



SIOMARA CRISTINA BROCH

**ASPECTOS TEÓRICOS E COMPUTACIONAIS
DAS ESTATÍSTICAS DO TESTE DE DUNNETT
NÃO-CENTRAL**

LAVRAS – MG

2013

SIOMARA CRISTINA BROCH

**ASPECTOS TEÓRICOS E COMPUTACIONAIS DAS ESTATÍSTICAS
DO TESTE DE DUNNETT NÃO-CENTRAL**

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 Doutora.

Orientador

Dr. Daniel Furtado Ferreira

LAVRAS – MG

2013

**Ficha Catalográfica Elaborada pela Divisão de Processos Técnicos da
Biblioteca da UFLA**

Broch, Siomara Cristina.

Aspectos teóricos e computacionais das estatísticas do teste de
Dunnett não-central / Siomara Cristina Broch. – Lavras : UFLA,
2013.

240 p. : il.

Tese (doutorado) – Universidade Federal de Lavras, 2012.

Orientador: Daniel Furtado Ferreira.

Bibliografia.

1. *T* multivariada não-central. 2. Comparações múltiplas com um
controle. 3. *Software R*. I. Universidade Federal de Lavras.

II. Título.

CDD – 519.5

SIOMARA CRISTINA BROCH

**ASPECTOS TEÓRICOS E COMPUTACIONAIS DAS ESTATÍSTICAS
DO TESTE DE DUNNETT NÃO-CENTRAL**

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 Doutora.

APROVADA em 20 de fevereiro de 2013.

Dr. Éric Batista Ferreira	UNIFAL
Dr. José Airton Rodrigues Nunes	UFLA
Dr. Lucas Monteiro Chaves	UFLA
Dr. Renato Ribeiro de Lima	UFLA

Dr. Daniel Furtado Ferreira
Orientador

**LAVRAS – MG
2013**

Aos meus queridos filhos Guilherme, Luiz Henrique e Gabriel,
Aos meus pais José (*in memoriam*) e Cristina,
Pelo amor incondicional.

DEDICO

AGRADECIMENTOS

À Deus, pela saúde, coragem e discernimento para vencer todas as dificuldades, incertezas e solidão que povoaram esses três anos de dedicação ao meu sonho.

Aos meus filhos, minha mãe, meus irmãos e demais familiares, por respeitarem minhas decisões e acreditarem no meu sucesso.

Ao professor e orientador Daniel Furtado Ferreira, pela atenção e paciência com que conduziu nossos estudos. Seu profissionalismo e dedicação são inspiração para seguir na profissão do magistério.

Aos colegas do curso de Pós-graduação em Estatística e Experimentação Agropecuária, pelo colegismo e amizade que nos uniu, seja nas horas de estudo ou no lazer. Foi um prazer conhecê-los e conviver com todos durante esse tempo.

Aos amigos e amigas que entenderam meus objetivos e, mesmo de longe, me incentivaram nesta jornada.

Em especial à Teresa, Alma Cristina, Cleonir, Leontina, Luis Felipe e Weber, pelo apoio constante.

Aos professores do Departamento de Ciências Exatas da UFLA, pelos ensinamentos, amizade e boa convivência.

A todos os funcionários do Departamento de Ciências Exatas, em especial a Josi, secretária da Pós-graduação, pelo suporte e atenção.

Sinceros agradecimentos a todos!

RESUMO

O teste de Dunnett é um teste utilizado para comparar simultaneamente a média de tratamentos em teste com a média de um tratamento controle, considerando que as amostras são aleatórias e independentes, oriundas de variáveis com distribuições normais. A limitação para o uso deste teste é a dificuldade em se obter as probabilidades da distribuição t multivariada e os valores dos quantis da estatística, pois o teste pode ser aplicado em situações balanceadas e não-balanceadas, unilateral ou bilateral, com infinitas possibilidades de correlações entre as comparações. Este trabