes.

Observou-se que os testes de MGM, t-bayesiano e de Scott-Knott apresentaram o maior poder, sendo o primeiro teste o que apresentou maior destaque. O teste de Tukey apresentou o pior desempenho. Em relação aos testes propostos, o teste de SNKM foi o que obteve o menor poder. Contudo, o poder desse teste foi superior ao poder do teste



* Resultado de Silva, Ferreira e Bearzoti (1999).

Figura 10 Poder dos testes de Scott-Knott, de MGR e de MGM, em porcentagem, sob H_1 completa, para detectar uma diferença entre médias de $2\sigma_{\bar{Y}}$, com $r = 4$ repetições, em função do número de tratamentos, para um $\alpha = 0,05$.



* Resultado de Silva, Ferreira e Bearzoti (1999) e ** Resultados de Percin e Malheiros (1989).

Figura 11 Poder dos testes de Duncan, de Scott-Knott, de MGM, de MGR, de SNK, de SNK modificado, de SNKM, t, t-bayesiana, de TM e de Tukey, em porcentagem, sob H_1 completa, para detectar uma diferença entre médias de $4\sigma_{\bar{Y}}$, com $r = 4$ repetições, em função do número de tratamentos, para um $\alpha = 0,05$.

de Tukey. Os demais testes apresentaram os valores do poder intermediário entre os do teste de MGM (teste com maior poder) e do teste de Tukey (teste com menor poder). A escolha do melhor teste avaliando apenas o poder, sob H_0 completa, é: para $n \leq 5$ o teste de Scott-Knott é indicado, e para $n > 5$, o teste de MGM é o mais indicado.

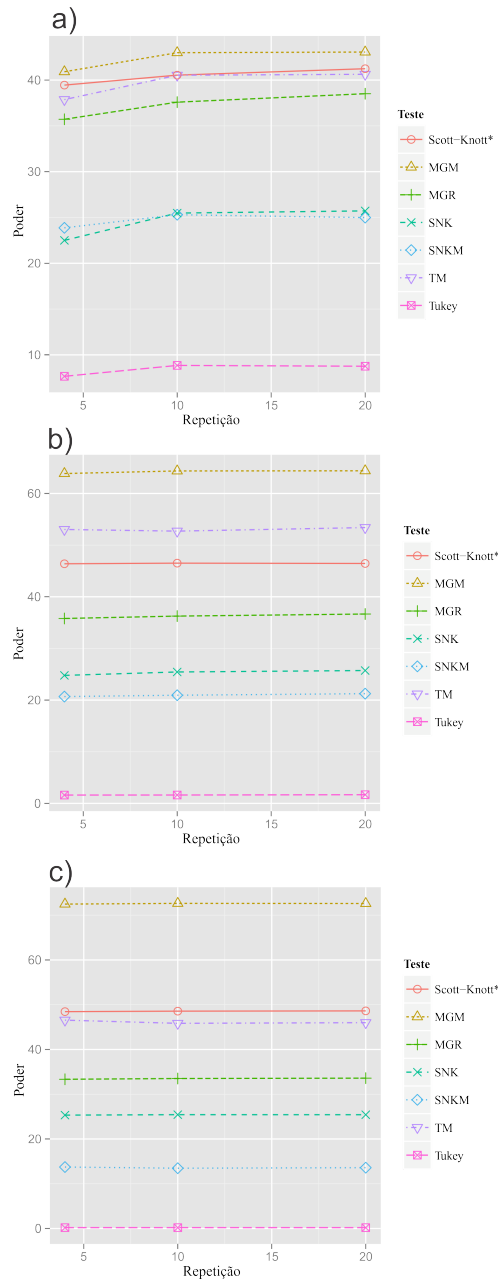
O número de repetições foi outro aspecto que influenciou o poder dos testes, porém não tão expressivo como no caso do número de tratamentos. Na Figura 12 é apresentado a avaliação de desempenho do poder em relação ao número de repetições. Essa avaliação foi analisada em três cenários: (a) 5 tratamentos, (b) 20 tratamentos e (c) 100 tratamentos, para uma diferença entre médias de $2\sigma_{\bar{X}}$ e $\alpha = 5\%$.

Para um pequeno número de tratamentos, Figura 12(a), houve um aumento no poder dos testes com o aumento do número de repetições, principalmente de 4 para 10 repetições, como verificado por Silva, Ferreira e Bearzoti (1999, p. 689) para o teste de Scott-Knott. Entretanto, quando o número de tratamentos aumentou ($n \geq 20$), Figuras 12(b) e 12(c), o poder dos testes praticamente não se alterou com o aumento do número de repetições. Isso pode ser devido a maior precisão da estimativa da variância residual, pois o aumento do número de tratamentos, independente do aumento do número de repetições, os graus de liberdade são elevados. Contudo, com um número de tratamentos pequeno, os graus de liberdade também são pequenos para um pequeno número de repetições, e elevados para um grande número de repetições. Dessa forma, observa-se o maior efeito do número de repetições para o poder nesta última situação, uma vez que a precisão do experimento foi fixada.

Portanto, para o estudo do poder considerando o número de repetições, o teste que obteve maior destaque foi o teste de MGM seguido dos testes de Scott-Knott e o teste de TM. O teste de Tukey obteve o pior desempenho. Considerando os testes propostos, o teste de SNKM teve o menor poder. Os testes de MGR e de SNK apresentaram o poder intermediário em relação ao poder dos testes analisados.

Outra avaliação de desempenho do poder foi com base na diferença entre médias. O poder dos testes aumentou rapidamente com o aumento da diferença entre médias. Silva, Ferreira e Bearzoti (1999, p. 689) e Perecin e Malheiros (1989, p. 49) mostraram que quando a magnitude entre médias foi igual ou superior a $6\sigma_{\bar{Y}}$, as porcentagens de decisões corretas dos testes avaliados foram altas.

Na Figura 13, foi apresentado o poder dos testes para as diferenças reais entre médias de 1 a $32\sigma_{\bar{Y}}$, com 4 repetições e $\alpha = 0,05$. O cenário foi dividido em relação ao número de tratamentos: (a) $n = 5$, (b) $n = 20$ e (c) $n = 100$. Quando $\delta \leq 6$, Fi-



* Resultado de Silva, Ferreira e Bearzoti (1999).

Figura 12 Poder dos testes de Scott-Knott, de MGM, de MGR, de SNK, de SNKM, de TM e de Tukey, em porcentagem, sob H_1 completa, para detectar uma diferença entre médias de $2\sigma_{\bar{Y}}$, com (a) $n = 5$, (b) $n = 20$ e (c) $n = 100$, em função do número de repetições, para um $\alpha = 0,05$.

gura 13(a), os testes de MGR (para $n = 5$) e de MGM (para $n = 20$ e 100) obtiveram poder superiores aos demais. Quando $\delta > 6$ quase todos os testes chegaram a 100% de poder (Figuras 13(b) e 13(c)), lembrando que os testes propostos convergiram mais lentamente para esse valor. Entretanto, os testes de TM e de SNKM tiveram pior desempenho para essa situação. Com 100 tratamentos, o poder desses testes não ultrapassaram 50%, mesmo quando a diferença real entre médias foi de $32\sigma_{\bar{Y}}$ (Figura 13(c)). É interessante observar o teste de SNK tende a ser ligeiramente superior ao teste de Tukey em todas as configurações, como também pode ser observado em Borges e Ferreira (2003, p. 78), e que esse teste convergiu mais rapidamente para 100%.

De modo que o poder dos testes propostos fosse comparado com o poder de outros testes encontrados na literatura, a diferença real entre médias foi considerada de 2 a $32\sigma_{\bar{Y}}$, para o número de tratamentos 5, 20 e 100, com 4 repetições e $\alpha = 0,05$, Figura 14. Esse cenário foi considerado, pois a avaliação de desempenho de diversos testes realizada por Percin e Malheiros (1989) e Silva, Ferreira e Bearzoti (1999), foi baseado nessa configuração, sendo base na comparação dos testes propostos desenvolvidos no presente estudo.

Para uma pequena diferença real entre médias, independente do tamanho de n , o teste de MGM obteve poder superior aos demais, sendo ainda mais acentuado à medida que n aumentava, principalmente do teste de Tukey, teste com menor poder para essa situação.

Os testes t-bayesiano e de Duncan, obtiveram destaque com relação ao poder, como já esperado, pois esses dois testes apresentam altas taxas de erro tipo I por experimento (BERNHARDSON, 1975; PERECIN; MALHEIROS, 1989), tratando-se de testes liberais. Sendo testes liberais, alta taxa erro tipo I por experimento implica em pequena taxa de erro tipo II, e por consequência alto poder. Com o aumento da diferença entre médias, esses dois testes convergiram mais rapidamente a 100%.

Ao contrário do teste de MGM, o teste de MGR apresentou o poder inferior ao do teste de Scott-Knott. Contudo, na avaliação de desempenho, foi classificado com o poder de teste intermediário com relação ao dos testes avaliados.

Os testes de SNK e de SNK modificado também apresentaram poder intermediário, porém inferiores ao teste de MGR para $n \leq 5$ e superiores ao teste de MGR para $n > 5$, Figura 14. O teste de TM apresentou o pior desempenho de avaliação do poder em relação aos testes propostos. Entretanto, comparando esse teste com base no teste original, o poder do teste de TM foi superior ao do teste de Tukey, exceto quando $\delta > 8\sigma_{\bar{Y}}$,

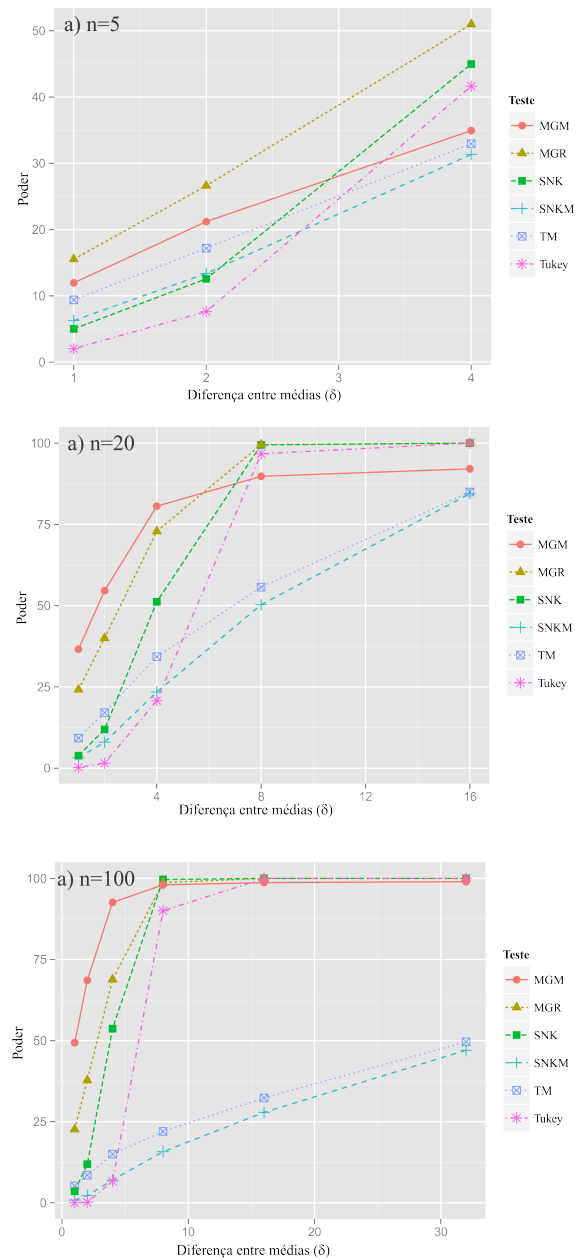


Figura 13 Poder dos testes de MGM, de MGR, de SNK, de SNKM, de TM e de Tukey, em porcentagem, sob H_1 completa, para detectar diferenças entre médias de 1 a $32 \sigma_{\bar{y}}$, considerando os tratamentos (a) $n = 5$, (b) $n = 20$ e (c) $n = 100$, 4 repetições, e um nível de significância de 5% de probabilidade.

tratando-se de um cenário não muito comum em situações reais.

Quando a diferença real entre médias aumentou, o poder dos testes t-bayesiano e de Duncan aumentou, apesar do teste de MGM ter mantido alto poder e controle do erro tipo I por experimento.

Portanto, como escolha de um bom teste, o teste de MGM ainda se mantém em destaque. Os testes de Scott-Knott e de SNK modificados obtiveram poder superior aos testes de MGR e de SNKM, respectivamente, com valores um pouco maiores com o aumento de n . O interessante é que em quase todas as configurações os testes de SNK e de SNK modificados foram praticamente iguais, exceto para a diferença entre médias entre 4 a $8\sigma_{\bar{y}}$.

Um aspecto muito relevante nos testes propostos (de MGR, de MGM, de TM e de SNKM) foi que apesar do poder ter apresentado uma convergência mais lenta para a porcentagem de decisões corretas máxima (100%), para pequenos valores de δ , esses testes superaram os testes originais em quase todas as configurações simuladas. Em experimentos reais, essa é a situação mais comum, Figura 13.

Na Figura 15, observa-se a configuração para a diferença real entre médias de 4 a $32\sigma_{\bar{y}}$, para $n = 5, 20$ e 100 tratamentos, com 4 repetições e $\alpha = 0,05$. Para esse cenário, a comparação dos testes propostos foi realizada com os testes de Tukey e de SNK. Independente do número tratamentos, os testes de MGM e de TM obtiveram poder superior, seguido dos testes de SNK, de SNKM, e de MGR. Mais uma vez, o teste com pior desempenho foi o de Tukey. Quando o valor inicial da diferença entre médias foi superior a $4\sigma_{\bar{y}}$, Tabelas 38 a 43, o poder dos testes convergiram rapidamente a 100%, já que a diferença real entre médias era muito grande.

No presente trabalho foi verificado que valores iniciais das diferenças reais entre médias, influenciaram o poder dos testes. Isso não foi verificado em nenhum outro trabalho. Considere o valor do poder do teste de MGM como exemplo. Na Tabela 21, o valor da diferença entre médias (δ) para o cenário $n = 5, r = 4$ e $\alpha = 0,05$, foi de 1 a 32. Na Tabela 27, o valor de δ foi de 2 a 32, e na Tabela 32, de 4 a 32. Observe que os valores de δ iniciais foram diferentes. Assim, para esses três cenários, considerando uma mesma diferença entre médias de $4\sigma_{\bar{y}}$, o poder para as três situações foi 34,94%, 69,91% e 89,58%, respectivamente, Figura 16. Isso mostra que o poder dos testes propostos e do teste de SNK, aumentou à medida que as as médias populacionais tornaram-se mais heterogêneas. Entretanto, isso acabou não acontecendo com o teste de MGR. Quando os valores iniciais de δ foram de 1 a 32 para 2 a 32 o poder desse teste aumentou ao avaliar

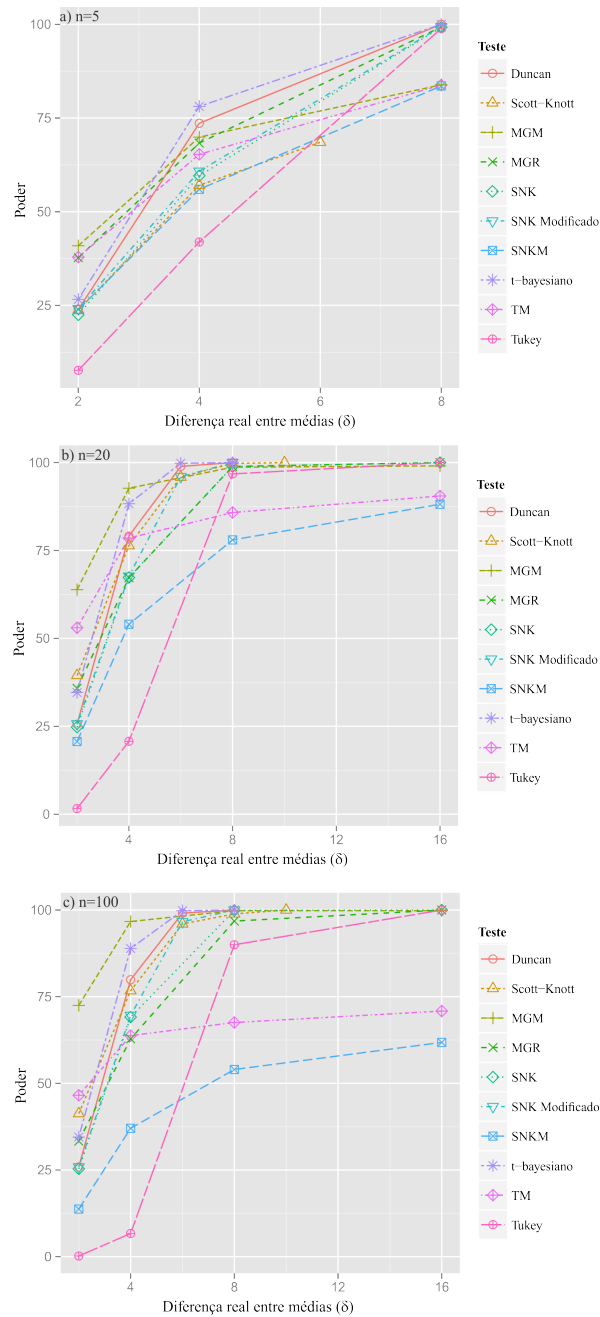


Figura 14 Poder dos testes de MGM, de MGR, de SNK, de SNKM, de TM, de Tukey, de Duncan, de SNK modificado e do t-bayesiano, em porcentagem, sob H_1 completa, para detectar diferenças entre médias de 2 a $32 \sigma_{\bar{Y}}$, considerando o número de tratamentos (a) $n = 5$, (b) $n = 20$ e (c) $n = 100$, 4 repetições, e um nível de significância de 5% de probabilidade.

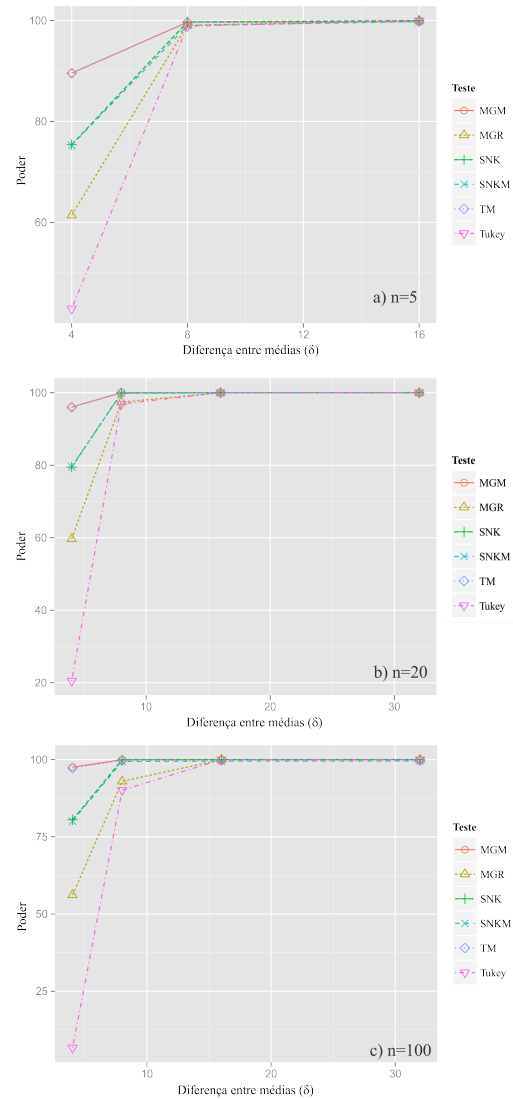


Figura 15 Poder dos testes de MGM, de MGR, de SNK, de SNKM, de TM, e de Tukey, em porcentagem, sob H_1 completa, para detectar diferenças entre médias de 4 a $32 \sigma_{\bar{Y}}$, considerando o número de tratamentos (a) $n = 5$, (b) $n = 20$ e (c) $n = 100$, 4 repetições, e um nível de significância de 5% de probabilidade.

uma mesma diferença entre médias ($4\sigma_{\bar{Y}}$). Entretanto, quando os valores iniciais de δ foram de 2 a 32 para 4 a 32 o poder desse teste decresceu. Assim, o que se observa é que o teste de MGR tende a ser mais poderoso quando avalia médias populações mais homogêneas do que em médias populações mais heterogêneas. Para o teste de Tukey, o poder tornou-se constante para uma mesma diferença entre médias populacionais cada vez mais heterogêneas. Isso pode ser explicado, pelo teste de Tukey ser muito conservativo. Um controle excessivo no erro tipo I, acaba influenciando no poder, como é preconizado na literatura.

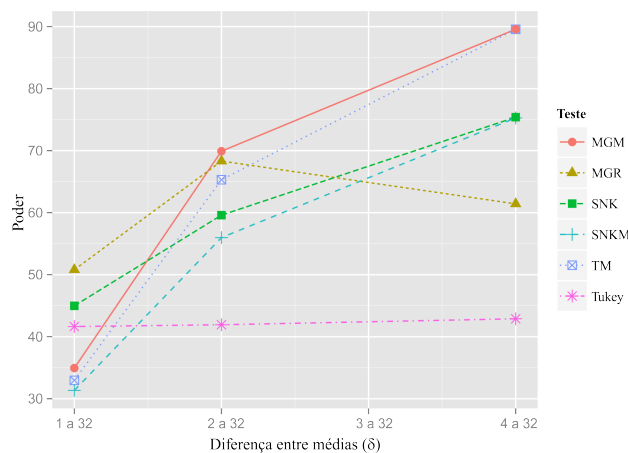


Figura 16 Poder dos testes de MGM, de MGR, de SNK, de SNKM, de TM e de Tukey, em porcentagem, sob H_1 completa, com relação aos valores iniciais das diferenças reais entre médias para $4\sigma_{\bar{Y}}$, para $n = 5$, $r = 4$ e $\alpha = 0,05$.

4.2.2 Poder sob H_0 parcial

No Apêndice C é apresentado as simulações sobre as porcentagens de decisões corretas dos testes dado que H_0 é falsa, sob H_0 parcial, Tabelas de 44 a 55. A avaliação levou em consideração o número de tratamentos, número de repetições, diferença entre médias e o nível de significância.

O número de tratamentos foi um ponto que influenciou o poder, sob H_0 parcial, apesar do número de repetições não ter apresentado tanta influência.

Na Figura 17, foi observado que o aumento do número de tratamentos (n), decresce o poder dos testes.

Entretanto, quando a diferença real entre médias foi $4\sigma_{\bar{Y}}$, Figura 17(c), o teste

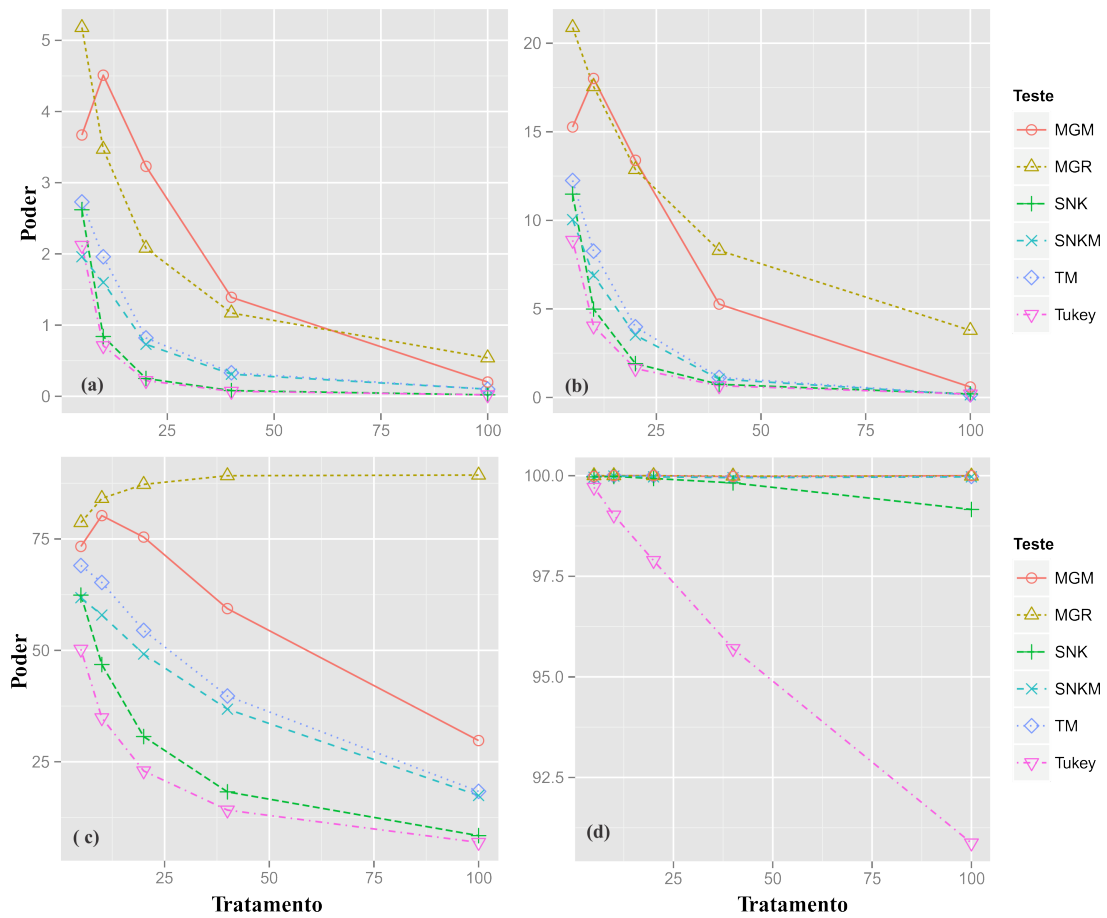


Figura 17 Poder dos testes de MGM, de MGR, de SNK, de SNKM, de TM, e de Tukey, em porcentagem, sob H_0 parcial, em função do número de tratamentos, para detectar diferenças entre médias de (a) $1\sigma_{\bar{Y}}$, (b) $2\sigma_{\bar{Y}}$, (c) $4\sigma_{\bar{Y}}$ e (d) $8\sigma_{\bar{Y}}$, com 4 repetições, e um nível de significância de 5% de probabilidade.

de MGR passou a aumentar o poder com o aumento de n , sendo o único teste a atingir poder em torno de 90% quando $n = 100$. Este teste e o teste de MGM obtiveram as maiores porcentagens de decisões corretas. Contudo, quando $\delta \leq 4$, os valores do poder não ultrapassaram 30%. Mesmo assim, o teste de Tukey obteve o pior desempenho em todas as situações. Com o aumento de n , seu poder chegou próximo de 0%. A partir de $\delta > 8$, quase todos os testes convergiram para o poder 100%.

Todos os testes propostos, sob H_0 parcial, obtiveram poder superior aos testes originais. Entretanto, o poder dos testes propostos, assim como o teste de SNK, passou a ter pouco significado prático, já que o erro tipo I por experimento de todos esses testes foi superior ao nível de significância adotado, sob H_0 parcial. Apenas os testes de Tukey e de Scheffé tiveram TEE idênticas ao nível nominal, como verificado em Carmer e Swanson (1973, p. 70). Entretanto, o poder desses chegou em determinadas situações a 0%.

Uma característica que pode ser melhorada nos testes propostos, para o controle do erro tipo I por experimento e alto poder, sob H_0 parcial, é tentar melhorar a contribuição que a média populacional desconhecida, influencia na DMS dos testes, já que a distribuição da midrange centrada em μ depende do parâmetro de locação.

4.3 Aplicações dos testes propostos

Após a discussão dos resultados da avaliação de desempenho dos testes propostos, será apresentado um exemplo para a aplicação do pacote `midrangeMCP`.

Antes de apresentar as rotinas do pacote, será mostrado os passos de cada teste de forma analítica, usando o Exemplo 3.1.

4.3.1 Teste de agrupamento de médias *Midrange* (MGM) - Solução analítica

Seguem os passos:

1. Inicialmente, ordena-se as médias dos tratamentos dados por:

3DOK13	3DOK4	COMPOS	3DOK7	3DOK5	3DOK1
13,26	14,64	18,70	19,92	23,98	28,82

2. Após isso, calcula-se as diferenças entre médias consecutivas e verifica-se a que obteve maior diferença:

$$\bar{Y}_{3DOK4} - \bar{Y}_{3DOK13} = 1,38$$

$$\bar{Y}_{COMPOS} - \bar{Y}_{3DOK4} = 4,06$$

$$\bar{Y}_{3DOK7} - \bar{Y}_{COMPOS} = 1,22$$

$$\bar{Y}_{3DOK5} - \bar{Y}_{3DOK7} = 4,06$$

$$\bar{Y}_{3DOK1} - \bar{Y}_{3DOK5} = 4,84$$

3. O passo anterior auxiliou a identificar o ponto de corte de dois potenciais grupos de médias de populações com médias populacionais diferentes, isto é:

$$\underbrace{\boxed{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4} \quad \bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7} \quad \bar{Y}_{3DOK5}}}_{\text{Grupo 1}} \boxplus \underbrace{\boxed{\bar{Y}_{3DOK1}}}_{\text{Grupo 2}}$$

Além do mais, o grupo com maior número de médias servirá para estimar a média populacional desconhecida, ou seja,

$$\bar{Y}^* = \frac{13,26 + 14,64 + 18,70 + 19,92 + 23,98}{5} = 18,10;$$

4. Calculando a estatística \bar{r}_6 , expressão (74):

$$\bar{r}_6 = \frac{13,26 + 28,82}{2} - 18,10 = 2,94;$$

5. Calculando Δ_6 expressa em (76), tem-se

$$\begin{aligned} \Delta_6 &= \bar{q}_{(0,05/2;6,24)} \sqrt{\frac{11,78867}{5}} + \frac{1}{\sqrt{6}} \sqrt{\frac{11,78867}{5}} \\ &= 1,004855 \sqrt{\frac{11,78867}{5}} + \frac{1}{\sqrt{6}} \sqrt{\frac{11,78867}{5}} \\ &= 2,88; \end{aligned}$$

6. Como $|\bar{r}_6| > |\Delta_6|$, então os dois grupos foram amostrados de populações com médias estatisticamente diferentes;
7. Assim, parte para a análise de cada grupo, lembrando que no grupo de maior média, como há apenas uma média, esta recebe a codificação de diferenciação dos grupos, isto é,

Tratamento	Média	Resultado
3DOK1	28,82	g_1
3DOK5	23,98	<i>análise</i>
3DOK7	19,92	<i>análise</i>
COMPOS	18,70	<i>análise</i>
3DOK4	14,64	<i>análise</i>
3DOK13	13,26	<i>análise</i>

8. Repetindo os passos de 1 a 4, porém, usando apenas os elementos do Grupo 1, segue que

3DOK13	3DOK4	COMPOS	3DOK7	3DOK5
13,26	14,64	18,70	19,92	23,98

$$\bar{Y}_{3DOK4} - \bar{Y}_{3DOK13} = 1,38$$

$$\bar{Y}_{COMPOS} - \bar{Y}_{3DOK4} = 4,06$$

$$\bar{Y}_{3DOK7} - \bar{Y}_{COMPOS} = 1,22$$

$$\bar{Y}_{3DOK5} - \bar{Y}_{3DOK7} = 4,06$$

Nesse caso houve um empate para k . Assim, poderá haver duas possibilidades para a separação de grupos:

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7} \quad \bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

ou

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4} \quad \bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

A escolha de k será determinada por:

$$\begin{aligned} k &= \max\{\min(k_1, m - k_1), \min(k_2, m - k_2)\} \\ &= \max\{\min(2, 3), \min(4, 1)\} \\ &= 2. \end{aligned}$$

Portanto, para $k = 2$ tem-se

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7} \quad \bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

$$\bar{r}_5 = \frac{13,26 + 23,98}{2} - 20,87 = -2,25;$$

9. Depois da quebra mais externa a Δ_6^* tem uma pequena alteração, em que é calculado da seguinte forma:

$$\begin{aligned} \Delta_6^* &= \bar{q}_{(0,05/2;6,24)} \sqrt{\frac{11,78867}{5}} \\ &= 1,004855 \sqrt{\frac{11,78867}{5}} \\ &= 1,542945. \end{aligned}$$

De agora em diante, todos os cálculos de \bar{r} serão comparados com Δ_6^* ;

10. Assim, como $|\bar{r}_5| > |\Delta_6^*|$ conclui-se que os grupos são estatisticamente diferentes. Por enquanto nenhuma letra foi considerada nessa segunda quebra, pois cada subgrupo ainda será avaliado;
11. A avaliação de cada subgrupo será reutilizando os passos 1, 2, 3, 4 e 9. Esse procedimento irá parar até que nenhum outro grupo seja dividido em dois novos grupos ou até quando os grupos contiverem apenas 1 média. Dessa forma, o resultado da análise fica:

Tratamento	Média	Resultado
3DOK1	28,82	g_1
3DOK5	23,98	g_2
3DOK7	19,92	g_3
COMPOS	18,70	g_3
3DOK4	14,64	g_4
3DOK13	13,26	g_4

4.3.2 Teste de agrupamento de médias *Range* (MGR) - Solução analítica

Comparando com a análise do teste de MGM, os passos de 1 a 3 são os mesmos. Para os passos seguintes seguem:

4. Calcular a estatística q_6 , expressa em (79):

$$q_6 = 28,82 - 13,26 = 15,56;$$

5. Calculando a diferença mínima significativa, expressão (80):

$$\Delta_6 = 4.372651 \sqrt{\frac{11,78867}{5}} = 6.714167;$$

6. Como $q_6 > \Delta_6$, então os dois grupos foram amostrados de populações com médias estatística diferentes;
7. Assim, parte para a análise de cada grupo, lembrando que no grupo de maior média, como há apenas uma média, esta recebe a primeira letra de diferenciação dos grupos, isto é,

Tratamento	Média	Resultado
3DOK1	28,82	g_1
3DOK5	23,98	<i>análise</i>
3DOK7	19,92	<i>análise</i>
COMPOS	18,70	<i>análise</i>
3DOK4	14,64	<i>análise</i>
3DOK13	13,26	<i>análise</i>

8. Repetindo os passos de 1 a 3, porém, usando apenas os elementos do grupo 1, tem-se

3DOK13	3DOK4	COMPOS	3DOK7	3DOK5
13,26	14,64	18,70	19,92	23,98

$$\bar{Y}_{3DOK4} - \bar{Y}_{3DOK13} = 1,38$$

$$\bar{Y}_{COMPOS} - \bar{Y}_{3DOK4} = 4,06$$

$$\bar{Y}_{3DOK7} - \bar{Y}_{COMPOS} = 1,22$$

$$\bar{Y}_{3DOK5} - \bar{Y}_{3DOK7} = 4,06$$

Nesse caso houve um empate para k . Assim, poderá haver duas possibilidades para a separação de grupos:

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7} \quad \bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

ou

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4} \bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

A escolha de k será determinada por:

$$\begin{aligned} k &= \max\{\min(k_1, m - k_1), \min(k_2, m - k_2)\} \\ &= \max\{\min(2,3), \min(4, 1)\} \\ &= 2. \end{aligned}$$

Portanto, para $k = 2$ tem-se

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7} \quad \bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

$$q_5 = 23,98 - 13,26 = 10,72;$$

9. Como $q_5 > \Delta_6$ conclui-se que os grupos são estatisticamente diferentes. Por enquanto nenhuma letra foi considerada, pois cada subgrupo será avaliado;
10. A avaliação de cada subgrupo será refazer todos os passos novamente em cada subgrupo. Esse procedimento irá parar até que nenhum outro grupo seja dividido em dois novos grupos ou até quando os grupos contiverem apenas 1 média. Dessa forma, o resultado da análise fica:

Tratamento	Média	Resultado
3DOK1	28,82	g_1
3DOK5	23,98	g_2
3DOK7	19,92	g_2
COMPOS	18,70	g_2
3DOK4	14,64	g_3
3DOK13	13,26	g_3

4.3.3 Teste de Student-Newman-Keuls *Midrange* (SNKM) - Solução analítica

Os passos para realizar o teste de SNKM é bem similar aos anteriores. O que irá diferenciar nos passos iniciais é que o maior “gap” servirá apenas para estimar a média

populacional, isto é, calcular \bar{Y}^* . Algo interessante também é que a DMS desse teste é alterada à medida que o número de médias se modifica nos grupos avaliados. Assim, seguem os passos:

1. Inicialmente, ordena-se as médias dos tratamentos dados por:

3DOK13	3DOK4	COMPOS	3DOK7	3DOK5	3DOK1
13,26	14,64	18,70	19,92	23,98	28,82

2. Após isso, calcula-se as diferenças entre médias consecutivas e verifica-se a que obteve maior diferença:

$$\bar{Y}_{3DOK4} - \bar{Y}_{3DOK13} = 1,38$$

$$\bar{Y}_{COMPOS} - \bar{Y}_{3DOK4} = 4,06$$

$$\bar{Y}_{3DOK7} - \bar{Y}_{COMPOS} = 1,22$$

$$\bar{Y}_{3DOK5} - \bar{Y}_{3DOK7} = 4,06$$

$$\bar{Y}_{3DOK1} - \bar{Y}_{3DOK5} = 4,84$$

3. O passo anterior auxiliou a identificar o ponto de corte de dois grupos de médias, isto é:

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4} \quad \bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7} \quad \bar{Y}_{3DOK5}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{3DOK1}}_{\text{Grupo 2}}$$

O grupo com maior número de médias servirá para estimar a média populacional desconhecida, ou seja,

$$\bar{Y}^* = \frac{13,26 + 14,64 + 18,70 + 19,92 + 23,98}{5} = 18,10;$$

4. Calculando a estatística \bar{r}_6 , expressão (74):

$$\bar{r}_6 = \frac{13,26 + 28,82}{2} - 18,10 = 2,94;$$

5. Calculando Δ_6 expressa em (76), tem-se

$$\begin{aligned} \Delta_6 &= \bar{q}_{(0,05/2;6,24)} \sqrt{\frac{11,78867}{5}} + \frac{1}{\sqrt{6}} \sqrt{\frac{11,78867}{5}} \\ &= 1.0049 \sqrt{\frac{11,78867}{5}} + \frac{1}{\sqrt{6}} \sqrt{\frac{11,78867}{5}} \end{aligned}$$

$$=1.9862;$$

6. Como $|\bar{r}_6| > |\Delta_6|$, então os dois grupos foram amostrados de populações com médias estatisticamente diferentes. O próximo passo é saber quais as médias são estatisticamente diferentes;
7. Deve-se fazer os mesmos passos anteriores considerando o primeiro grupo de 5 médias. Assim, segue que:
- 7.1. Inicialmente, ordena-se as cinco primeiras médias dos tratamentos dados por:

3DOK13	3DOK4	COMPOS	3DOK7	3DOK5
13,26	14,64	18,70	19,92	23,98

- 7.2. Após isso, calcula-se as diferenças entre médias consecutivas e verifica-se a que obteve maior diferença:

$$\bar{Y}_{3DOK4} - \bar{Y}_{3DOK13} = 1,38$$

$$\bar{Y}_{COMPOS} - \bar{Y}_{3DOK4} = 4,06$$

$$\bar{Y}_{3DOK7} - \bar{Y}_{COMPOS} = 1,22$$

$$\bar{Y}_{3DOK5} - \bar{Y}_{3DOK7} = 4,06$$

Nesse caso houve um empate para k . Assim, poderá haver duas possibilidades para a separação de grupos:

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7} \quad \bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

ou

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4} \quad \bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

A escolha de k será determinada por:

$$\begin{aligned} k &= \max\{\min(k_1, m - k_1), \min(k_2, m - k_2)\} \\ &= \max\{\min(2, 3), \min(4, 1)\} \\ &= 2. \end{aligned}$$

Portanto, para $k = 2$ tem-se

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7} \quad \bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

$$\bar{r}_5 = \frac{13,26 + 23,98}{2} - 20,87 = -2,25;$$

7.3. Calculando \bar{Y}^* :

$$\bar{Y}^* = \frac{18,70 + 19,92 + 23,98}{3} = 20,87;$$

7.4. Calculando Δ_5^* , expressão (78),

$$\begin{aligned} \Delta_5^* &= \bar{q}_{(0,05/2;5,24)} \sqrt{\frac{11,78867}{5}} \\ &= 1,0555 \sqrt{\frac{11,78867}{5}} \\ &= 2,0640. \end{aligned}$$

7.5. Calculando \bar{r}_5

$$\bar{r}_5 = \frac{13,26 + 23,98}{2} - 20,87 = -2,25;$$

7.6. $|\bar{r}_5| > |\Delta_5^*|$, conclui-se que as médias são estatisticamente diferentes. Por enquanto nenhuma letra foi considerada, pois cada subgrupo será avaliado;

8. Fazendo os mesmos passos anteriores considerando o segundo grupo de 5 médias, segue que:

8.1. Inicialmente, ordena-se as cinco primeiras médias dos tratamentos dados por:

3DOK4	COMPOS	3DOK7	3DOK5	3DOK1
14,64	18,70	19,92	23,98	28,82

8.2. Após isso, calcula-se as diferenças entre médias consecutivas e verifica-se a que obteve maior diferença:

$$\bar{Y}_{COMPOS} - \bar{Y}_{3DOK4} = 4,06$$

$$\bar{Y}_{3DOK7} - \bar{Y}_{COMPOS} = 1,22$$

$$\bar{Y}_{3DOK5} - \bar{Y}_{3DOK7} = 4,06$$

$$\bar{Y}_{3DOK1} - \bar{Y}_{3DOK5} = 4,84$$

Para o cálculo de \bar{Y}^* os grupos se subdividiram em :

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4} \quad \bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

8.3. Calculando \bar{Y}^* :

$$\bar{Y}^* = \frac{14,64 + 18,70 + 19,92 + 23,98}{4} = 19,31;$$

8.4. Calculando Δ_5^* , expressão (78),

$$\begin{aligned} \Delta_5^* &= \bar{q}_{(0,05/2;5,24)} \sqrt{\frac{11,78867}{5}} \\ &= 1,0555 \sqrt{\frac{11,78867}{5}} \\ &= 2,0640. \end{aligned}$$

8.5. Calculando \bar{r}_5

$$\bar{r}_5 = \frac{14,64 + 28,82}{2} - 19,31 = 2,42;$$

8.6. $|\bar{r}_5| > |\Delta_5^*|$, conclui-se que as médias são estatisticamente diferentes. Por enquanto nenhuma letra foi considerada, pois cada subgrupo será avaliado;

9. Até agora nada pode-se afirmar sobre as médias. Assim, o procedimento segue avaliando o grupo de 4 médias, depois o grupo com 3 médias até duas a duas. O processo se encerra caso não haja diferença entre médias ou até a comparação duas a duas;
10. Dessa forma, o resultado da análise ficou da seguinte forma:

Tratamento	Média	Resultado
3DOK1	28,82	g_1
3DOK5	23,98	g_2
3DOK7	19,92	g_2
COMPOS	18,70	g_2 g_3
3DOK4	14,64	g_2 g_3
3DOK13	13,26	g_3

4.3.4 Teste de Tukey *Midrange* (TM) - Solução analítica

O procedimento para realizar o teste de Tukey *Midrange* é similar ao do teste de SNKM. O que diferencia é que a DMS é fixa independente do número de médias envolvido. Seguem os passos para o teste de TM:

1. Inicialmente, ordena-se as médias dos tratamentos dados por:

3DOK13	3DOK4	COMPOS	3DOK7	3DOK5	3DOK1
13,26	14,64	18,70	19,92	23,98	28,82

2. Após isso, calcula-se as diferenças entre médias consecutivas e verifica-se a que obteve maior diferença:

$$\bar{Y}_{3DOK4} - \bar{Y}_{3DOK13} = 1,38$$

$$\bar{Y}_{COMPOS} - \bar{Y}_{3DOK4} = 4,06$$

$$\bar{Y}_{3DOK7} - \bar{Y}_{COMPOS} = 1,22$$

$$\bar{Y}_{3DOK5} - \bar{Y}_{3DOK7} = 4,06$$

$$\bar{Y}_{3DOK1} - \bar{Y}_{3DOK5} = 4,84$$

3. O passo anterior auxiliou a identificar o ponto de corte de dois grupos de médias, isto é:

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4} \quad \bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7} \quad \bar{Y}_{3DOK5}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{3DOK1}}_{\text{Grupo 2}}$$

O grupo com maior número de médias servirá para estimar a média populacional desconhecida, ou seja,

$$\bar{Y}^* = \frac{13,26 + 14,64 + 18,70 + 19,92 + 23,98}{5} = 18,10;$$

4. Calculando a estatística \bar{r}_6 , expressão (74):

$$\bar{r}_6 = \frac{13,26 + 28,82}{2} - 18,10 = 2,94;$$

5. Calculando Δ_6 expressa em (76), tem-se

$$\begin{aligned} \Delta_6 &= \bar{q}_{(0,05/2;6,28)} \sqrt{\frac{11,78867}{5}} + \frac{1}{\sqrt{6}} \sqrt{\frac{11,78867}{5}} \\ &= 1,004855 \sqrt{\frac{11,78867}{5}} + \frac{1}{\sqrt{6}} \sqrt{\frac{11,78867}{5}} \\ &= 1,986204; \end{aligned}$$

6. Como $|\bar{r}_6| > |\Delta_6|$, então os dois grupos foram amostrados de populações com médias estatisticamente diferentes. O próximo passo é saber quais as médias são estatisticamente diferentes;
7. Deve-se fazer os mesmos passos anteriores considerando o primeiro grupo de 5 médias. Assim, segue que:

- 7.1. Inicialmente, ordena-se as cinco primeiras médias dos tratamentos dados por:

3DOK13	3DOK4	COMPOS	3DOK7	3DOK5
13,26	14,64	18,70	19,92	23,98

- 7.2. Após isso, calcula-se as diferenças entre médias consecutivas e verifica-se a que obteve maior diferença:

$$\bar{Y}_{3DOK4} - \bar{Y}_{3DOK13} = 1,38$$

$$\bar{Y}_{COMPOS} - \bar{Y}_{3DOK4} = 4,06$$

$$\bar{Y}_{3DOK7} - \bar{Y}_{COMPOS} = 1,22$$

$$\bar{Y}_{3DOK5} - \bar{Y}_{3DOK7} = 4,06$$

Nesse caso houve um empate para k . Assim, poderá haver duas possibilidades para a separação de grupos:

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7} \quad \bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

ou

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4} \bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

A escolha de k será determinada por:

$$\begin{aligned} k &= \max\{\min(k_1, m - k_1), \min(k_2, m - k_2)\} \\ &= \max\{\min(2,3), \min(4, 1)\} \\ &= 2. \end{aligned}$$

Portanto, para $k = 2$ tem-se

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7} \quad \bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

$$\bar{r}_5 = \frac{13,26 + 23,98}{2} - 20,87 = -2,25;$$

7.3. Calculando \bar{Y}^* :

$$\bar{Y}^* = \frac{18,70 + 19,92 + 23,98}{3} = 20,87;$$

7.4. Calculando Δ_5^* , expressão (77),

$$\begin{aligned} \Delta_5^* &= \bar{q}_{(0,05/2;6,28)} \sqrt{\frac{11,78867}{5}} \\ &= 1,004855 \sqrt{\frac{11,78867}{5}} \\ &= 1,542945. \end{aligned}$$

7.5. Calculando \bar{r}_5

$$\bar{r}_5 = \frac{13,26 + 23,98}{2} - 20,87 = -2,25;$$

7.6. $|\bar{r}_5| > |\Delta_5^*|$, conclui-se que as médias são estatisticamente diferentes.

8. Deve-se fazer os mesmos passos anteriores considerando o segundo grupo de 5

médias. Assim, segue que:

8.1. Inicialmente, ordena-se as cinco primeiras médias dos tratamentos dados por:

3DOK4	COMPOS	3DOK7	3DOK5	3DOK1
14,64	18,70	19,92	23,98	28,82

8.2. Após isso, calcula-se as diferenças entre médias consecutivas e verifica-se a que obteve maior diferença:

$$\bar{Y}_{COMPOS} - \bar{Y}_{3DOK4} = 4,06$$

$$\bar{Y}_{3DOK7} - \bar{Y}_{COMPOS} = 1,22$$

$$\bar{Y}_{3DOK5} - \bar{Y}_{3DOK7} = 4,06$$

$$\bar{Y}_{3DOK1} - \bar{Y}_{3DOK5} = 4,84$$

Para o cálculo de \bar{Y}^* os grupos se subdividiram em :

$$\underbrace{\bar{Y}_{3DOK13} \quad \bar{Y}_{3DOK4} \quad \bar{Y}_{COMPOS} \quad \bar{Y}_{3DOK7}}_{\text{Grupo 1}} \boxplus \underbrace{\bar{Y}_{3DOK5}}_{\text{Grupo 2}}$$

8.3. Calculando \bar{Y}^* :

$$\bar{Y}^* = \frac{14,64 + 18,70 + 19,92 + 23,98}{4} = 19,31;$$

8.4. Calculando Δ_5^* , expressão (77),

$$\begin{aligned} \Delta_5^* &= \bar{q}_{(0,05/2;6,28)} \sqrt{\frac{11,78867}{5}} \\ &= 1,004855 \sqrt{\frac{11,78867}{5}} \\ &= 1,542945. \end{aligned}$$

8.5. Calculando \bar{r}_5

$$\bar{r}_5 = \frac{14,64 + 28,82}{2} - 19,31 = 2,42;$$

8.6. $|\bar{r}_5| > |\Delta_5|$, conclui-se que as médias são estatisticamente diferentes;

9. Até agora nada pode-se afirmar sobre as médias. Assim, o procedimento segue avaliando o grupo de 4 médias, depois o grupo com 3 médias até duas a duas. O

processo se encerra caso não haja diferença entre médias ou até a comparação duas a duas;

10. Dessa forma, o resultado da análise ficou da seguinte forma:

Tratamento	Média	Resultado
3DOK1	28,82	g_1
3DOK5	23,98	g_2
3DOK7	19,92	g_3
COMPOS	18,70	g_3
3DOK4	14,64	g_4
3DOK13	13,26	g_4

4.3.5 Usando o pacote midrangeMCP

A primeira função do pacote `midrangeMCP` a ser apresentada é `MRTTest`. Inicialmente, criou-se um arquivo que armazenou o banco de dados do Exemplo 3.1, intitulado `datast1980.txt`. No R:

```
dados <- read.table("datast1980.txt", h =T)
dados$trt <- factor(dados$trt)
attach(dados) # Quebrando o objeto dados
```

```
> dados
  trt  y
1   1 19.4
2   1 32.6
3   1 27.0
4   1 32.1
5   1 33.0
6   2 17.7
7   2 24.8
8   2 27.9
9   2 25.2
10  2 24.3
11  3 17.0
12  3 19.4
13  3  9.1
```

14	3	11.9
15	3	15.8
16	4	20.7
17	4	21.0
18	4	20.5
19	4	18.8
20	4	18.6
21	5	14.3
22	5	14.4
23	5	11.8
24	5	11.6
25	5	14.2
26	6	17.3
27	6	19.4
28	6	19.1
29	6	16.9
30	6	20.8

A codificação dos tratamentos no banco de dados foi da seguinte forma:

Tratamento	Codificação
3DOK1	1
3DOK5	2
3DOK4	3
3DOK7	4
3DOK13	5
COMPOS	6

Realizando a análise de variância para obter o quadrado médio e o grau de liberdade do resíduo, tem-se:

```
anava <- aov(y~trt)
```

Como visto na subseção 3.3.2, os argumentos `y` e `trt` da função `MRtest` têm três tipos de entrada:

- `y` como objeto da função `aov`. Assim, para o argumento `trt` basta informar o nome da coluna que armazena os níveis do tratamento:


```
MRtest(y = anava, trt = "trt")
```

MCP's based on distributions of the studentized midrange and range

```
Study: aov(y ~ trt) ~ "trt"
```

Summary:

	Means	std	r	Min	Max
1	28.82	5.80	5	19.4	33.0
2	23.98	3.78	5	17.7	27.9
3	14.64	4.12	5	9.1	19.4
4	19.92	1.13	5	18.6	21.0
5	13.26	1.43	5	11.6	14.4
6	18.70	1.60	5	16.9	20.8

Scott-Knott Midrange Test

Statistics:

Exp.Mean	CV	MSerror	Df	n	Stud.Mid	Ext.DMS	Int.DMS
19.89	17.27	11.78867	24	6	1.004855	1.986204	1.542945

Groups:

	Means	Groups
1	28.82	g1
2	23.98	g2
4	19.92	g3
6	18.70	g3
3	14.64	g4
5	13.26	g4

Mean Grouping Range Test

Statistics:

Exp.Mean	CV	MSerror	Df	n	Stud.Range	DMS
19.88667	17.26515	11.78867	24	6	4.372651	6.714167

Groups:

	Means	Groups
1	28.82	g1
2	23.98	g2
4	19.92	g2
6	18.70	g2
3	14.64	g3
5	13.26	g3

SNK Midrange Test

Statistics:

	Exp.Mean	CV	MSerror	Df	n	Stud.Mid	DMS
comp1	19.8867	17.2651	11.7887	24	6	1.0049	1.9862
comp2	19.8867	17.2651	11.7887	24	5	1.0555	2.0640
comp3	19.8867	17.2651	11.7887	24	4	1.1278	2.1750
comp4	19.8867	17.2651	11.7887	24	3	1.2422	2.3506
comp5	19.8867	17.2651	11.7887	24	2	1.4594	2.6841

Groups:

	Means	Groups
1	28.82	g1
2	23.98	g1g2
4	19.92	g2
6	18.70	g2g3
3	14.64	g2g3
5	13.26	g3

Tukey Midrange Test

Statistics:

```
Exp.Mean   CV  MSerror  Df n  Stud.Mid  Ext.DMS  Int.DMS
19.89      17.27 11.78867 24 6  1.004855  1.986204  1.542945
```

Groups:

	Means	Groups
1	28.82	g1
2	23.98	g2
4	19.92	g3
6	18.70	g3
3	14.64	g4
5	13.26	g4

- `y` como vetor da variável resposta e o argumento `trt` como vetor dos tratamentos. Assim, deve-se informar na função `MRtest` o valor do quadrado médio e o grau de liberdade do resíduo nos argumentos `mserror` e `dferror`, respectivamente.

Dessa forma, tem-se:

```
anava <- aov(y~trt)
glerror <- df.residual(anava)
qmerror <- deviance(anava) / glerror

# Aplicando os testes
MRtest(y = y,
       trt = trt,
       dferror = glerror,
       mserror = qmerror)
```

MCP's based on distributions of the studentized midrange and range

Study: y ~ trt

Summary:

	Means	std	r	Min	Max
1	28.82	5.80	5	19.4	33.0

```

2 23.98 3.78 5 17.7 27.9
3 14.64 4.12 5 9.1 19.4
4 19.92 1.13 5 18.6 21.0
5 13.26 1.43 5 11.6 14.4
6 18.70 1.60 5 16.9 20.8

```

Scott-Knott Midrange Test

Statistics:

```

Exp.Mean  CV  MSerror Df n Stud.Mid  Ext.DMS  Int.DMS
19.89     17.27 11.78867 24 6 1.004855 1.986204 1.542945

```

Groups:

```

Means  Groups
1 28.82  g1
2 23.98  g2
4 19.92  g3
6 18.70  g3
3 14.64  g4
5 13.26  g4

```

Mean Grouping Range Test

Statistics:

```

Exp.Mean      CV  MSerror Df n Stud.Range      DMS
19.88667 17.26515 11.78867 24 6 4.372651 6.714167

```

Groups:

```

Means  Groups
1 28.82  g1
2 23.98  g2
4 19.92  g2
6 18.70  g2
3 14.64  g3

```

5 13.26 g3

SNK Midrange Test

Statistics:

	Exp.Mean	CV	MSerror	Df	n	Stud.Mid	DMS
comp1	19.8867	17.2651	11.7887	24	6	1.0049	1.9862
comp2	19.8867	17.2651	11.7887	24	5	1.0555	2.0640
comp3	19.8867	17.2651	11.7887	24	4	1.1278	2.1750
comp4	19.8867	17.2651	11.7887	24	3	1.2422	2.3506
comp5	19.8867	17.2651	11.7887	24	2	1.4594	2.6841

Groups:

	Means	Groups
1	28.82	g1
2	23.98	g1g2
4	19.92	g2
6	18.70	g2g3
3	14.64	g2g3
5	13.26	g3

Tukey Midrange Test

Statistics:

Exp.Mean	CV	MSerror	Df	n	Stud.Mid	Ext.DMS	Int.DMS
19.89	17.27	11.78867	24	6	1.004855	1.986204	1.542945

Groups:

	Means	Groups
1	28.82	g1
2	23.98	g2
4	19.92	g3
6	18.70	g3
3	14.64	g4

```
5 13.26      g4
```

- `y` como vetor de médias dos tratamentos e o argumento `trt` como vetor dos tratamentos. Assim, deve-se informar na função `MRtest` que o argumento `y` é um vetor de médias, isto é, `ismean = TRUE`. Deve-se também informar o número de repetições, o valor do quadrado médio e o grau de liberdade do resíduo nos argumentos `replication`, `mserror` e `dferror`, respectivamente. Suponha que no exemplo 3.1, as informações sobre o experimento fossem apenas: a análise de variância, as médias e o número de repetições dos tratamentos, isto é:

```
anova(anova)
```

```
Analysis of Variance Table
```

```
Response: y
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
trt	5	847.05	169.409	14.37	1.485e-06 ***
Residuals	24	282.93	11.789		

```
# Média dos tratamentos
mean.trt <- c(28.82, 23.98, 14.64, 19.92, 13.26, 18.70)

# Tratamentos
treat <- as.factor(1:6)

# Núm. de repetições
repet <- 5

# Grau de liberdade do resíduo (obtido da ANAVA)
dferror <- 24

# Quadrado médio do resíduo (obtido da ANAVA)
mserror <- 11.78867
```

Com essas informações a função `MRtest` também realizará os testes quando `ismean = TRUE`. Observe os resultados abaixo.

```
# Analise da funcao MRtest - usando as medias

MRtest(y = mean.trt,
       trt = treat,
       dferror,
       mserror,
       replication = repet,
       ismean = TRUE)
```

MCP's based on distributions of the studentized midrange
and range

Study: $y \sim \text{trt}$

Summary:

	Means	r
1	28.82	5
2	23.98	5
3	14.64	5
4	19.92	5
5	13.26	5
6	18.70	5

Mean Grouping Midrange Test

Statistics:

Exp.Mean	CV	MSerror	Df	n	Stud.Mid	Ext.DMS	Int.DMS
19.89	17.27	11.78867	24	6	1.004855	1.986204	1.542945

Groups:

Means	Groups
1 28.82	g1
2 23.98	g2
4 19.92	g3

```

6 18.70    g3
3 14.64    g4
5 13.26    g4

```

Mean Grouping Range Test

Statistics:

Exp.Mean	CV	MSerror	Df	n	Stud.Range	DMS
19.88667	17.26515	11.78867	24	6	4.372651	6.714167

Groups:

Means	Groups
1 28.82	g1
2 23.98	g2
4 19.92	g2
6 18.70	g2
3 14.64	g3
5 13.26	g3

SNK Midrange Test

Statistics:

Exp.Mean	CV	MSerror	Df	n	Stud.Mid	DMS
comp1	19.8867	17.2651	11.7887	24	6	1.0049 1.9862
comp2	19.8867	17.2651	11.7887	24	5	1.0555 2.0640
comp3	19.8867	17.2651	11.7887	24	4	1.1278 2.1750
comp4	19.8867	17.2651	11.7887	24	3	1.2422 2.3506
comp5	19.8867	17.2651	11.7887	24	2	1.4594 2.6841

Groups:

Means	Groups
1 28.82	g1
2 23.98	g1g2
4 19.92	g2


```

6 18.70      g2g3
3 14.64      g2g3
5 13.26      g3

Tukey Midrange Test

Statistics:
Exp.Mean   CV  MSerror  Df n  Stud.Mid  Ext.DMS  Int.DMS
19.89     17.27 11.78867 24 6  1.004855  1.986204  1.542945

Groups:
Means Groups
1 28.82      g1
2 23.98      g2
4 19.92      g3
6 18.70      g3
3 14.64      g4
5 13.26      g4

```

Entretanto, o usuário pode não está interessado em realizar todos os testes. Dessa forma pode-se usar o argumento `MCP` para selecionar os testes desejados. Por exemplo, realizando os testes de `MGM` e de `MGR`, tem-se

```

MRtest(y = anava,
       trt = "trt",
       MCP = c("MGM", "MGR"))

```

MCP's based on distributions of the studentized midrange and range

```
Study: aov(y ~ trt) ~ "trt"
```

Summary:

```
Means  std r  Min  Max
1 28.82 5.80 5 19.4 33.0
```

```

2 23.98 3.78 5 17.7 27.9
3 14.64 4.12 5 9.1 19.4
4 19.92 1.13 5 18.6 21.0
5 13.26 1.43 5 11.6 14.4
6 18.70 1.60 5 16.9 20.8

```

Mean Grouping Midrange Test

Statistics:

```

Exp.Mean  CV  MSerror Df n Stud.Mid  Ext.DMS  Int.DMS
19.89     17.27 11.78867 24 6 1.004855 1.986204 1.542945

```

Groups:

Means Groups

```

1 28.82      g1
2 23.98      g2
4 19.92      g3
6 18.70      g3
3 14.64      g4
5 13.26      g4

```

Mean Grouping Range Test

Statistics:

```

Exp.Mean      CV  MSerror Df n Stud.Range      DMS
19.88667 17.26515 11.78867 24 6 4.372651 6.714167

```

Groups:

Means Groups

```

1 28.82      g1
2 23.98      g2
4 19.92      g2
6 18.70      g2
3 14.64      g3

```

```
5 13.26      g3
```

Mesmo armazenando os resultados da função `MRtest` em um objeto, estes são impressos no console do R. Entretanto, esses resultados podem ser recuperados. Veja a seguir.

```
teste.TM <- MRtest(y = anava,
                  trt = "trt",
                  MCP = c("TM"))
```

```
MCP's based on distributions of the studentized midrange
and range
```

```
Study:  anava ~ "trt"
```

```
Summary:
```

	Means	std	r	Min	Max
1	28.82	5.80	5	19.4	33.0
2	23.98	3.78	5	17.7	27.9
3	14.64	4.12	5	9.1	19.4
4	19.92	1.13	5	18.6	21.0
5	13.26	1.43	5	11.6	14.4
6	18.70	1.60	5	16.9	20.8

```
Tukey Midrange Test
```

```
Statistics:
```

Exp.Mean	CV	MSerror	Df	n	Stud.Mid	Ext.DMS	Int.DMS
19.89	17.27	11.78867	24	6	1.004855	1.986204	1.542945

```
Groups:
```

	Means	Groups
1	28.82	g1
2	23.98	g2
4	19.92	g3

6	18.70	g3
3	14.64	g4
5	13.26	g4

Para recuperá-los no objeto `teste.TM` basta usar, após o nome do objeto, o símbolo `$`, como apresentado na Figura 18.

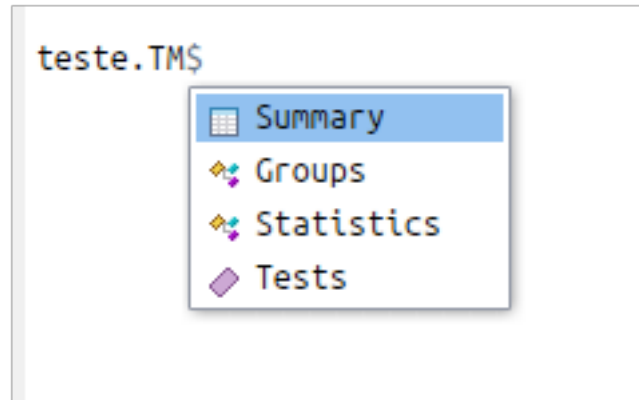


Figura 18 Como acessar as opções de um objeto que recebe a função `MRtest`.

Caso deseje obter apenas os resultados do teste escolhido, basta usar:

```
teste.TM$Groups
```

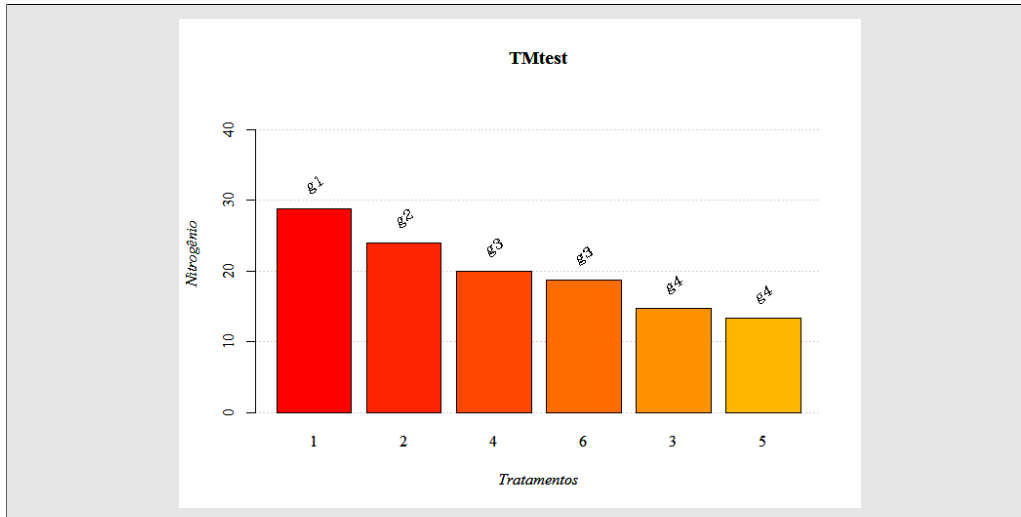
```
$group.TMtest
```

	Means	Groups
1	28.82	g1
2	23.98	g2
4	19.92	g3
6	18.70	g3
3	14.64	g4
5	13.26	g4

Para mais detalhes, ver a subseção 3.3.2. Após ter criado um objeto para a função `MRtest`, as outras funções do pacote `midrangeMCP`, `MRbarplot` e `MRwrite`, ficam fáceis de serem utilizadas.

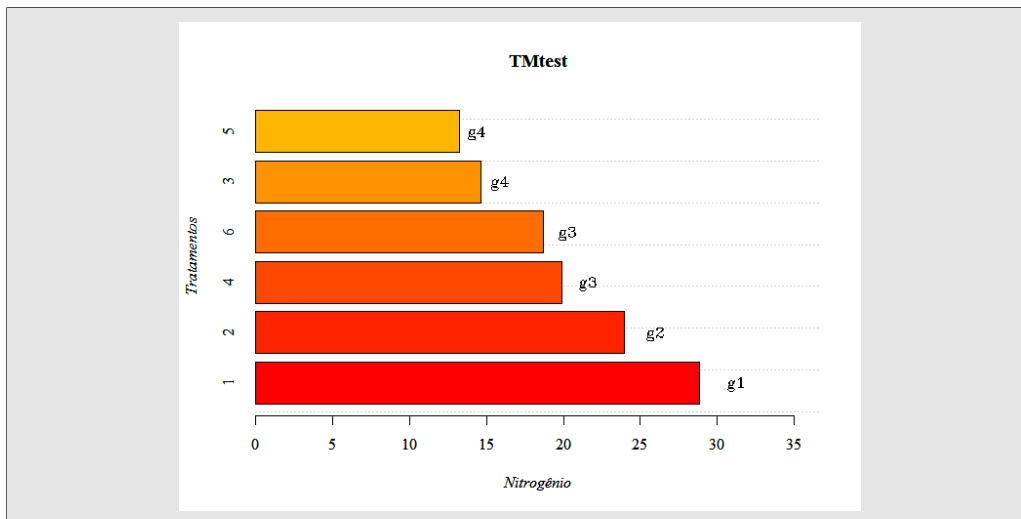
A seguir, será usado a função `MRbarplot` para o objeto `teste.TM`.

```
MRbarplot(teste.TM, xlab = "Tratamentos",
          ylab = "Nitrogênio")
```



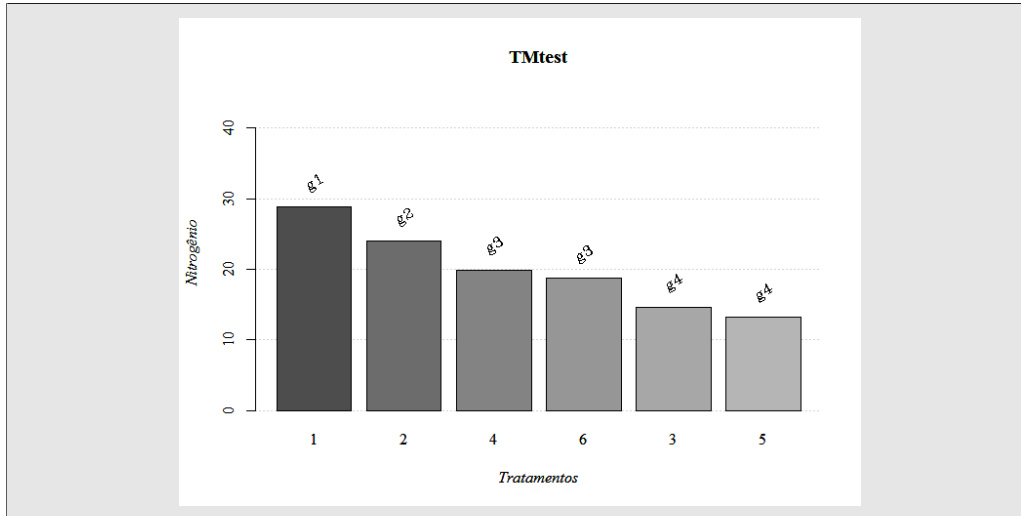
Caso deseje as barras na vertical, basta alterar o argumento `horiz` na função `MRbarplot`.

```
MRbarplot(teste.TM, xlab = "Tratamentos",
          ylab = "Nitrogênio", horiz = TRUE)
```



Para publicação de artigos algumas revistas cobram por figuras coloridas. Para plotar o gráfico dos resultados dos testes com o degradê nos tons de preto e cinza, basta alterar o argumento `col` na função `MRbarplot`.

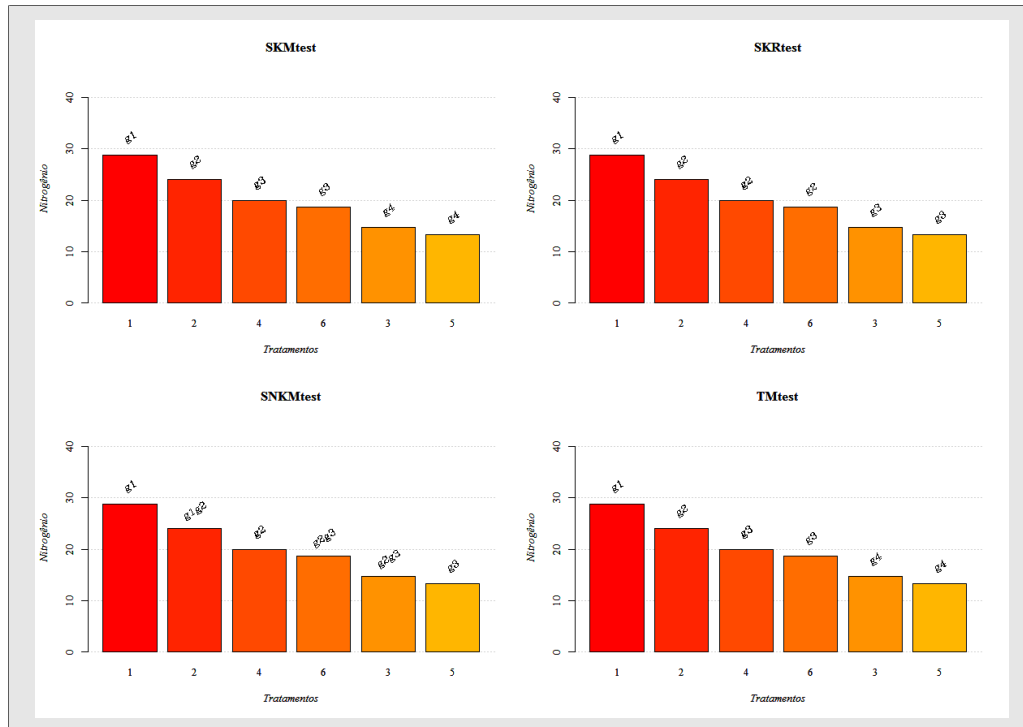
```
MRbarplot(teste.TM, xlab = "Tratamentos",
           ylab = "Nitrogênio", col = gray.colors(10))
```



Entretanto, se na função `MRtest` o usuário escolher mais de um teste (usando o argumento `MCP`), a função `MRbarplot` plotará o resultado de todos os testes escolhidos. Ver exemplo abaixo.

```
# Por "default" MCP = "all", isto é, todos os testes
# serão analisados
teste.all <- MRtest(y = anava, trt = "trt")

# Plotando o resultado dos testes
MRbarplot(teste.TM, xlab = "Tratamentos",
           ylab = "Nitrogênio", horiz = TRUE)
```

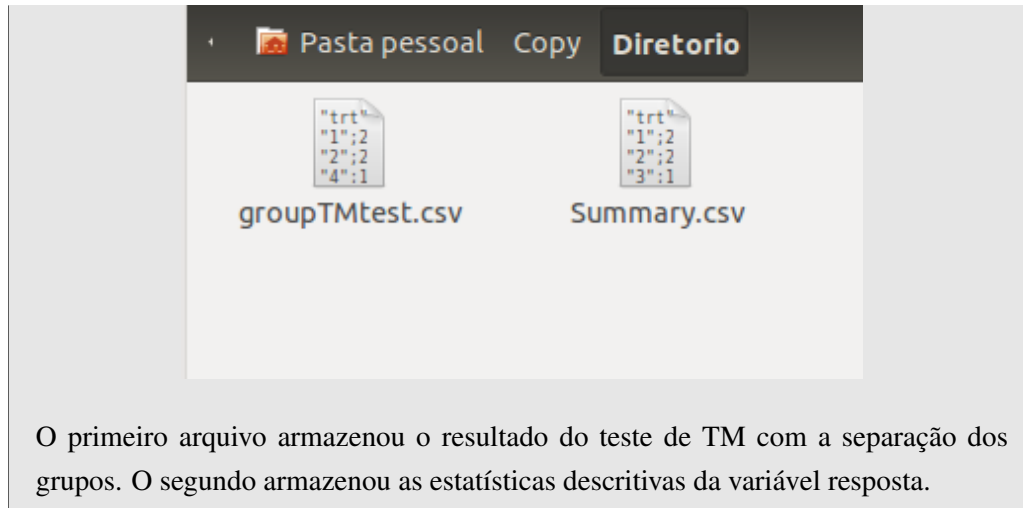


A última função apresentada do pacote `midrangeMCP` é `MRwrite`, tendo por objetivo exportar os principais resultados da análise dos testes. Por *default* essa função sempre exportará os resultados para um arquivo com a extensão “<arquivo>.csv”. Lembrando que os arquivos sempre serão exportados para o diretório de trabalho do R. Para a escolha do diretório no R, basta usar a função `setwd`. A seguir um exemplo para a função `MRwrite` usando o objeto `teste.TM`.

```
# (Sistema Operacional Linux - Ubuntu 12.04)
# Mudando o diretorio
setwd("~/Copy/Diretorio")

# Exportando os resultados de "teste.TM"
MRwrite(teste.TM)
```

```
See your files in Directory
Format: csv
Os arquivos exportados:
```

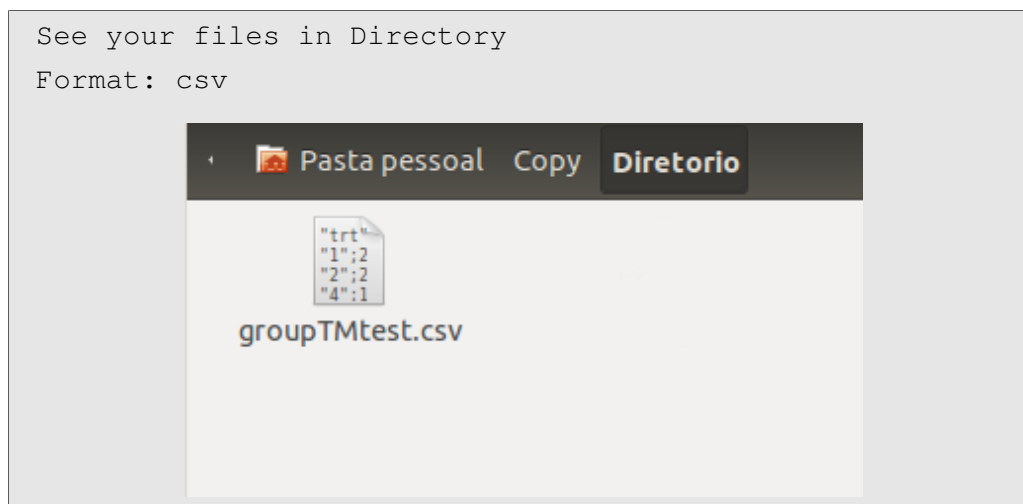


O primeiro arquivo armazenou o resultado do teste de TM com a separação dos grupos. O segundo armazenou as estatísticas descritivas da variável resposta.

Caso não deseje exportar esses dois arquivos, basta usar o argumento `dataMR`. Por exemplo, se o usuário opta por simplesmente exportar o resultado da análise, basta usar:

```
# (Sistema Operacional Linux - Ubuntu 12.04)
# Mudando o diretorio
setwd("~/Copy/Diretorio")

# Exportando os resultados de "teste.TM"
MRwrite(teste.TM, dataMR = "groups")
```



As outras extensões possíveis são: “<arquivo>.txt” e “<arquivo>.xlsx”. Para detalhes, ver a subseção 3.3.4. Agora, para quem usa \LaTeX é possível exportar o código em formato

de tabela desses resultados.

```
# (Sistema Operacional Linux - Ubuntu 12.04)
# Mudando o diretorio
setwd("~/Copy/Diretorio")

# Exportando os resultados de "teste.TM"
MRwrite(teste.TM, dataMR = "groups", extension = "latex")
```

```
% latex table generated in R 3.2.2
% by xtable 1.7-4 package
\begin{table}[ht]
\centering
\begin{tabular}{lrl}
\hline
trt & Means & Groups \\
\hline
1 & 28.82 & g1 \\
2 & 23.98 & g2 \\
4 & 19.92 & g3 \\
6 & 18.70 & g3 \\
3 & 14.64 & g4 \\
5 & 13.26 & g4 \\
\hline
\end{tabular}
\end{table}
```

See yours tables in Console

Format: latex

Basta copiar o código e inserir em algum editor de texto do \LaTeX . O resultado fica dessa forma:

trt	Means	Groups
1	28.82	g1
2	23.98	g2
4	19.92	g3
6	18.70	g3
3	14.64	g4
5	13.26	g4

4.3.6 Comparando os resultados dos testes

Usando o Exemplo 3.1, na Tabela 9 será apresentado os resultados dos testes propostos (TM e SNKM) comparando com os testes originais (teste de Tukey e de SNK).

Tabela 9 Resultados da análise do Exemplo 3.1 dos testes de TM e de SNKM, de Tukey, de SNK, de MGM, de MGR e de Scott-Knott.

Trat.	Médias	Testes						
		TM	Tukey	SNKM	SNK	MGM	MGR	Scott-Knott
1	28,82	g1	a	g1	a	g1	g1	a
2	23,98	g2	ab	g1g2	b	g2	g2	b
4	19,92	g3	bc	g2	bc	g3	g2	c
6	18,70	g3	bc	g2g3	bc	g3	g2	c
3	14,64	g4	c	g2g3	cd	g4	g3	d
5	13,26	g4	c	g3	d	g4	g3	d

O teste de TM apresentou uma maior separação de grupos, resultando assim, numa melhor interpretação prática para os tratamentos. Observe que o teste de TM, para esse exemplo, conseguiu separar os grupos de médias sem ambiguidade e igual aos resultados dos testes de MGM e de MGR. Algo que não ocorreu com o teste de Tukey original. O resultado do teste de SNKM diferiu do resultado do teste de SNK. Portanto, de modo prático os testes de TM, MGM e MGR conseguem informar que o efeito da bactéria 3DOK1 produziu um maior teor de nitrogênio que as demais.

Utilizando agora o Exemplo 3.2, será apresentado como proceder a realização dos testes, utilizando o pacote `MidrangeMCP`, uma entrada de dados diferente da feita no Exemplo 3.1. Os dados analisados pelo pacote serão as médias dos genótipos, com informações adicionais sobre o quadrado médio do resíduo do experimento e seu grau de

liberdade, e o número de repetições das médias.

No R, os genótipos e suas respectivas médias do período de floração foram armazenadas no objeto `dados`, como observado abaixo.

```
dados
```

```

      genotipo    flow
1      BR501 83.2543
2      BR505 79.9100
3      BR506 84.3286
4      BR507 85.1686
5      BR508 83.7000
6     BRS511 81.4200
7   CMSXS629 83.2671
8   CMSXS630 85.7471
9   CMSXS633 80.9100
10  CMSXS635 82.4800
11  CMSXS636 75.4286
12  CMSXS637 79.5886
13  CMSXS639 78.1500
14  CMSXS643 87.5143
15  CMSXS644 82.3543
16  CMSXS646 78.5914
17  CMSXS647 77.2871
18  CMSXS648 81.1243
19     BRS601 78.3300
20 SUGARGRAZE 75.4471
21     V82391 75.3614
22     V82392 72.5243
23     V82393 73.8300
24  XBSW80007 75.1486
25  XBSW80140 79.3500

```

Posteriormente, usou-se a função `MRtest` com o argumento `ismean = TRUE` para informar que o argumento `y` está recebendo as médias dos genótipos.

```

y <- dados$genotipo
trt <- dados$flow
MRtest(y, trt, dferror = 252, mserror = 6.3078,
       replication = 21, alpha = 0.05,
       MCP = c("MGM", "MGR"), ismean = TRUE)

```

Mean Grouping Midrange Test

Statistics:

Exp.Mean	CV	MSerror	Df	n	Stud.Mid	Ext.DMS	Int.DMS
80.01	3.14	6.3078	252	25	0.72453	0.475	0.397

Groups:

Means Groups

CMSXS643	87.51	g1
CMSXS630	85.75	g2
BR507	85.17	g2
BR506	84.33	g3
BR508	83.70	g3
CMSXS629	83.27	g3
BR501	83.25	g3
CMSXS635	82.48	g3
CMSXS644	82.35	g3
BRS511	81.42	g4
CMSXS648	81.12	g4
CMSXS633	80.91	g4
BR505	79.91	g5
CMSXS637	79.59	g5
XBSW80140	79.35	g5
CMSXS646	78.59	g5
BRS601	78.33	g5
CMSXS639	78.15	g5
CMSXS647	77.29	g5
SUGARGRAZE	75.45	g6

CMSXS636	75.43	g6
V82391	75.36	g6
XBSW80007	75.15	g6
V82393	73.83	g7
V82392	72.52	g8

Mean Grouping Range Test

Statistics:

Exp.Mean	CV	MSerror	Df	n	Stud.Range	DMS
80.01	3.14	6.3078	252	25	5.231999	2.88

Groups:

Means Groups

CMSXS643	87.51	g1
CMSXS630	85.75	g2
BR507	85.17	g2
BR506	84.33	g3
BR508	83.70	g3
CMSXS629	83.27	g3
BR501	83.25	g3
CMSXS635	82.48	g3
CMSXS644	82.35	g3
BRS511	81.42	g4
CMSXS648	81.12	g4
CMSXS633	80.91	g4
BR505	79.91	g5
CMSXS637	79.59	g5
XBSW80140	79.35	g5
CMSXS646	78.59	g5
BRS601	78.33	g5
CMSXS639	78.15	g5
CMSXS647	77.29	g5
SUGARGRAZE	75.45	g6

CMSXS636	75.43	g6
V82391	75.36	g6
XBSW80007	75.15	g6
V82393	73.83	g7
V82392	72.52	g7

Tabela 10 Resultados dos testes de MGM, de MGR e de Scott-Knott avaliando os 25 genótipos de sorgo apresentado no Exemplo 3.2.

Genótipo	Médias	Diferença entre médias	Testes		
			MGM	MGR	Scott-Knott
CMSXS643	87.51	-	g1	g1	A
CMSXS630	85,75	1,77	g2	g2	B
BR507	85,17	0,58	g2	g2	B
BR506	84,33	0,84	g3	g3	B
BR508	83,70	0,63	g3	g3	C
CMSXS629	83,27	0,43	g3	g3	C
BR501	83,25	0,01	g3	g3	C
CMSXS635	82,48	0,77	g3	g3	C
CMSXS644	82,35	0,13	g3	g3	C
BRS511	81,42	0,93	g4	g4	D
CMSXS648	81,12	0,30	g4	g4	D
CMSXS633	80,91	0,21	g4	g4	D
BR505	79,91	1,00	g5	g5	E
CMSXS637	79,59	0,32	g5	g5	E
XBSW80140	79,35	0,24	g5	g5	E
CMSXS646	78,59	0,76	g5	g5	E
BRS601	78,33	0,26	g5	g5	E
CMSXS639	78,15	0,18	g5	g5	E
CMSXS647	77,29	0,86	g5	g5	E
SUGARGRAZE	75,45	1,84	g6	g6	F
CMSXS636	75,43	0,02	g6	g6	F
V82391	75,36	0,07	g6	g6	F
XBSW80007	75,15	0,21	g6	g6	F
V82393	73,83	1,32	g7	g7	G
V82392	72,52	1,31	g8	g7	G

Na Tabela 10, os resultados dos testes são apresentados para a comparação. Nessa tabela também é enfatizado as diferenças consecutivas entre as médias ordenadas para auxílio na comparação dos resultados dos testes. Um outro aspecto é o destaque nas linhas em que um dos testes separou o grupo de médias.

Os resultados mostram que os testes propostos (MGM e MGR) apresentaram uma maior separação dos grupos de médias do que o teste de Scott-Knott de modo mais coerente. Os testes propostos apresentaram resultados bem semelhantes. As médias foram ordenadas para facilitar a discussão. Observe a diferença dos resultados dos testes nos primeiros grupos de médias. As médias dos genótipos BR507 e BR506 foram considerados estatisticamente iguais pelo teste de Scott-Knott, mas diferentes pelo teste de MGM e de MGR. Posteriormente, as médias dos genótipos BR506 e BR508 foram consideradas estatisticamente iguais pelos testes propostos, porém diferentes pelo teste de Scott-Knott. Ocorre uma incoerência no teste de Scott-Knott, muito comum na prática. Perceba a diferença $\bar{Y}_{BR507} - \bar{Y}_{BR506} = 0,84$. O valor de 0,84 entre essas duas médias não foi o suficiente para o teste de Scott-Knott detectar que estas são amostradas de populações com médias diferentes. Entretanto, esse mesmo teste detectou que a diferença $\bar{Y}_{BR506} - \bar{Y}_{BR508} = 0,63$ era significativa, e portanto, o efeitos médio desses genótipos são diferentes. Isso se deve a filosofia como o teste de Scott-Knott foi desenvolvido. A separação de grupos ocorre pela razão de verossimilhança entre grupos. As diferenças entre médias limitantes de cada grupo podem muitas vezes ser menor do que as diferenças entre médias consecutivas dentro dos grupos.

Ao contrário do teste de Scott-Knott, os testes de MGM e de MGR são mais coerentes quanto a este aspecto. A diferença entre os genótipos BR506 e BR508 de 0,63 não foi suficiente para os testes propostos avaliarem esses dois genótipos como estatisticamente diferentes. Entretanto, para a diferença maior entre os genótipos BR507 e BR506 de 0,84 estes foram estatisticamente diferentes.

Entretanto, em uma situação o teste de MGR também não se livrou desse aspecto. Verificando a diferença entre os genótipos V82393 e V82392 que foi de 1,31, o teste de MGR não detectou diferença entre essas médias, como também não foi verificado pelo teste de Scott-Knott. Esse questionamento se deve ao fato da diferença entre os genótipos BR507 e BR506 de 0,84 ter sido detectado como uma diferença significativa pelo teste de MGR. Já para o teste de MGM isso não ocorre, a diferença para os genótipos V82393 e V82392 de 1,31 foi detectado como genótipos estatisticamente diferentes. Apenas em uma situação nenhum dos testes detectou significância numa diferença de 0,86 (diferença entre os genótipos CMSXS639 e CMSXS642). A menor diferença significativa detectada para os testes de MGM e de MGR foi de 0,84, e para o teste de Scott-Knott foi de 0,63. Assim, todos os testes acima desses valores deveriam também detectar diferença. Vale lembrar que para os testes propostos, os valores 0,84 e 0,86 estão bem próximos, sendo

um limiar para esses testes detectarem a significância na diferença entre as médias.

Em todas as outras situações em que o teste de Scott-Knott diferenciou os grupos de médias os testes de MGM e de MGR conseguiram também detectar. Levando em consideração que o teste de MGM refinou ainda mais a separação de grupos.

A maior separação de modo mais coerente ocorre nos testes de MGM e de MGR devido o desenvolvimento de como os testes foram propostos. A separação dos grupos desses testes leva em consideração a maior diferença consecutiva entre médias, e isso foi determinante para que não houvesse a incoerência que muitas vezes ocorre no teste de Scott-Knott.

5 CONCLUSÃO

Os testes propostos de MGM, de MGR e de TM apresentaram desempenho superior aos testes originais de Tukey, de SNK e de Skott-Knott na maioria das avaliações feitas, exceto para o erro tipo I por experimento sob H_0 parcial. Até o momento, os únicos testes que apresentam esse controle são o teste de Tukey e o teste de Scheffé, muito embora apresentem baixo poder. Apesar de o poder, sob H_0 parcial, ter sido baixo para todos os testes, o poder dos testes propostos nesse estudo foram superiores ao poder dos testes originais. O teste de SNKM foi o único a apresentar baixo desempenho quando comparado com o teste original (SNK).

O teste de TM, de MGM e de MGR apresentaram uma maior separação de grupos que os testes de Tukey e de Scott-Knott, levando em consideração que além da ausência de ambiguidade nos resultados dos testes de MGM e MGR, separaram os grupos de médias de forma mais coerente que a do teste Scott-Knott, uma vez que os testes propostos se baseiam na máxima diferença entre médias consecutivas para separação de grupos.

A vantagem de utilização dos testes é o uso do pacote `midrangeMCP` desenvolvido no R, que além de realizar os testes, permite uma apresentação dos resultados por meio de gráficos e até mesmo sua exportação para de banco de dados externo, como MS Excel, códigos em \LaTeX , entre outros.

REFERÊNCIAS

BANZATTO, D. A.; KRONKA, S. do N. **Experimentação Agrícola**. 4. ed. Jaboticabal: Funep, 2006. 237 p.

BATISTA, B. D. de O. **Distribuição exata da *midrange* estudentizada externamente da normal e desenvolvimento de uma biblioteca R utilizando quadratura gaussiana**. 97 p. Dissertação (Mestrado) — Universidade Federal de Lavras, Lavras-MG, 2012.

BATISTA, B. D. de O.; FERREIRA, D. F. SMR: An R package for computing the externally studentized normal midrange distribution. **The R Journal**, v. 6, n. 2, p. 123–136, dez. 2014.

BATISTA, B. D. de O.; FERREIRA, D. F. **SMR**: Externally studentized midrange distribution. Vienna, Austria, 2014. R package version 2.0.1. Disponível em: <<http://cran.r-project.org/web/packages/SMR/index.html>>.

BATISTA, B. D. de O.; FERREIRA, D. F. **midrangeMCP**: Multiple comparisons procedures based on studentized midrange and range distributions. Vienna, Austria, 2016. R package version 1.2. Disponível em: <<http://CRAN.R-project.org/package=midrangeMCP>>.

BENJAMINI, Y.; BRETZ, F.; SARKAR, S. (Ed.). **Recent Developments in Multiple Comparison Procedures**. Beachwood, Ohio, USA: Institute of Mathematical Statistics, 2004. 156 p.

BERNHARDSON, C. S. 375: Type I error rates when multiple comparison procedures follow a significant F test of anova. **Biometrics**, International Biometric Society, v. 31, n. 1, p. 229–232, 1975.

BOARDMAN, T. J.; MOFFITT, D. R. Graphical monte carlo type I error rates for multiple comparison procedures. **Biometrics**, International Biometric Society, v. 27, n. 3, p. 738–744, 1971.

BORGES, L. C.; FERREIRA, D. F. Poder e taxas de erro tipo I dos testes Scott-Knott, Tukey e Student-Newman-Keuls sob distribuições normal e não normais dos resíduos. **Revista Matemática e Estatística**, v. 21, n. 1, p. 67–83, 2003.

BRETZ, F.; HOTHORN, T.; WESTFALL, P. **Multiple comparisons using R**. New York: CRC Press, 2011. 182 p.

CARMER, S. G.; SWANSON, M. R. An evaluation of ten pairwise multiple comparison procedures by monte carlo methods. **Journal of the American Statistical Association**, v. 68, n. 341, p. 66–74, 1973.

DAHL, D. B. **xtable**: Export tables to latex or html. Vienna, Austria, 2014. R package version 1.7-4. Disponível em: <<http://CRAN.R-project.org/package=xtable>>.

DAVID, H. A.; HARTLEY, H. O.; PEARSON, E. S. The distribution of the ratio, in a single normal sample, of range to standard deviation. **Biometrika**, v. 41, n. 3/4, p. 482–493, 1954.

DAVID, H. A.; NAGARAJA, H. N. **Order Statistics**. Canada: John Wiley & Sons, 2003. 458 p.

DICKHAUS, T. **Simultaneous Statistical Inference**: with applications in the life sciences. Berlin: Springer, 2014. 180 p.

DUNCAN, D. **Significance Tests for Differences Between Ranked Variates Drawn from Normal Populations**. 67 p. Tese (Unpublished doctoral thesis) — Iowa State College, Ames, Iowa, 1947.

DUNCAN, D. A significance test for differences between ranked treatments in an analysis of variance. **Virginia Journal of Science**, v. 2, p. 171–189, 1951.

DUNCAN, D. B. On the properties of the multiple comparisons test. **Virginia Journal of Science**, v. 3, p. 50–67, 1952.

DUNCAN, D. B. Multiple range and multiple F tests. **Biometrics**, v. 11, p. 1–42, 1955.

DUNCAN, D. B. A bayesian approach to multiple comparisons. **Technometrics**, v. 7, n. 2, p. 171–222, 1965.

DUNNETT, C. W. Multiple comparisons with a standard. **Proc. 9th. Ann. Convention American Society for Quality Control**, p. 485–492, May 1955.

EDWARDS, A. W. F.; CAVALLI-SFORZA, L. L. A method for cluster analysis. **Biometrics**, v. 21, n. 2, p. 362–375, 1965.

FERREIRA, D. F. **Estatística Básica**. 2. ed. Lavras: Editora UFLA, 2009. 664 p.

FERREIRA, D. F. **Estatística computacional em Java**. Lavras: Editora UFLA, 2013. 695 p.

FIGUEIREDO, U. J. de et al. Adaptability and stability of genotypes of sweet sorghum by ggebiplot and toler methods. **Genetics and Molecular Research**, v. 14, n. 3, p. 11211–11221, 2015.

FISHER, R. A. The logic of inductive inference. **Journal of the Royal Statistical Society**, v. 98, n. 1, p. 39–82, 1935.

FISHER, W. D. On grouping for maximum homogeneity. **Journal of the American Statistical Association**, v. 53, n. 284, p. 789–798, 1958.

GALTON, F. F. The most suitable proportion between the value of first and second prizes. **Biometrika**, v. 1, n. 4, p. 385–399, 1902.

GEARY, R. C. Moments of the ratio of the mean deviation to the standard deviation for normal samples. **Biometrika**, v. 28, n. 3/4, p. 295–307, 1936.

GELMAN, A.; HILL, J.; YAJIMA, M. Why we (usually) don't have to worry about multiple comparisons. **Journal of Research on Educational Effectiveness**, v. 5, p. 189–211, 2012.

GRAYBILL, F. **An introduction to linear statistical models**. New York: McGraw-Hill,

1961. 463 p.

GUMBEL, E. J. The distribution of the range. **The Annals of Mathematical Statistics**, v. 18, n. 3, p. 384–412, 1947.

GUMBEL, E. J. **Statistics of Extremes**. New York: Columbia University Press, 1958. 375 p.

GUMBEL, E. J.; CARLSON, P. G.; MUSTAFI, C. K. A note on midrange. **The Annals of Mathematical Statistics**, v. 36, n. 3, p. 1052–1054, 1965.

HARTLEY, H. O. The range in random samples. **Biometrika**, v. 32, n. 3/4, p. 334–348, 1942.

HARTLEY, H. O. Some recent developments in analysis of variance. **Communications on Pure and Applied Mathematics**, III, p. 47–72, 1955.

HINKELMANN, K.; KEMPTHORNE, O. **Design and Analysis of Experiments: Introduction to experimental design**. 2. ed. New Jersey: Wiley, 2007. I. 631 p.

HOCHBERG, Y.; TAMHANE, A. **Multiple comparison procedures**. New York: Wiley, 1987. 450 p.

HOCKING, R. R. **Methods and Applications of Linear Models: Regression and the analysis of variance**. 2. ed. New Jersey: Wiley, 2003. 741 p.

HSU, J. C. **Multiple comparisons: Theory and methods**. London: Chapman & Hall/CRC, 1996. 277 p.

IRWIN, J. O. On a criterion for the rejection of outlying observations. **Biometrika**, v. 17, n. 3/4, p. 238–250, 1925.

KAMAT, A. R. Some properties of estimates for the standard deviation based on deviations from the mean and variate differences. **Journal of the Royal Statistical Society**, v. 15, n. 2, p. 233–240, 1953.

KEULS, M. The use of the “studentized range” in connection with an analysis of variance. **Euphytica**, v. 1, p. 112–122, 1952.

KRAMER, C. Y. Extension of multiple range tests to group means with unequal numbers of replications. **Biometrics**, v. 12, n. 3, p. 307–310, 1956.

LEEMIS, L. M.; TRIVEDI, K. S. A comparison of approximate interval estimators for the bernoulli parameter. **The American Statistician**, v. 50, n. 1, p. 63–68, 1996.

MCKAY, A. T.; PEARSON, E. S. A note on the distribution of range in samples of n . **Biometrika**, v. 25, n. 3/4, p. 415–420, 1933.

MILLER, R. **Simultaneous statistical inference**. 1. ed. New York: McGraw-Hill, 1966.

MILLER, R. Developments in multiple comparisons 1966-1976. **Journal of the American Statistical Association**, v. 72, n. 360, p. 779–788, 1977.

MILLER, R. **Simultaneous statistical inference**. 2. ed. New York: Springer-Verlag, 1981. 299 p.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the Theory of Statistics**. New York: McGraw-Hill, 1974. 564 p.

NEWMAN, D. The distribution of range in samples from a normal population, expressed in terms of an independent estimate of standard deviation. **Biometrika**, v. 31, n. 1/2, p. 20–30, 1939.

OLIVEIRA, I. R. C.; FERREIRA, D. F. Multivariate extension of chi-squared univariate normality test. **Journal of Statistical Computation and Simulation**, v. 80, n. 5, p. 513–526, 2010.

O’NEILL, R.; WETHERILL, G. B. The present state of multiple comparison methods. **Journal of the Royal Statistical Society**, v. 33, n. 2, p. 218–250, 1971.

PEARSON, E. S. A further note on the distribution of range in samples taken from a

normal population. **Biometrika**, v. 18, n. 1/2, p. 173–194, 1926.

PEARSON, E. S. The percentage limits for the distribution of range in samples from a normal population. **Biometrika**, v. 24, n. 3/4, p. 404–417, 1932.

PEARSON, E. S. Student as statistician. **Biometrika**, v. 30, n. 3/4, p. 210–250, 1939.

PEARSON, E. S.; HAINES, J. The use of range in place of standard deviation in small samples. **Supplement to the Journal of the Royal Statistical Society**, v. 2, n. 1, p. 83–98, 1935.

PEARSON, E. S.; HARTLEY, H. O. The probability integral of the range in samples of n observations from a normal population. **Biometrika**, v. 32, n. 3/4, p. 301–310, 1942.

PEARSON, E. S.; HARTLEY, H. O. Tables of the probability integral of the studentized range. **Biometrika**, v. 33, n. 1, p. 89–99, 1943.

PEARSON, K. Note on Francis Galton's problem. **Biometrika**, v. 1, p. 390–339, 1902.

PERECIN, D.; MALHEIROS, E. B. Uma avaliação de seis procedimentos para comparações múltiplas. In: ESCOLA SUPERIOR DE LAVRAS. 3^o **Simpósio de Estatística aplicada à Experimentação Agonômica**. Lavras, MG, 1989. p. 66.

R CORE TEAM. **R**: A language and environment for statistical computing. Vienna, Austria, 2015. Disponível em: <<https://www.R-project.org/>>.

RAMALHO, M. A. P.; FERREIRA, D. F.; OLIVEIRA, A. C. de. **Experimentação em genética e melhoramento de plantas**. 2. ed. Lavras: Editora UFLA, 2005. 322 p.

RAO, C. R.; SWARUPCHAND, U. Multiple comparison procedure - a note and a bibliography. **Journal of Statistics**, v. 16, p. 66–109, 2009.

RIDER, P. R. The midrange of a sample as an estimator of the population midrange. **Journal of the American Statistical Association**, v. 52, n. 280, p. 537–542, Dec. 1957.

SCHEFFÉ, H. A method for judging all contrasts in the analysis of variance. **Biometrika**, v. 40, n. 1/2, p. 87–104, 1953.

SCHWARTZ, M. **WriteXLS**: Cross-platform perl based r function to create excel 2003 (xls) and excel 2007 (xlsx) files. Vienna, Austria, 2015. R package version 3.6.1. Disponível em: <<http://CRAN.R-project.org/package=WriteXLS>>.

SCOTT, A. J.; KNOTT, M. A cluster analysis method for grouping means in the analysis of variance. **Biometrics**, International Biometric Society, v. 30, n. 3, p. 507–512, 1974.

SEARLE, S. R. **Linear models for unbalanced data**. New York: Wiley, 1987. 536 p.

SILVA, E. C. da; FERREIRA, D. F.; BEARZOTI, E. Avaliação do poder e taxas de erro tipo I do teste de Scott-Knott por meio do método de monte carlo. **Ciência e Agrotecnologia**, v. 23, n. 3, p. 687–696, jul./set. 1999.

STEEL, R.; TORRIE, J. **Principles and Procedures of Statistics**: A biometrical approach. 2. ed. New York: McGraw-Hill, 1980. 666 p.

STUDENT. Errors in routine analysis. **Biometrika**, v. 19, p. 151–164, 1927.

TIPPET, L. H. C. On the extreme individuals and the range of samples taken from a normal population. **Biometrika**, v. 17, p. 365–387, 1925.

TUKEY, J. W. Comparing individual means in the analysis of variance. **Biometrics**, v. 5, p. 99–114, 1949.

TUKEY, J. W. Allowances for various types of error rates. Unpublished invited address presented at Blacksburg meeting of Institute of Mathematical Statistics and Biometric Society. 1952.

TUKEY, J. W. The problem of multiple comparisons. Unpublished memorandum in private circulation. 1953.

TUKEY, J. W. The philosophy of multiple comparisons. **Statistical Science**, v. 6, n. 1,

p. 100–116, 1991.

WALLER, R. A.; DUNCAN, D. B. A Bayes rule for the symmetric multiple comparisons problems. **Journal of the American Statistical Association**, v. 64, n. 328, p. 1484–1503, dez. 1969.

WILCOX, R. R. New designs in analysis of variance. **Annu. Rev. Psychol.**, International Biometric Society, v. 38, p. 29–60, 1987.

WORKING, H.; HOTELLING, H. Applications of the theory of error to the interpretation of trends. **Journal of the American Statistical Association**, v. 24, n. 165, p. 73–85, 1929.

ZIEFFLER, A. S.; HARRING, J. R.; LONG, J. D. **Comparing Groups: Randomization and bootstrap methods using R**. New Jersey: Wiley, 2011. 298 p.

APÊNDICE A - Erro tipo I sob H_0 parcial

Tabela 11 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 5\%$ de probabilidade e $\delta = 8\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Trat	Rep	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	2,820 ⁻⁻	9,720 ⁺⁺	2,980 ⁻⁻	31,780 ⁺⁺	17,380 ⁺⁺	31,780 ⁺⁺
	10	2,780 ⁻⁻	10,900 ⁺⁺	2,900 ⁻⁻	34,580 ⁺⁺	18,820 ⁺⁺	34,580 ⁺⁺
	20	2,240 ⁻⁻	9,400 ⁺⁺	2,400 ⁻⁻	35,740 ⁺⁺	18,480 ⁺⁺	35,740 ⁺⁺
10	4	3,440 ⁻⁻	10,380 ⁺⁺	3,660 ⁻⁻	55,100 ⁺⁺	39,200 ⁺⁺	55,100 ⁺⁺
	10	2,440 ⁻⁻	9,360 ⁺⁺	2,500 ⁻⁻	56,220 ⁺⁺	40,200 ⁺⁺	56,220 ⁺⁺
	20	2,680 ⁻⁻	9,880 ⁺⁺	2,740 ⁻⁻	57,540 ⁺⁺	40,100 ⁺⁺	57,540 ⁺⁺
20	4	2,880 ⁻⁻	9,220 ⁺⁺	2,940 ⁻⁻	66,980 ⁺⁺	57,720 ⁺⁺	67,020 ⁺⁺
	10	2,880 ⁻⁻	9,660 ⁺⁺	3,080 ⁻⁻	67,420 ⁺⁺	58,120 ⁺⁺	67,440 ⁺⁺
	20	2,700 ⁻⁻	9,560 ⁺⁺	2,840 ⁻⁻	68,480 ⁺⁺	58,780 ⁺⁺	68,480 ⁺⁺
40	4	3,480 ⁻⁻	10,380 ⁺⁺	3,560 ⁻⁻	64,400 ⁺⁺	60,020 ⁺⁺	64,440 ⁺⁺
	10	3,080 ⁻⁻	9,740 ⁺⁺	3,220 ⁻⁻	66,120 ⁺⁺	61,240 ⁺⁺	66,180 ⁺⁺
	20	2,980 ⁻⁻	9,440 ⁺⁺	3,200 ⁻⁻	66,240 ⁺⁺	61,680 ⁺⁺	66,280 ⁺⁺
100	4	3,160 ⁻⁻	10,000 ⁺⁺	3,420 ⁻⁻	58,960 ⁺⁺	56,640 ⁺⁺	59,100 ⁺⁺
	10	3,280 ⁻⁻	9,420 ⁺⁺	3,500 ⁻⁻	59,280 ⁺⁺	56,580 ⁺⁺	59,520 ⁺⁺
	20	3,320 ⁻⁻	10,560 ⁺⁺	3,700 ⁻⁻	58,020 ⁺⁺	55,580 ⁺⁺	58,280 ⁺⁺

Tabela 12 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 5\%$ de probabilidade e $\delta = 16\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Trat	Rep	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	2,640 ⁻⁻	9,140 ⁺⁺	2,640 ⁻⁻	32,960 ⁺⁺	18,700 ⁺⁺	32,960 ⁺⁺
	10	2,820 ⁻⁻	10,340 ⁺⁺	2,820 ⁻⁻	34,240 ⁺⁺	18,120 ⁺⁺	34,240 ⁺⁺
	20	2,460 ⁻⁻	10,140 ⁺⁺	2,460 ⁻⁻	34,860 ⁺⁺	17,700 ⁺⁺	34,860 ⁺⁺
10	4	2,700 ⁻⁻	9,040 ⁺⁺	2,700 ⁻⁻	53,880 ⁺⁺	37,600 ⁺⁺	53,880 ⁺⁺
	10	2,560 ⁻⁻	9,580 ⁺⁺	2,560 ⁻⁻	55,740 ⁺⁺	40,280 ⁺⁺	55,740 ⁺⁺
	20	2,920 ⁻⁻	10,000 ⁺⁺	2,920 ⁻⁻	57,780 ⁺⁺	40,660 ⁺⁺	57,780 ⁺⁺
20	4	2,940 ⁻⁻	9,420 ⁺⁺	3,120 ⁻⁻	67,840 ⁺⁺	59,140 ⁺⁺	67,840 ⁺⁺
	10	2,780 ⁻⁻	9,740 ⁺⁺	2,780 ⁻⁻	68,500 ⁺⁺	58,840 ⁺⁺	68,500 ⁺⁺
	20	2,840 ⁻⁻	10,000 ⁺⁺	2,840 ⁻⁻	68,760 ⁺⁺	58,700 ⁺⁺	68,760 ⁺⁺
40	4	3,120 ⁻⁻	9,220 ⁺⁺	3,120 ⁻⁻	65,900 ⁺⁺	61,280 ⁺⁺	65,900 ⁺⁺
	10	3,280 ⁻⁻	10,200 ⁺⁺	3,280 ⁻⁻	65,820 ⁺⁺	60,760 ⁺⁺	65,820 ⁺⁺
	20	3,100 ⁻⁻	9,960 ⁺⁺	3,100 ⁻⁻	66,760 ⁺⁺	61,620 ⁺⁺	66,760 ⁺⁺
100	4	2,940 ⁻⁻	9,720 ⁺⁺	2,940 ⁻⁻	58,840 ⁺⁺	56,400 ⁺⁺	58,840 ⁺⁺
	10	3,440 ⁻⁻	9,760 ⁺⁺	3,440 ⁻⁻	59,900 ⁺⁺	57,700 ⁺⁺	59,900 ⁺⁺
	20	3,320 ⁻⁻	10,040 ⁺⁺	3,320 ⁻⁻	60,080 ⁺⁺	57,640 ⁺⁺	60,080 ⁺⁺

Tabela 13 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 5\%$ de probabilidade e $\delta = 32\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Trat	Rep	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	2,300 ⁻⁻	9,740 ⁺⁺	2,300 ⁻⁻	31,780 ⁺⁺	17,320 ⁺⁺	31,780 ⁺⁺
	10	2,460 ⁻⁻	9,520 ⁺⁺	2,460 ⁻⁻	33,840 ⁺⁺	18,120 ⁺⁺	33,840 ⁺⁺
	20	2,700 ⁻⁻	9,780 ⁺⁺	2,700 ⁻⁻	36,000 ⁺⁺	19,900 ⁺⁺	36,000 ⁺⁺
10	4	2,800 ⁻⁻	9,860 ⁺⁺	2,800 ⁻⁻	54,340 ⁺⁺	38,200 ⁺⁺	54,340
	10	2,720 ⁻⁻	9,520 ⁺⁺	2,720 ⁻⁻	57,200 ⁺⁺	39,780 ⁺⁺	57,200 ⁺⁺
	20	2,780 ⁻⁻	1,040 ⁺⁺	2,780 ⁻⁻	56,560 ⁺⁺	39,480 ⁺⁺	56,560 ⁺⁺
20	4	2,740 ⁻⁻	8,520 ⁺⁺	2,740 ⁻⁻	67,440 ⁺⁺	58,200 ⁺⁺	67,440 ⁺⁺
	10	2,720 ⁻⁻	9,520 ⁺⁺	2,720 ⁻⁻	67,300 ⁺⁺	57,140 ⁺⁺	67,300 ⁺⁺
	20	3,180 ⁻⁻	10,520 ⁺⁺	3,180 ⁻⁻	68,520 ⁺⁺	59,180 ⁺⁺	68,520 ⁺⁺
40	4	2,940 ⁻⁻	9,380 ⁺⁺	2,940 ⁻⁻	65,600 ⁺⁺	60,920 ⁺⁺	65,600 ⁻⁻
	10	3,280 ⁻⁻	1,010 ⁺⁺	3,280 ⁻⁻	66,440 ⁺⁺	61,980 ⁺⁺	66,440 ⁺⁺
	20	2,980 ⁻⁻	9,660 ⁺⁺	2,980 ⁻⁻	66,320 ⁺⁺	61,140 ⁺⁺	66,320 ⁺⁺
100	4	2,700 ⁻⁻	9,560 ⁺⁺	2,700 ⁻⁻	59,360 ⁺⁺	56,920 ⁺⁺	59,360 ⁺⁺
	10	3,080 ⁻⁻	9,720 ⁺⁺	3,080 ⁻⁻	58,740 ⁺⁺	56,020 ⁺⁺	58,740 ⁺⁺
	20	3,140 ⁻⁻	10,260 ⁺⁺	3,140 ⁻⁻	59,620 ⁺⁺	57,440 ⁺⁺	59,620 ⁺⁺

Tabela 14 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 1\%$ de probabilidade e $\delta = 1\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Trat	Rep	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	0,580 ⁻⁻	0,760	2,520 ⁺⁺	0,840	0,640 ⁻⁻	1,240
	10	0,360 ⁻⁻	0,500 ⁻⁻	2,280 ⁺⁺	0,580 ⁻⁻	0,520 ⁻⁻	0,900
	20	0,580 ⁻⁻	0,640 ⁻⁻	2,640 ⁺⁺	0,720	0,580 ⁻⁻	0,840
10	4	0,640 ⁻⁻	0,760	3,020 ⁺⁺	1,660 ⁺⁺	1,560 ⁺⁺	2,460 ⁺⁺
	10	0,500 ⁻⁻	0,620 ⁻⁻	3,580 ⁺⁺	1,440 ⁺⁺	1,280	2,000 ⁺⁺
	20	0,480 ⁻⁻	0,560 ⁻⁻	3,440 ⁺⁺	1,360	1,280	1,880 ⁺⁺
20	4	0,640 ⁻⁻	0,700	3,560 ⁺⁺	1,260	1,220	1,720 ⁻⁻
	10	0,560 ⁻⁻	0,600 ⁻⁻	3,920 ⁺⁺	1,340	1,280	1,620 ⁺⁺
	20	0,660 ⁻⁻	0,740	3,880 ⁺⁺	1,340	1,300	1,660 ⁺⁺
40	4	0,600 ⁻⁻	0,660 ⁻⁻	4,300 ⁺⁺	0,880	0,880	0,980
	10	0,700	0,740	4,660 ⁺⁺	0,900	0,900	0,920
	20	0,620 ⁻⁻	0,700	4,380 ⁺⁺	0,940	0,940	1,080
100	4	0,640 ⁻⁻	0,640 ⁻⁻	4,700 ⁺⁺	1,060	1,060	1,060
	10	0,840	0,840	5,640 ⁺⁺	0,800	0,800	0,800
	20	0,680	0,680	5,000 ⁺⁺	0,940	0,940	0,940

Tabela 15 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 1\%$ de probabilidade e $\delta = 2\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Trat	Rep	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	0,320 ⁻⁻	0,880	6,320 ⁺⁺	2,020 ⁺⁺	1,440 ⁺⁺	2,700 ⁺⁺
	10	0,560 ⁻⁻	1,020	8,480 ⁺⁺	1,900 ⁺⁺	1,080	2,760 ⁺⁺
	20	0,420 ⁻⁻	0,800	9,480 ⁺⁺	2,260 ⁺⁺	1,340	3,140 ⁺⁺
10	4	0,440 ⁻⁻	0,860	12,180 ⁺⁺	2,800 ⁺⁺	2,180 ⁺⁺	7,060 ⁺⁺
	10	0,600 ⁻⁻	1,000	16,220 ⁺⁺	2,900 ⁺⁺	2,140 ⁺⁺	7,400 ⁺⁺
	20	0,640 ⁻⁻	0,980	9,480 ⁺⁺	2,560 ⁺⁺	2,000 ⁺⁺	7,280 ⁺⁺
20	4	0,580 ⁻⁻	0,760	18,300 ⁺⁺	2,320 ⁺⁺	0,720	5,330 ⁺⁺
	10	0,640 ⁻⁻	0,900	21,160 ⁺⁺	2,740 ⁺⁺	2,320 ⁺⁺	8,680 ⁺⁺
	20	0,660 ⁻⁻	0,840	21,840 ⁺⁺	2,880 ⁺⁺	2,620 ⁺⁺	8,040 ⁺⁺
40	4	0,600 ⁻⁻	0,700	24,960 ⁺⁺	2,140 ⁺⁺	2,040 ⁺⁺	4,780 ⁺⁺
	10	0,640 ⁻⁻	0,740	27,220 ⁺⁺	1,920 ⁺⁺	1,840 ⁺⁺	4,340 ⁺⁺
	20	0,560 ⁻⁻	0,720	28,860 ⁺⁺	2,240 ⁺⁺	2,220 ⁺⁺	4,860 ⁺⁺
100	4	0,440 ⁻⁻	0,520 ⁻⁻	33,600 ⁺⁺	1,320	1,320	1,600 ⁺⁺
	10	0,800	0,840	36,780 ⁺⁺	1,140	1,140	1,300
	20	0,680	0,740	38,400 ⁺⁺	1,280	1,280	1,460 ⁺⁺

Tabela 16 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 1\%$ de probabilidade e $\delta = 4\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Trat	Rep	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	0,420 ⁻⁻	1,660 ⁺⁺	12,200 ⁺⁺	6,400 ⁺⁺	2,760 ⁺⁺	8,460 ⁺⁺
	10	0,560 ⁻⁻	1,880 ⁺⁺	18,120 ⁺⁺	6,340 ⁺⁺	1,960 ⁺⁺	8,380 ⁺⁺
	20	0,420 ⁻⁻	2,000 ⁺⁺	18,240 ⁺⁺	6,900 ⁺⁺	2,140 ⁺⁺	9,420 ⁺⁺
10	4	0,560 ⁻⁻	1,580 ⁺⁺	28,300 ⁺⁺	7,180 ⁺⁺	3,360 ⁺⁺	25,960 ⁺⁺
	10	0,520 ⁻⁻	1,880 ⁺⁺	34,200 ⁺⁺	7,160 ⁺⁺	3,020 ⁺⁺	28,580 ⁺⁺
	20	0,400 ⁻⁻	1,880 ⁺⁺	36,220 ⁺⁺	7,240 ⁺⁺	3,140 ⁺⁺	29,040 ⁺⁺
20	4	0,740	1,820 ⁺⁺	49,360 ⁺⁺	5,920 ⁺⁺	2,840 ⁺⁺	47,840 ⁺⁺
	10	0,640 ⁻⁻	1,620 ⁺⁺	54,100 ⁺⁺	5,980 ⁺⁺	2,780 ⁺⁺	49,400 ⁺⁺
	20	0,600 ⁻⁻	1,620 ⁺⁺	55,280 ⁺⁺	5,380 ⁺⁺	2,400 ⁺⁺	49,700 ⁺⁺
40	4	0,580 ⁻⁻	1,540 ⁺⁺	74,820 ⁺⁺	3,140 ⁺⁺	1,740 ⁺⁺	50,440 ⁺⁺
	10	0,460 ⁻⁻	1,820 ⁺⁺	77,300 ⁺⁺	3,500 ⁺⁺	2,440 ⁺⁺	49,640 ⁺⁺
	20	0,700	1,600 ⁺⁺	76,940 ⁺⁺	3,440 ⁺⁺	2,160 ⁺⁺	50,960 ⁺⁺
100	4	0,460 ⁻⁻	1,420	96,260 ⁺⁺	1,280	0,940	27,280 ⁺⁺
	10	0,520 ⁻⁻	1,260	96,800 ⁺⁺	1,600 ⁺⁺	1,260	27,640 ⁺⁺
	20	0,600 ⁻⁻	1,540 ⁺⁺	97,120 ⁺⁺	1,700 ⁺⁺	1,300	28,480 ⁺⁺

Tabela 17 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 1\%$ de probabilidade e $\delta = 8\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Trat	Rep	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	0,500 ⁻⁻	1,780 ⁺⁺	0,880	11,500 ⁺⁺	4,380 ⁺⁺	11,500 ⁺⁺
	10	0,440 ⁻⁻	2,340 ⁺⁺	0,800	13,380 ⁺⁺	4,760 ⁺⁺	13,400 ⁺⁺
	20	0,400 ⁻⁻	1,940 ⁺⁺	0,600 ⁻⁻	14,260 ⁺⁺	5,000 ⁺⁺	14,260 ⁺⁺
10	4	0,480 ⁻⁻	1,780 ⁺⁺	0,680	26,740 ⁺⁺	15,420 ⁺⁺	26,820 ⁺⁺
	10	0,620 ⁻⁻	1,920 ⁺⁺	0,740	28,200 ⁺⁺	16,320 ⁺⁺	28,300 ⁺⁺
	20	0,480 ⁻⁻	2,000 ⁺⁺	0,600 ⁻⁻	28,660 ⁺⁺	16,720 ⁺⁺	28,740 ⁺⁺
20	4	0,400 ⁻⁻	1,580 ⁺⁺	0,460 ⁻⁻	47,060 ⁺⁺	38,380 ⁺⁺	47,200 ⁺⁺
	10	0,560 ⁻⁻	2,300 ⁺⁺	0,820	46,420 ⁺⁺	37,940 ⁺⁺	46,540 ⁺⁺
	20	0,620 ⁻⁻	2,180 ⁺⁺	0,760	46,480 ⁺⁺	37,900 ⁺⁺	46,560 ⁺⁺
40	4	0,580 ⁻⁻	1,720 ⁺⁺	0,840	51,980 ⁺⁺	48,960 ⁺⁺	52,360 ⁺⁺
	10	0,580 ⁻⁻	2,060 ⁺⁺	0,840	52,300 ⁺⁺	49,300 ⁺⁺	52,000 ⁺⁺
	20	0,680	2,380 ⁺⁺	0,840	52,680 ⁺⁺	49,940 ⁺⁺	53,040 ⁺⁺
100	4	0,420 ⁻⁻	1,760 ⁺⁺	1,000	50,980 ⁺⁺	50,000 ⁺⁺	51,720 ⁺⁺
	10	0,620 ⁻⁻	2,040 ⁺⁺	1,060	51,460 ⁺⁺	50,400 ⁺⁺	52,360 ⁺⁺
	20	0,560 ⁻⁻	1,980 ⁺⁺	1,080	52,620 ⁺⁺	51,460 ⁺⁺	53,540 ⁺⁺

Tabela 18 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 1\%$ de probabilidade e $\delta = 16\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Trat	Rep	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	0,440 ⁻⁻	1,920 ⁺⁺	0,440 ⁻⁻	11,940 ⁺⁺	4,420 ⁺⁺	11,940 ⁺⁺
	10	0,600 ⁻⁻	2,240 ⁺⁺	0,600 ⁻⁻	13,740 ⁺⁺	4,780 ⁺⁺	13,740 ⁺⁺
	20	0,380 ⁻⁻	2,220 ⁺⁺	0,380 ⁻⁻	13,740 ⁺⁺	5,020 ⁺⁺	13,740 ⁺⁺
10	4	0,580 ⁻⁻	2,340 ⁺⁺	0,580 ⁻⁻	26,680 ⁺⁺	15,900 ⁺⁺	26,680 ⁺⁺
	10	0,660 ⁻⁻	2,240 ⁺⁺	0,660 ⁻⁻	28,760 ⁺⁺	16,300 ⁺⁺	28,760 ⁺⁺
	20	0,320 ⁻⁻	2,100 ⁺⁺	0,320 ⁻⁻	28,560 ⁺⁺	16,580 ⁺⁺	28,560 ⁺⁺
20	4	0,740	2,420 ⁺⁺	0,740	45,960 ⁺⁺	35,800 ⁺⁺	45,960 ⁺⁺
	10	0,600 ⁻⁻	1,820 ⁺⁺	0,600 ⁻⁻	45,920 ⁺⁺	37,640 ⁺⁺	45,920 ⁺⁺
	20	0,440 ⁻⁻	1,820 ⁺⁺	0,440 ⁻⁻	47,140 ⁺⁺	38,720 ⁺⁺	47,140 ⁺⁺
40	4	0,480 ⁻⁻	1,940 ⁺⁺	0,480 ⁻⁻	51,760 ⁺⁺	48,760 ⁺⁺	51,760 ⁺⁺
	10	0,740	2,140 ⁺⁺	0,740	53,140 ⁺⁺	50,340 ⁺⁺	53,140 ⁺⁺
	20	0,400 ⁻⁻	1,800 ⁺⁺	0,400 ⁻⁻	52,680 ⁺⁺	49,780 ⁺⁺	52,680 ⁺⁺
100	4	0,540 ⁻⁻	1,840 ⁺⁺	0,540 ⁻⁻	50,600 ⁺⁺	49,740 ⁺⁺	50,600 ⁺⁺
	10	0,460 ⁻⁻	1,980 ⁺⁺	0,460 ⁻⁻	52,140 ⁺⁺	51,340 ⁺⁺	52,140 ⁺⁺
	20	0,580 ⁻⁻	1,960 ⁺⁺	0,580 ⁻⁻	52,480 ⁺⁺	51,660 ⁺⁺	52,480 ⁺⁺

Tabela 19 Taxa de erro tipo I por experimento, em porcentagem, dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância $\alpha = 1\%$ de probabilidade e $\delta = 32\sigma_{\bar{Y}}$, avaliado pelo teste binomial exato com coeficiente de confiança de 99% de probabilidade.

Trat	Rep	Testes avaliados					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	0,500 ⁻⁻	2,080 ⁺⁺	0,500 ⁻⁻	12,480 ⁺⁺	4,860 ⁺⁺	12,480 ⁺⁺
	10	0,440 ⁻⁻	1,840 ⁺⁺	0,440 ⁻⁻	12,980 ⁺⁺	5,160 ⁺⁺	12,980 ⁺⁺
	20	0,440 ⁻⁻	1,980 ⁺⁺	0,440 ⁻⁻	14,620 ⁺⁺	5,280 ⁺⁺	14,620 ⁺⁺
10	4	0,580 ⁻⁻	1,920 ⁺⁺	0,580 ⁻⁻	27,040 ⁺⁺	15,480 ⁺⁺	27,040 ⁺⁺
	10	0,420 ⁻⁻	2,180 ⁺⁺	0,420 ⁻⁻	27,340 ⁺⁺	15,660 ⁺⁺	27,340 ⁺⁺
	20	0,460 ⁻⁻	1,700 ⁺⁺	0,460 ⁻⁻	29,600 ⁺⁺	17,580 ⁺⁺	29,600 ⁺⁺
20	4	0,220 ⁻⁻	1,600 ⁺⁺	0,220 ⁻⁻	46,540 ⁺⁺	38,340 ⁺⁺	46,540 ⁺⁺
	10	0,520 ⁻⁻	1,940 ⁺⁺	0,520 ⁻⁻	47,700 ⁺⁺	39,460 ⁺⁺	47,700 ⁺⁺
	20	0,500 ⁻⁻	1,860 ⁺⁺	0,500 ⁻⁻	47,460 ⁺⁺	38,840 ⁺⁺	47,460 ⁺⁺
40	4	0,460 ⁻⁻	1,700 ⁺⁺	0,460 ⁻⁻	53,240 ⁺⁺	49,940 ⁺⁺	53,240 ⁺⁺
	10	0,580 ⁻⁻	2,400 ⁺⁺	0,580 ⁻⁻	53,120 ⁺⁺	49,720 ⁺⁺	53,120 ⁺⁺
	20	0,520 ⁻⁻	1,900 ⁺⁺	0,520 ⁻⁻	53,300 ⁺⁺	50,420 ⁺⁺	53,300 ⁺⁺
100	4	0,540 ⁻⁻	1,880 ⁺⁺	0,540 ⁻⁻	52,340 ⁺⁺	51,440 ⁺⁺	52,340 ⁺⁺
	10	0,560 ⁻⁻	2,080 ⁺⁺	0,560 ⁻⁻	52,980 ⁺⁺	51,900 ⁺⁺	52,980 ⁺⁺
	20	0,600 ⁻⁻	1,880 ⁺⁺	0,600 ⁻⁻	52,160 ⁺⁺	51,440 ⁺⁺	52,160 ⁺⁺

APÊNDICE B - Poder sob H_1 completa

Tabela 20 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com erro padrão ($1\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 1% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	0,510	2,270	19,740	-	-	-
	10	0,180	1,080	14,270	92,110	-	-
	20	0,050	0,390	9,370	90,920	100,000	-
	40	0,020	0,170	5,860	87,630	100,000	100,000
	100	0,000	0,040	2,810	80,850	100,000	100,000
SNK	5	1,120	3,530	20,900	-	-	-
	10	0,900	3,470	24,150	94,630	-	-
	20	0,780	3,410	27,670	97,170	100,000	-
	40	0,720	3,500	29,480	98,070	100,000	100,000
	100	0,700	3,480	30,670	98,450	100,000	100,000
MGR	5	7,050	11,670	24,220	-	-	-
	10	19,870	33,070	62,930	98,050	-	-
	20	21,360	35,930	66,960	98,660	100,000	-
	40	21,740	36,330	66,920	98,380	100,000	100,000
	100	21,130	35,540	65,490	97,570	100,000	100,000
TM	5	2,380	4,990	11,420	-	-	-
	10	3,080	7,230	21,190	55,110	-	-
	20	2,140	5,140	14,950	36,130	74,250	-
	40	1,530	3,410	9,340	21,120	41,680	80,600
	100	0,910	1,930	4,890	10,440	20,040	37,300
SNKM	5	1,150	3,390	10,460	-	-	-
	10	1,230	3,810	15,700	53,570	-	-
	20	0,610	2,090	8,890	31,470	73,480	-
	40	0,290	0,990	4,540	16,350	38,970	80,170
	100	0,100	0,360	1,780	6,830	17,070	35,360
MGM	5	3,960	7,790	13,160	-	-	-
	10	15,260	26,200	47,850	63,200	-	-
	20	24,890	40,120	66,170	80,800	86,070	-
	40	32,180	49,670	77,480	89,490	92,650	94,650
	100	38,160	56,990	85,160	95,150	96,910	97,550

Tabela 21 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com erro padrão ($1\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 5% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	2,040	7,630	41,640	-	-	-
	10	0,700	3,580	30,960	98,130	-	-
	20	0,220	1,510	20,760	96,730	100,000	-
	40	0,070	0,620	13,080	94,690	100,000	100,000
	100	0,020	0,190	6,690	90,060	100,000	100,000
SNK	5	5,020	12,570	44,980	-	-	-
	10	4,150	11,930	48,350	99,090	-	-
	20	3,840	11,920	51,180	99,470	100,000	-
	40	3,590	11,760	52,590	99,620	100,000	100,000
	100	3,550	11,910	53,710	99,700	100,000	100,000
MGR	5	15,540	25,620	50,800	-	-	-
	10	23,690	39,270	72,520	99,620	-	-
	20	24,190	40,020	72,870	99,560	100,000	-
	40	23,630	39,280	71,380	99,330	100,000	100,000
	100	22,650	37,790	68,850	98,710	100,000	100,000
TM	5	9,390	17,200	32,960	-	-	-
	10	11,120	20,980	43,480	72,460	-	-
	20	9,250	17,080	34,290	55,690	84,960	-
	40	7,380	12,800	24,330	37,660	57,240	88,100
	100	5,190	8,520	14,990	22,010	32,310	49,620
SNKM	5	6,270	13,390	31,340	-	-	-
	10	5,070	12,560	34,900	71,480	-	-
	20	3,190	8,000	23,420	50,290	84,470	-
	40	1,830	4,640	14,140	30,880	54,120	87,850
	100	0,840	2,230	7,060	15,760	27,880	47,030
MGM	5	11,970	21,210	34,940	-	-	-
	10	25,910	41,000	65,540	78,030	-	-
	20	36,610	54,610	80,600	89,790	92,090	-
	40	43,360	62,150	87,280	94,480	95,910	96,480
	100	49,360	68,600	92,620	98,010	98,720	98,980

Tabela 22 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com erro padrão ($1\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 1% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	0,510	2,810	28,580	-	-	-
	10	0,130	1,280	18,520	96,670	-	-
	20	0,050	0,450	11,130	93,950	100,000	-
	40	0,020	0,200	6,590	89,950	100,000	100,000
	100	0,000	0,050	3,040	82,470	100,000	100,000
SNK	5	1,130	4,370	30,320	-	-	-
	10	0,860	3,910	29,900	97,760	-	-
	20	0,770	3,630	30,180	98,300	100,000	-
	40	0,710	3,550	31,000	98,490	100,000	100,000
	100	0,700	3,560	31,310	98,630	100,000	100,000
MGR	5	10,480	17,010	36,320	-	-	-
	10	21,310	35,570	67,350	99,330	-	-
	20	22,000	36,900	68,520	99,160	100,000	-
	40	22,060	36,770	67,780	98,630	100,000	100,000
	100	21,390	35,830	65,930	97,770	100,000	100,000
TM	5	2,880	6,090	14,100	-	-	-
	10	3,170	7,550	22,070	56,390	-	-
	20	2,280	5,310	15,330	36,510	73,640	-
	40	1,460	3,410	9,340	21,120	41,680	80,600
	100	0,920	1,940	4,930	10,520	20,030	37,160
SNKM	5	1,540	4,210	12,880	-	-	-
	10	1,330	4,040	16,360	54,760	-	-
	20	0,620	2,150	9,080	31,700	72,970	-
	40	0,290	0,990	4,520	16,410	39,530	79,870
	100	0,110	0,370	1,790	6,920	17,100	35,200
MGM	5	4,730	9,490	16,300	-	-	-
	10	15,460	26,530	48,320	64,280	-	-
	20	25,360	40,710	66,700	80,530	85,100	-
	40	32,130	49,670	77,480	89,490	92,650	94,110
	100	38,170	57,150	85,140	94,960	96,710	97,380

Tabela 23 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com erro padrão ($1\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 5% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	2,290	8,830	49,700	-	-	-
	10	0,720	3,910	34,390	98,840	-	-
	20	0,230	1,520	22,530	97,770	100,000	-
	40	0,070	0,660	13,980	95,690	100,000	100,000
	100	0,020	0,190	6,910	90,880	100,000	100,000
SNK	5	5,430	14,540	53,600	-	-	-
	10	4,340	12,380	52,650	99,380	-	-
	20	3,800	12,240	53,580	99,660	100,000	-
	40	3,680	12,110	53,950	99,720	100,000	100,000
	100	3,560	12,030	54,310	99,730	100,000	100,000
MGR	5	18,340	31,050	61,080	-	-	-
	10	24,510	40,200	74,320	99,780	-	-
	20	24,590	40,710	73,890	99,720	100,000	-
	40	23,890	39,620	71,950	99,440	100,000	100,000
	100	22,870	38,060	69,280	98,840	100,000	100,000
TM	5	9,430	17,650	34,200	-	-	-
	10	11,290	22,040	45,880	75,790	-	-
	20	9,570	17,640	35,220	56,650	85,530	-
	40	7,190	12,720	24,090	37,200	56,650	88,050
	100	5,110	8,370	14,750	21,740	32,140	49,370
SNKM	5	6,070	13,690	32,240	-	-	-
	10	5,150	13,130	37,120	74,670	-	-
	20	3,270	8,210	23,970	51,140	84,940	-
	40	1,750	4,580	13,980	30,590	53,540	87,780
	100	0,820	2,150	6,830	15,450	27,630	46,820
MGM	5	12,110	22,090	36,700	-	-	-
	10	26,860	42,740	68,500	81,330	-	-
	20	37,080	55,030	81,130	90,430	92,780	-
	40	43,660	62,490	87,750	94,770	96,100	96,760
	100	49,400	68,530	92,410	97,780	98,510	98,800

Tabela 24 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com erro padrão ($1\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 1% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	0,510	3,070	31,220	-	-	-
	10	0,150	1,200	19,330	97,190	-	-
	20	0,050	0,480	11,700	94,560	100,000	-
	40	0,010	0,190	6,740	9,550	100,000	100,000
	100	0,000	0,050	3,060	82,940	100,000	100,000
SNK	5	1,080	4,810	33,280	-	-	-
	10	0,910	3,770	30,600	98,190	-	-
	20	0,730	3,750	31,240	98,490	100,000	-
	40	0,740	3,660	31,320	98,610	100,000	100,000
	100	0,690	3,590	31,570	98,680	100,000	100,000
MGR	5	11,340	18,710	39,400	-	-	-
	10	21,610	36,000	67,870	99,420	-	-
	20	22,300	37,110	69,010	99,240	100,000	-
	40	22,170	36,990	68,120	98,730	100,000	100,000
	100	21,330	35,820	66,000	97,860	100,000	100,000
TM	5	2,580	5,510	13,460	-	-	-
	10	3,160	7,640	22,470	56,860	-	-
	20	2,210	5,310	15,460	37,060	74,730	-
	40	1,560	3,520	9,680	21,790	43,150	80,580
	100	0,930	1,920	4,890	10,450	20,010	37,320
SNKM	5	1,350	3,850	12,200	-	-	-
	10	1,250	3,990	16,660	55,410	-	-
	20	0,620	2,020	9,080	32,030	74,000	-
	40	0,310	1,000	4,620	16,760	40,270	80,160
	100	0,100	0,360	1,750	6,840	17,050	35,260
MGM	5	4,370	8,780	15,380	-	-	-
	10	15,810	27,280	49,510	65,280	-	-
	20	25,530	40,620	66,670	81,150	85,880	-
	40	32,660	50,290	78,340	90,160	92,830	94,140
	100	38,230	57,100	95,230	95,140	96,880	97,520

Tabela 25 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com erro padrão ($1\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 5% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	2,150	8,950	50,060	-	-	-
	10	0,780	3,890	35,300	99,290	-	-
	20	0,250	1,670	23,330	98,070	100,000	-
	40	0,080	0,650	14,160	96,020	100,000	100,000
	100	0,010	0,190	6,970	91,060	100,000	100,000
SNK	5	5,110	13,990	54,180	-	-	-
	10	4,280	12,670	53,840	99,730	-	-
	20	3,950	12,470	54,080	99,690	100,000	-
	40	3,730	12,180	54,290	99,720	100,000	100,000
	100	3,580	12,030	54,460	99,740	100,000	100,000
MGR	5	18,690	30,920	61,640	-	-	-
	10	24,490	40,750	75,100	99,890	-	-
	20	24,510	40,550	73,970	99,720	100,000	-
	40	23,810	39,750	71,980	99,500	100,000	100,000
	100	22,880	38,020	69,240	98,850	100,000	100,000
TM	5	9,140	17,490	34,660	-	-	-
	10	10,920	21,350	44,610	75,190	-	-
	20	9,390	17,600	35,310	56,980	85,580	-
	40	7,430	13,040	24,850	38,410	58,120	89,080
	100	4,890	8,080	14,330	21,250	31,500	48,510
SNKM	5	5,870	13,510	32,720	-	-	-
	10	4,980	12,350	36,120	74,000	-	-
	20	3,220	8,230	24,090	51,590	85,140	-
	40	1,780	4,650	14,180	31,270	54,780	88,780
	100	0,820	2,140	6,830	15,340	27,220	45,960
MGM	5	12,180	22,250	36,920	-	-	-
	10	27,290	43,280	68,190	80,670	-	-
	20	37,060	55,360	81,280	90,310	92,550	-
	40	44,110	63,170	88,710	95,810	97,030	97,480
	100	49,400	68,460	92,450	97,580	98,580	98,830

Tabela 26 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com dois erros padrões ($2\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 1% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
TR	5	-	2,400	19,260	92,740	-	-
	10	-	1,000	14,180	92,340	100,000	-
	20	-	0,430	9,670	90,850	100,000	100,000
	40	-	0,170	5,750	87,370	100,000	100,000
	100	-	0,050	2,830	80,870	100,000	100,000
SNKR	5	-	7,470	31,510	93,440	-	-
	10	-	8,220	38,420	97,680	100,000	-
	20	-	8,850	42,260	98,780	100,000	100,000
	40	-	9,040	44,210	99,170	100,000	100,000
	100	-	9,310	45,450	99,350	100,000	100,000
MGR	5	-	28,520	55,910	94,040	-	-
	10	-	32,200	61,850	97,420	100,000	-
	20	-	32,670	62,240	96,980	100,000	100,000
	40	-	32,410	61,590	96,230	100,000	100,000
	100	-	31,510	59,830	94,890	100,000	100,000
TM	5	-	15,990	33,570	59,560	-	-
	10	-	22,480	46,860	69,820	88,660	-
	20	-	25,090	48,380	64,140	77,250	94,490
	40	-	24,090	43,030	53,410	62,160	76,320
	100	-	21,170	35,020	41,130	46,710	55,970
SNKM	5	-	7,970	24,490	58,820	-	-
	10	-	7,240	26,490	61,180	88,300	-
	20	-	5,810	22,240	50,020	72,670	94,290
	40	-	3,960	16,210	37,240	53,800	62,160
	100	-	2,600	10,760	24,820	35,910	49,080
MGM	5	-	21,810	44,120	60,260	-	-
	10	-	39,310	69,780	87,200	91,190	-
	20	-	49,390	81,730	95,110	97,050	97,680
	40	-	55,580	87,490	97,890	98,720	98,980
	100	-	60,960	91,720	99,230	99,660	99,770

Tabela 27 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com dois erros padrões ($2\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 5% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	7,660	41,930	98,880	-	-
	10	-	3,350	29,840	98,060	100,000	-
	20	-	1,610	20,740	96,770	100,000	100,000
	40	-	0,680	13,440	94,820	100,000	100,000
	100	-	0,190	6,690	89,980	100,000	100,000
SNK	5	-	22,500	59,600	99,220	-	-
	10	-	23,440	64,530	99,710	100,000	-
	20	-	24,780	67,310	99,850	100,000	100,000
	40	-	25,170	68,410	99,880	100,000	100,000
	100	-	25,330	69,110	99,890	100,000	100,000
MGR	5	-	35,720	68,320	99,320	-	-
	10	-	36,480	68,790	99,280	100,000	-
	20	-	35,800	67,260	98,900	100,000	100,000
	40	-	34,890	65,420	98,240	100,000	100,000
	100	-	33,340	62,840	96,810	100,000	100,000
TM	5	-	37,870	65,310	83,800	-	-
	10	-	49,420	79,750	91,370	96,320	-
	20	-	53,030	78,520	85,800	90,500	97,760
	40	-	50,520	72,430	77,680	81,710	89,240
	100	-	46,550	63,780	67,550	70,850	76,770
SNKM	5	-	23,870	55,980	83,560	-	-
	10	-	23,620	59,560	86,980	96,160	-
	20	-	20,680	54,020	78,010	88,140	97,690
	40	-	17,300	46,170	67,090	76,470	87,550
	100	-	13,730	37,030	54,000	61,810	71,330
MGM	5	-	40,910	69,910	83,880	-	-
	10	-	56,000	86,610	95,970	97,180	-
	20	-	63,860	92,710	98,670	99,070	99,220
	40	-	68,540	95,080	99,430	99,580	99,640
	100	-	72,480	96,690	99,820	99,820	99,890

Tabela 28 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com dois erros padrões ($2\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 1% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	2,980	27,840	98,120	-	-
	10	-	1,150	18,390	96,720	100,000	-
	20	-	0,450	11,060	93,880	100,000	100,000
	40	-	0,200	6,610	90,060	100,000	100,000
	100	-	0,050	3,060	82,460	100,000	100,000
SNK	5	-	9,170	41,110	98,540	-	-
	10	-	9,250	44,090	99,090	100,000	-
	20	-	9,110	44,950	99,310	100,000	100,000
	40	-	9,420	45,700	99,350	100,000	100,000
	100	-	9,510	46,020	99,380	100,000	100,000
MGR	5	-	31,820	62,620	99,000	-	-
	10	-	33,530	64,230	98,690	100,000	-
	20	-	33,450	63,700	97,890	100,000	100,000
	40	-	32,820	62,250	96,800	100,000	100,000
	100	-	31,690	60,140	95,110	100,000	100,000
TM	5	-	19,310	40,690	68,680	-	-
	10	-	24,690	51,210	74,150	90,550	-
	20	-	26,370	50,500	65,760	77,970	94,740
	40	-	24,780	44,260	54,570	63,280	77,320
	100	-	21,770	35,850	42,020	47,620	56,810
SNKM	5	-	9,550	30,080	68,200	-	-
	10	-	7,990	28,790	65,150	90,210	-
	20	-	5,800	82,980	95,950	97,460	97,920
	40	-	4,000	16,630	38,020	54,670	73,960
	100	-	2,650	10,960	25,350	36,670	49,850
MGM	5	-	25,340	50,690	69,180	-	-
	10	-	41,210	72,500	89,490	92,690	-
	20	-	50,380	82,980	95,950	97,460	97,920
	40	-	56,130	88,140	98,140	98,970	99,250
	100	-	61,010	91,700	99,110	99,530	99,660

Tabela 29 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com dois erros padrões ($2\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 5% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	8,850	50,280	99,660	-	-
	10	-	3,860	34,330	99,090	100,000	-
	20	-	1,620	22,700	97,850	100,000	100,000
	40	-	0,650	13,970	95,670	100,000	100,000
	100	-	0,190	6,950	90,840	100,000	100,000
SNK	5	-	25,480	66,670	99,900	-	-
	10	-	25,330	68,240	99,840	100,000	-
	20	-	25,440	68,850	99,890	100,000	100,000
	40	-	25,250	69,030	99,890	100,000	100,000
	100	-	25,430	69,390	99,900	100,000	100,000
MGR	5	-	37,590	71,730	99,920	-	-
	10	-	37,340	70,340	99,680	100,000	-
	20	-	36,260	67,980	99,170	100,000	100,000
	40	-	35,160	65,940	98,420	100,000	100,000
	100	-	33,510	63,170	97,040	100,000	100,000
TM	5	-	40,530	70,010	86,860	-	-
	10	-	51,000	81,730	92,490	96,960	-
	20	-	52,710	79,290	86,490	91,090	98,100
	40	-	50,900	72,900	78,070	82,060	89,270
	100	-	45,840	62,680	66,330	69,590	75,450
SNKM	5	-	25,300	60,090	86,660	-	-
	10	-	24,440	62,050	88,970	96,860	-
	20	-	20,940	55,130	79,330	89,030	98,000
	40	-	17,410	46,740	67,690	76,730	87,360
	100	-	13,470	36,540	53,290	61,050	70,440
MGM	5	-	42,990	73,040	86,960	-	-
	10	-	57,240	87,960	96,920	97,720	-
	20	-	64,360	92,870	98,850	99,180	99,320
	40	-	68,830	95,260	99,490	99,650	99,730
	100	-	72,640	96,680	99,740	99,810	99,850

Tabela 30 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com dois erros padrões ($2\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 1% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	3,050	31,320	98,860	-	-
	10	-	1,210	19,410	97,140	100,000	-
	20	-	0,490	11,850	94,650	100,000	100,000
	40	-	0,200	6,820	90,520	100,000	100,000
	100	-	0,050	3,090	82,850	100,000	100,000
SNK	5	-	9,660	44,640	99,220	-	-
	10	-	9,430	44,870	99,240	100,000	-
	20	-	9,550	45,690	99,310	100,000	100,000
	40	-	9,440	45,810	99,380	100,000	100,000
	100	-	9,430	46,070	99,390	100,000	100,000
MGR	5	-	32,480	63,830	99,520	-	-
	10	-	33,830	64,640	98,880	100,000	-
	20	-	33,660	63,870	97,980	100,000	100,000
	40	-	32,930	62,400	96,900	100,000	100,000
	100	-	31,710	60,710	95,190	100,000	100,000
TM	5	-	19,470	41,530	68,960	-	-
	10	-	25,310	52,670	75,470	91,530	-
	20	-	26,030	50,280	65,600	77,860	94,480
	40	-	26,540	51,200	66,790	79,000	95,270
	100	-	21,050	34,770	40,790	46,310	55,500
SNKM	5	-	9,160	30,330	68,400	-	-
	10	-	8,000	29,740	66,720	91,270	-
	20	-	5,820	22,900	51,500	73,140	94,480
	40	-	5,880	23,470	52,780	74,660	95,070
	100	-	2,480	10,560	24,560	35,590	48,650
MGM	5	-	25,580	51,160	69,420	-	-
	10	-	41,480	73,220	90,070	93,460	-
	20	-	50,670	83,330	96,100	97,680	98,180
	40	-	50,440	83,050	96,020	97,600	98,180
	100	-	61,150	91,770	99,170	99,580	99,690

Tabela 31 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com dois erros padrões ($2\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 5% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	8,780	52,150	99,820	-	-
	10	-	4,020	35,580	99,230	100,00	-
	20	-	1,680	23,470	98,090	100,00	100,00
	40	-	0,690	14,310	95,970	100,00	100,00
	100	-	0,190	6,920	91,010	100,00	100,00
SNK	5	-	25,720	68,330	99,880	-	-
	10	-	25,710	69,240	99,870	100,00	-
	20	-	25,720	69,430	99,890	100,00	100,00
	40	-	25,470	69,390	99,910	100,00	100,00
	100	-	25,420	69,460	99,900	100,00	100,00
MGR	5	-	38,150	72,470	99,900	-	-
	10	-	37,420	70,410	99,700	100,00	-
	20	-	36,660	68,530	99,270	100,00	100,00
	40	-	35,240	66,010	98,540	100,00	100,00
	100	-	33,590	63,310	97,080	100,00	100,00
TM	5	-	40,630	70,770	87,000	-	-
	10	-	51,500	81,960	92,510	96,620	-
	20	-	53,400	80,130	87,180	91,720	98,320
	40	-	53,160	79,730	86,780	91,320	98,340
	100	-	45,990	62,990	66,700	70,000	75,950
SNKM	5	-	24,990	61,070	86,800	-	-
	10	-	24,900	62,340	88,970	96,560	-
	20	-	21,240	55,720	79,850	89,390	98,220
	40	-	21,110	55,610	79,650	89,120	98,240
	100	-	13,590	36,850	53,760	61,570	71,020
MGM	5	-	43,060	73,500	87,060	-	-
	10	-	57,230	87,640	96,530	97,360	-
	20	-	64,390	92,990	98,770	99,100	99,320
	40	-	64,430	92,940	98,700	99,080	99,310
	100	-	72,610	96,750	99,840	99,900	99,900

Tabela 32 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com quatro erros padrões ($4\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 1% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	-	20,020	92,730	100,000	-
	10	-	-	14,430	92,290	100,000	100,000
	20	-	-	9,320	90,760	100,000	100,000
	40	-	-	5,840	87,440	100,000	100,000
	100	-	-	2,860	80,910	100,000	100,000
SNK	5	-	-	47,790	97,530	100,000	-
	10	-	-	53,980	99,100	100,000	100,000
	20	-	-	56,840	99,500	100,000	100,000
	40	-	-	58,250	99,590	100,000	100,000
	100	-	-	59,400	80,910	100,000	100,000
MGR	5	-	-	54,120	94,850	100,000	-
	10	-	-	55,790	94,540	100,000	100,000
	20	-	-	55,990	93,600	100,000	100,000
	40	-	-	54,710	91,740	100,000	100,000
	100	-	-	59,400	99,670	100,000	100,000
TM	5	-	-	73,820	94,640	97,500	-
	10	-	-	86,300	99,400	99,660	99,790
	20	-	-	91,000	99,470	99,530	99,640
	40	-	-	92,500	98,990	99,060	99,200
	100	-	-	93,930	98,800	98,850	98,970
SNKM	5	-	-	47,740	89,150	97,500	-
	10	-	-	53,200	95,750	99,460	99,780
	20	-	-	56,410	97,050	99,140	99,470
	40	-	-	57,290	96,620	98,280	98,720
	100	-	-	57,490	95,950	97,420	97,880
MGM	5	-	-	74,340	95,050	97,500	-
	10	-	-	86,520	99,540	99,790	99,820
	20	-	-	91,470	99,940	99,960	99,970
	40	-	-	93,450	99,990	100,000	100,000
	100	-	-	95,100	100,000	100,000	100,000

Tabela 33 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com quatro erros padrões ($4\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 5% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	-	42,890	98,890	100,000	-
	10	-	-	30,630	98,070	100,000	100,000
	20	-	-	20,500	96,900	100,000	100,000
	40	-	-	13,230	94,650	100,000	100,000
	100	-	-	6,660	90,130	100,000	100,000
SNK	5	-	-	75,400	99,730	100,000	-
	10	-	-	78,010	99,910	100,000	100,000
	20	-	-	79,470	99,930	100,000	100,000
	40	-	-	80,050	99,950	100,000	100,000
	100	-	-	80,520	99,950	100,000	100,000
MGR	5	-	-	61,420	99,080	100,000	-
	10	-	-	60,010	98,360	100,000	100,000
	20	-	-	59,670	97,400	100,000	100,000
	40	-	-	58,590	95,800	100,000	100,000
	100	-	-	56,160	92,990	100,000	100,000
TM	5	-	-	89,570	99,670	99,780	-
	10	-	-	94,270	99,910	99,920	99,920
	20	-	-	95,990	99,960	99,960	99,970
	40	-	-	96,680	99,860	99,870	99,880
	100	-	-	97,290	99,730	99,730	99,750
SNKM	5	-	-	75,280	99,030	99,780	-
	10	-	-	78,280	99,710	99,920	99,920
	20	-	-	79,560	99,810	99,960	99,970
	40	-	-	79,840	99,640	99,770	99,810
	100	-	-	80,090	99,400	99,520	99,570
MGM	5	-	-	89,580	99,670	99,780	-
	10	-	-	94,270	99,910	99,920	99,920
	20	-	-	96,030	100,000	100,000	100,000
	40	-	-	96,810	100,000	100,000	100,000
	100	-	-	97,550	100,000	100,000	100,000

Tabela 34 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com quatro erros padrões ($4\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 1% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	-	28,170	98,090	100,000	-
	10	-	-	18,050	96,600	100,000	100,000
	20	-	-	11,140	93,870	100,000	100,000
	40	-	-	6,660	90,050	100,000	100,000
	100	-	-	3,040	82,490	100,000	100,000
SNK	5	-	-	55,960	99,310	100,000	-
	10	-	-	57,740	99,600	100,000	100,000
	20	-	-	58,890	99,670	100,000	100,000
	40	-	-	59,400	99,680	100,000	100,000
	100	-	-	59,780	99,700	100,000	100,000
MGR	5	-	-	56,850	98,290	100,000	-
	10	-	-	57,960	97,140	100,000	100,000
	20	-	-	57,330	95,180	100,000	100,000
	40	-	-	55,810	92,900	100,000	100,000
	100	-	-	52,650	89,410	100,000	100,000
TM	5	-	-	79,900	98,370	99,260	-
	10	-	-	88,300	99,760	99,880	99,910
	20	-	-	91,570	99,620	99,680	99,790
	40	-	-	92,840	99,160	99,210	99,330
	100	-	-	93,900	98,770	98,820	98,940
SNKM	5	-	-	55,990	95,930	99,260	-
	10	-	-	57,940	97,880	99,830	99,910
	20	-	-	58,580	97,850	99,410	99,650
	40	-	-	58,450	97,110	98,490	98,850
	100	-	-	57,980	96,190	97,530	97,970
MGM	5	-	-	80,000	98,370	99,260	-
	10	-	-	88,370	99,800	99,920	99,920
	20	-	-	91,890	99,960	99,990	100,000
	40	-	-	93,620	99,970	99,970	99,970
	100	-	-	95,090	100,000	100,000	100,000

Tabela 35 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com quatro erros padrões ($4\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 5% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	-	50,190	99,680	100,000	-
	10	-	-	34,410	99,020	100,000	100,000
	20	-	-	22,540	97,920	100,000	100,000
	40	-	-	14,030	95,660	100,000	100,000
	100	-	-	6,930	90,840	100,000	100,000
SNK	5	-	-	79,570	99,910	100,000	-
	10	-	-	79,940	99,940	100,000	100,000
	20	-	-	80,530	99,950	100,000	100,000
	40	-	-	80,460	99,950	100,000	100,000
	100	-	-	80,690	99,960	100,000	100,000
MGR	5	-	-	62,940	99,690	100,000	-
	10	-	-	60,660	99,110	100,000	100,000
	20	-	-	60,630	98,170	100,000	100,000
	40	-	-	59,150	96,390	100,000	100,000
	100	-	-	56,520	93,300	100,000	100,000
TM	5	-	-	91,220	99,920	99,960	-
	10	-	-	94,740	99,970	99,980	99,980
	20	-	-	95,980	99,940	99,940	99,960
	40	-	-	96,130	99,960	99,970	99,980
	100	-	-	97,330	99,730	99,740	99,770
SNKM	5	-	-	79,110	99,650	99,960	-
	10	-	-	79,990	99,860	99,980	99,980
	20	-	-	80,210	99,830	99,930	99,960
	40	-	-	80,280	99,840	99,950	99,970
	100	-	-	80,220	99,430	99,540	99,600
MGM	5	-	-	91,230	99,920	99,960	-
	10	-	-	94,750	99,970	99,980	99,980
	20	-	-	96,030	100,000	100,000	100,000
	40	-	-	96,160	100,000	100,000	100,000
	100	-	-	97,590	100,000	100,000	100,000

Tabela 36 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com quatro erros padrões ($4\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 1% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	-	30,670	98,760	100,000	-
	10	-	-	19,320	97,260	100,000	100,000
	20	-	-	11,920	94,610	100,000	100,000
	40	-	-	6,780	90,620	100,000	100,000
	100	-	-	3,110	82,940	100,000	100,000
SNK	5	-	-	57,810	99,590	100,000	-
	10	-	-	58,780	99,640	100,000	100,000
	20	-	-	59,310	99,670	100,000	100,000
	40	-	-	59,680	99,670	100,000	100,000
	100	-	-	59,910	99,710	100,000	100,000
MGR	5	-	-	57,160	98,870	100,000	-
	10	-	-	58,450	97,620	100,000	100,000
	20	-	-	57,820	95,680	100,000	100,000
	40	-	-	56,100	93,190	100,000	100,000
	100	-	-	52,770	89,500	100,000	100,000
TM	5	-	-	82,080	99,100	99,600	-
	10	-	-	58,940	98,270	99,870	99,920
	20	-	-	91,740	99,720	99,750	99,810
	40	-	-	92,960	99,210	99,270	99,380
	100	-	-	94,190	99,030	99,070	99,150
SNKM	5	-	-	57,730	97,330	99,600	-
	10	-	-	58,940	98,270	99,870	99,920
	20	-	-	59,170	98,240	99,590	99,740
	40	-	-	58,570	97,300	98,570	98,920
	100	-	-	58,400	96,670	97,950	98,330
MGM	5	-	-	82,150	99,110	99,600	-
	10	-	-	88,630	99,830	99,910	99,920
	20	-	-	91,980	99,960	99,980	99,980
	40	-	-	92,700	99,980	99,990	100,000
	100	-	-	95,120	100,000	100,000	100,000

Tabela 37 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com quatro erros padrões ($4\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 5% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	-	52,270	99,770	100,000	-
	10	-	-	35,890	99,190	100,000	-
	20	-	-	23,160	98,010	100,000	100,000
	40	-	-	14,230	95,980	100,000	100,000
	100	-	-	6,970	91,080	100,000	100,000
SNK	5	-	-	80,240	99,930	100,000	-
	10	-	-	80,340	99,930	100,000	100,000
	20	-	-	80,540	99,950	100,000	100,000
	40	-	-	80,570	99,950	100,000	100,000
	100	-	-	80,700	99,960	100,000	100,000
MGR	5	-	-	63,050	99,780	100,000	-
	10	-	-	60,770	99,250	100,000	100,000
	20	-	-	60,350	98,230	100,000	100,000
	40	-	-	59,420	96,620	100,000	100,000
	100	-	-	56,590	93,430	100,000	100,000
TM	5	-	-	91,480	99,980	100,000	-
	10	-	-	94,660	99,990	100,000	100,000
	20	-	-	96,060	99,970	99,970	99,970
	40	-	-	96,820	99,850	99,860	99,870
	100	-	-	97,250	99,670	99,680	99,710
SNKM	5	-	-	79,590	99,810	100,000	-
	10	-	-	80,140	99,870	100,000	100,000
	20	-	-	80,480	99,840	99,940	99,950
	40	-	-	80,460	99,670	99,770	99,810
	100	-	-	80,210	99,360	99,480	99,540
MGM	5	-	-	91,490	99,980	100,000	-
	10	-	-	94,670	99,990	100,000	100,000
	20	-	-	96,090	100,000	100,000	100,000
	40	-	-	96,970	100,000	100,000	100,000
	100	-	-	97,570	100,000	100,000	100,000

Tabela 38 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com oito erros padrões ($8\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 1% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	-	-	93,040	100,000	100,000
	10	-	-	-	92,360	100,000	100,000
	20	-	-	-	90,560	100,000	100,000
	40	-	-	-	87,580	100,000	100,000
	100	-	-	-	80,820	100,000	100,000
SNK	5	-	-	-	99,320	100,000	100,000
	10	-	-	-	99,670	100,000	100,000
	20	-	-	-	99,830	100,000	100,000
	40	-	-	-	99,850	100,000	100,000
	100	-	-	-	99,880	100,000	100,000
MGR	5	-	-	-	99,420	100,000	100,000
	10	-	-	-	92,610	100,000	100,000
	20	-	-	-	90,800	100,000	100,000
	40	-	-	-	87,920	100,000	100,000
	100	-	-	-	81,760	100,000	100,000
TM	5	-	-	-	99,960	100,000	100,000
	10	-	-	-	99,990	100,000	100,000
	20	-	-	-	100,000	100,000	100,000
	40	-	-	-	100,000	100,000	100,000
	100	-	-	-	100,000	100,000	100,000
SNKM	5	-	-	-	99,120	100,000	100,000
	10	-	-	-	99,740	100,000	100,000
	20	-	-	-	99,830	100,000	100,000
	40	-	-	-	99,860	100,000	100,000
	100	-	-	-	99,890	100,000	100,000
MGM	5	-	-	-	99,960	100,000	100,000
	10	-	-	-	99,990	100,000	100,000
	20	-	-	-	100,000	100,000	100,000
	40	-	-	-	100,000	100,000	100,000
	100	-	-	-	100,000	100,000	100,000

Tabela 39 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com oito erros padrões ($8\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 4 ($r = 4$), sob H_1 completa, ao nível de significância de 5% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	-	-	98,820	100,000	100,000
	10	-	-	-	98,070	100,000	100,000
	20	-	-	-	96,920	100,000	100,000
	40	-	-	-	94,820	100,000	100,000
	100	-	-	-	89,960	100,000	100,000
SNK	5	-	-	-	99,970	100,000	100,000
	10	-	-	-	99,980	100,000	100,000
	20	-	-	-	99,980	100,000	100,000
	40	-	-	-	99,990	100,000	100,000
	100	-	-	-	99,990	100,000	100,000
MGR	5	-	-	-	98,830	100,000	100,000
	10	-	-	-	98,070	100,000	100,000
	20	-	-	-	96,930	100,000	100,000
	40	-	-	-	94,840	100,000	100,000
	100	-	-	-	90,070	100,000	100,000
TM	5	-	-	-	100,000	100,000	100,000
	10	-	-	-	100,000	100,000	100,000
	20	-	-	-	100,000	100,000	100,000
	40	-	-	-	100,000	100,000	100,000
	100	-	-	-	100,000	100,000	100,000
SNKM	5	-	-	-	99,960	100,000	100,000
	10	-	-	-	99,980	100,000	100,000
	20	-	-	-	99,980	100,000	100,000
	40	-	-	-	99,980	100,000	100,000
	100	-	-	-	99,990	100,000	100,000
MGM	5	-	-	-	100,000	100,000	100,000
	10	-	-	-	100,000	100,000	100,000
	20	-	-	-	100,000	100,000	100,000
	40	-	-	-	100,000	100,000	100,000
	100	-	-	-	100,000	100,000	100,000

Tabela 40 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com oito erros padrões ($8\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 1% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	-	-	98,120	100,000	100,000
	10	-	-	-	96,560	100,000	100,000
	20	-	-	-	94,040	100,000	100,000
	40	-	-	-	90,050	100,000	100,000
	100	-	-	-	82,390	100,000	100,000
SNK	5	-	-	-	99,810	100,000	100,000
	10	-	-	-	99,860	100,000	100,000
	20	-	-	-	99,870	100,000	100,000
	40	-	-	-	99,890	100,000	100,000
	100	-	-	-	99,900	100,000	100,000
MGR	5	-	-	-	98,130	100,000	100,000
	10	-	-	-	96,570	100,000	100,000
	20	-	-	-	94,050	100,000	100,000
	40	-	-	-	90,150	100,000	100,000
	100	-	-	-	82,990	100,000	100,000
TM	5	-	-	-	99,980	100,000	100,000
	10	-	-	-	100,000	100,000	100,000
	20	-	-	-	100,000	100,000	100,000
	40	-	-	-	100,000	100,000	100,000
	100	-	-	-	100,000	100,000	100,000
SNKM	5	-	-	-	99,810	100,000	100,000
	10	-	-	-	99,840	100,000	100,000
	20	-	-	-	99,880	100,000	100,000
	40	-	-	-	99,890	100,000	100,000
	100	-	-	-	99,890	100,000	100,000
MGM	5	-	-	-	99,980	100,000	100,000
	10	-	-	-	100,000	100,000	100,000
	20	-	-	-	100,000	100,000	100,000
	40	-	-	-	100,000	100,000	100,000
	100	-	-	-	100,000	100,000	100,000

Tabela 41 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com oito erros padrões ($8\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 10 ($r = 10$), sob H_1 completa, ao nível de significância de 5% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	-	-	99,650	100,000	100,000
	10	-	-	-	99,040	100,000	100,000
	20	-	-	-	97,870	100,000	100,000
	40	-	-	-	95,870	100,000	100,000
	100	-	-	-	90,900	100,000	100,000
SNK	5	-	-	-	99,990	100,000	100,000
	10	-	-	-	99,980	100,000	100,000
	20	-	-	-	99,990	100,000	100,000
	40	-	-	-	99,990	100,000	100,000
	100	-	-	-	99,990	100,000	100,000
MGR	5	-	-	-	99,650	100,000	100,000
	10	-	-	-	99,040	100,000	100,000
	20	-	-	-	97,870	100,000	100,000
	40	-	-	-	95,870	100,000	100,000
	100	-	-	-	90,960	100,000	100,000
TM	5	-	-	-	100,000	100,000	100,000
	10	-	-	-	100,000	100,000	100,000
	20	-	-	-	100,000	100,000	100,000
	40	-	-	-	100,000	100,000	100,000
	100	-	-	-	100,000	100,000	100,000
SNKM	5	-	-	-	99,980	100,000	100,000
	10	-	-	-	99,990	100,000	100,000
	20	-	-	-	99,990	100,000	100,000
	40	-	-	-	99,990	100,000	100,000
	100	-	-	-	99,990	100,000	100,000
MGM	5	-	-	-	100,000	100,000	100,000
	10	-	-	-	100,000	100,000	100,000
	20	-	-	-	100,000	100,000	100,000
	40	-	-	-	100,000	100,000	100,000
	100	-	-	-	100,000	100,000	100,000

Tabela 42 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com oito erros padrões ($8\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 1% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	-	-	98,810	100,000	100,000
	10	-	-	-	97,190	100,000	100,000
	20	-	-	-	95,700	100,000	100,000
	40	-	-	-	90,570	100,000	100,000
	100	-	-	-	82,920	100,000	100,000
SNK	5	-	-	-	99,860	100,000	100,000
	10	-	-	-	99,860	100,000	100,000
	20	-	-	-	99,900	100,000	100,000
	40	-	-	-	99,890	100,000	100,000
	100	-	-	-	99,900	100,000	100,000
MGR	5	-	-	-	98,810	100,000	100,000
	10	-	-	-	97,200	100,000	100,000
	20	-	-	-	94,710	100,000	100,000
	40	-	-	-	90,640	100,000	100,000
	100	-	-	-	83,430	100,000	100,000
TM	5	-	-	-	100,000	100,000	100,000
	10	-	-	-	100,000	100,000	100,000
	20	-	-	-	100,000	100,000	100,000
	40	-	-	-	100,000	100,000	100,000
	100	-	-	-	100,000	100,000	100,000
SNKM	5	-	-	-	99,840	100,000	100,000
	10	-	-	-	99,880	100,000	100,000
	20	-	-	-	99,880	100,000	100,000
	40	-	-	-	99,890	100,000	100,000
	100	-	-	-	99,900	100,000	100,000
MGM	5	-	-	-	100,000	100,000	100,000
	10	-	-	-	100,000	100,000	100,000
	20	-	-	-	100,000	100,000	100,000
	40	-	-	-	100,000	100,000	100,000
	100	-	-	-	100,000	100,000	100,000

Tabela 43 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias iniciando com oito erros padrões ($8\sigma_{\bar{Y}}$) a $32\sigma_{\bar{Y}}$, em função do número de tratamentos e do número de repetições igual a 20 ($r = 20$), sob H_1 completa, ao nível de significância de 5% de probabilidade.

Teste	Nº Trat	Diferença real entre médias					
		$1\sigma_{\bar{Y}}$	$2\sigma_{\bar{Y}}$	$4\sigma_{\bar{Y}}$	$8\sigma_{\bar{Y}}$	$16\sigma_{\bar{Y}}$	$32\sigma_{\bar{Y}}$
Tukey	5	-	-	-	99,710	100,000	100,000
	10	-	-	-	99,230	100,000	100,000
	20	-	-	-	98,200	100,000	100,000
	40	-	-	-	95,940	100,000	100,000
	100	-	-	-	91,120	100,000	100,000
SNK	5	-	-	-	99,980	100,000	100,000
	10	-	-	-	99,990	100,000	100,000
	20	-	-	-	99,990	100,000	100,000
	40	-	-	-	99,990	100,000	100,000
	100	-	-	-	99,990	100,000	100,000
MGR	5	-	-	-	99,710	100,000	100,000
	10	-	-	-	99,230	100,000	100,000
	20	-	-	-	98,200	100,000	100,000
	40	-	-	-	95,940	100,000	100,000
	100	-	-	-	91,170	100,000	100,000
TM	5	-	-	-	100,000	100,000	100,000
	10	-	-	-	100,000	100,000	100,000
	20	-	-	-	100,000	100,000	100,000
	40	-	-	-	100,000	100,000	100,000
	100	-	-	-	100,000	100,000	100,000
SNKM	5	-	-	-	99,970	100,000	100,000
	10	-	-	-	99,980	100,000	100,000
	20	-	-	-	99,990	100,000	100,000
	40	-	-	-	99,990	100,000	100,000
	100	-	-	-	99,990	100,000	100,000
MGM	5	-	-	-	100,000	100,000	100,000
	10	-	-	-	100,000	100,000	100,000
	20	-	-	-	100,000	100,000	100,000
	40	-	-	-	100,000	100,000	100,000
	100	-	-	-	100,000	100,000	100,000

APÊNDICE C - Poder sob H_0 parcial

Tabela 44 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de um erro padrão ($1\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 1% de probabilidade.

Trat	Rep	Poder dos testes					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	0,490	0,610	1,280	0,560	0,480	0,940
	10	0,470	0,540	1,360	0,540	0,430	0,860
	20	0,470	0,530	1,410	0,540	0,410	0,760
10	4	0,170	0,200	0,810	0,390	0,330	1,290
	10	0,180	0,200	0,960	0,310	0,260	1,190
	20	0,160	0,180	0,920	0,260	0,220	0,980
20	4	0,050	0,060	0,510	0,130	0,120	0,700
	10	0,050	0,050	0,510	0,120	0,100	0,690
	20	0,050	0,060	0,480	0,120	0,110	0,650
40	4	0,020	0,020	0,270	0,040	0,040	0,230
	10	0,020	0,020	0,290	0,050	0,040	0,140
	20	0,020	0,020	0,270	0,050	0,040	0,210
100	4	0,640	0,640	0,120	0,020	0,020	0,030
	10	0,000	0,000	0,130	0,800	0,800	0,800
	20	0,000	0,000	0,120	0,020	0,020	0,060

Tabela 45 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de um erro padrão ($1\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 5% de probabilidade.

Trat	Rep	Poder dos testes					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	2,100	2,720	4,990	3,230	2,600	4,090
	10	2,120	2,620	5,180	2,730	1,960	3,670
	20	2,040	2,650	5,440	2,820	2,240	3,560
10	4	0,680	0,800	3,330	2,110	1,760	4,840
	10	0,710	0,840	3,470	1,960	1,600	4,510
	20	0,710	0,830	3,640	1,710	1,430	4,260
20	4	0,240	0,280	1,990	0,930	0,830	3,300
	10	0,220	0,250	2,080	0,820	0,730	3,230
	20	0,220	0,250	2,130	0,770	0,680	3,040
40	4	0,080	0,080	1,150	0,310	0,290	1,210
	10	0,070	0,080	1,170	0,330	0,310	1,390
	20	0,070	0,080	1,050	0,300	0,290	1,180
100	4	0,020	0,020	0,520	0,100	0,100	0,210
	10	0,020	0,020	0,540	0,100	0,100	0,200
	20	0,020	0,020	0,510	0,090	0,090	0,170

Tabela 46 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de dois erros padrões ($2\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 1% de probabilidade.

Trat	Rep	Poder dos testes					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	2,140	2,610	5,260	3,310	2,630	4,900
	10	2,610	3,090	6,900	3,270	2,450	4,990
	20	3,040	3,600	8,240	3,810	2,840	5,690
10	4	1,040	1,230	4,940	2,150	1,780	7,420
	10	1,260	1,450	6,480	2,180	1,770	7,260
	20	1,220	1,390	6,350	2,070	1,700	7,770
20	4	0,440	0,510	3,940	0,830	0,720	5,330
	10	0,450	0,510	4,320	0,970	0,860	6,050
	20	0,460	0,510	4,390	0,870	0,780	5,490
40	4	0,180	0,200	2,460	0,290	0,270	2,730
	10	0,180	0,190	2,630	0,230	0,220	2,650
	20	0,180	0,200	2,790	0,290	0,270	2,820
100	4	0,050	0,050	1,200	0,040	0,040	0,350
	10	0,020	0,050	1,290	0,030	0,030	0,190
	20	0,050	0,050	1,320	0,030	0,030	0,270

Tabela 47 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de dois erros padrões ($2\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 5% de probabilidade.

Trat	Rep	Poder dos testes					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	7,470	9,510	16,400	12,630	10,340	15,880
	10	8,870	11,480	20,890	12,250	10,040	15,270
	20	9,170	11,610	21,220	12,150	9,820	15,020
10	4	3,470	4,370	14,580	9,130	7,670	18,770
	10	4,030	4,990	17,550	8,260	6,890	18,010
	20	3,960	4,910	17,600	8,410	7,000	18,180
20	4	1,520	1,820	11,530	4,330	3,830	13,740
	10	1,630	1,910	12,870	3,990	3,520	13,390
	20	1,620	1,910	13,030	3,690	3,250	12,860
40	4	0,610	0,690	7,670	1,340	1,250	6,170
	10	0,660	0,740	8,300	1,130	1,040	5,270
	20	0,660	0,740	7,980	1,280	1,180	6,160
100	4	0,190	0,210	3,590	0,180	0,180	0,740
	10	0,190	0,200	3,790	0,150	0,150	0,590
	20	0,190	0,200	3,730	0,170	0,160	0,650

Tabela 48 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de quatro erros padrões ($4\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 1% de probabilidade.

Poder dos testes							
Trat	Rep	Tukey	SNK	MGR	TM	SNKM	MGM
5	4	19,690	25,480	41,890	38,620	31,000	47,040
	10	27,890	34,860	57,560	45,180	36,700	53,930
	20	31,540	39,200	63,220	48,570	39,510	56,800
10	4	14,350	18,930	57,720	41,720	34,140	69,600
	10	18,520	23,940	70,490	44,290	36,670	70,750
	20	19,490	25,100	72,520	45,340	37,550	72,190
20	4	9,370	11,960	69,950	35,600	30,540	68,470
	10	11,060	13,970	77,390	36,200	31,200	68,780
	20	11,850	14,930	79,590	35,660	30,810	68,330
40	4	5,810	7,120	77,100	26,150	23,460	55,800
	10	6,570	8,010	81,080	25,830	23,120	54,730
	20	6,660	8,110	82,490	26,540	23,790	55,570
100	4	2,840	3,310	78,380	11,350	10,540	26,200
	10	3,020	3,510	82,100	11,570	10,750	26,540
	20	3,080	3,580	82,350	12,000	11,140	27,280

Tabela 49 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de quatro erros padrões ($4\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 5% de probabilidade.

Poder dos testes							
Trat	Rep	Tukey	SNK	MGR	TM	SNKM	MGM
5	4	43,480	54,970	71,850	67,700	60,620	72,640
	10	50,250	62,400	78,680	69,040	61,820	73,320
	20	52,070	63,930	80,240	70,670	64,060	75,050
10	4	30,460	41,270	79,650	65,100	57,720	81,250
	10	34,870	46,820	84,100	65,240	57,920	80,240
	20	35,950	47,890	84,470	64,980	57,790	80,120
20	4	20,690	27,930	84,530	53,840	48,530	75,270
	10	22,950	30,660	87,250	54,490	49,170	75,420
	20	23,210	31,010	87,620	54,900	49,600	76,090
40	4	13,120	17,000	87,490	40,000	37,070	60,290
	10	14,160	18,250	89,200	39,760	36,810	59,370
	20	14,220	18,310	89,280	39,840	36,860	60,140
100	4	6,650	8,130	88,780	18,260	17,320	29,860
	10	6,900	8,420	89,340	18,360	17,430	29,760
	20	6,880	8,380	89,930	18,650	17,660	29,870

Tabela 50 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de oito erros padrões ($8\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_0 parcial, ao nível de significância de 1% de probabilidade.

Trat	Rep	Poder dos testes					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	92,510	97,480	99,310	99,870	99,140	99,920
	10	98,150	99,540	99,920	99,980	99,780	100,000
	20	98,800	99,810	99,980	99,940	99,770	99,950
10	4	92,410	98,600	99,980	99,920	99,590	99,970
	10	96,530	99,580	99,980	99,960	99,760	100,000
	20	97,370	99,680	99,980	99,970	99,800	100,000
20	4	90,800	98,660	100,000	99,940	99,760	99,990
	10	93,960	99,300	99,980	99,940	99,810	99,970
	20	94,640	99,440	99,990	99,980	99,840	100,000
40	4	87,610	97,790	99,990	99,970	99,860	99,990
	10	89,950	98,540	99,990	99,980	99,880	99,990
	20	90,590	98,690	99,990	99,960	99,870	99,990
100	4	80,940	94,550	99,990	99,910	99,850	99,950
	10	82,490	95,420	99,990	99,920	99,850	99,950
	20	82,840	95,580	99,990	99,960	99,900	100,000

Tabela 51 Poder dos testes de Tukey, de SNK, de MGR, de TM, de SNKM e de MGM, em percentagem, para detectar uma diferença entre médias de oito erros padrões ($8\sigma_{\bar{Y}}$), em função do número de tratamentos e do número de repetições, sob H_1 parcial, ao nível de significância de 5% de probabilidade.

Trat	Rep	Poder dos testes					
		Tukey	SNK	MGR	TM	SNKM	MGM
5	4	98,900	99,860	99,950	99,980	99,920	99,990
	10	99,710	99,970	100,000	99,980	99,960	99,980
	20	99,790	99,980	99,990	99,970	99,950	99,970
10	4	98,090	99,860	99,980	99,980	99,960	99,990
	10	99,090	99,980	100,000	100,000	99,990	100,000
	20	99,260	99,980	99,990	99,980	99,960	99,980
20	4	96,860	99,870	100,000	99,990	99,970	100,000
	10	97,890	99,930	100,000	100,000	99,970	100,000
	20	98,080	99,940	100,000	99,980	99,960	100,000
40	4	95,750	99,750	99,990	99,990	99,980	100,000
	10	95,700	99,820	99,990	99,970	99,950	99,980
	20	95,970	99,830	100,000	100,000	99,980	100,000
100	4	90,010	98,960	99,990	99,980	99,970	100,000
	10	90,870	99,160	99,990	99,980	99,970	100,000
	20	91,060	99,180	100,000	99,960	99,950	99,980

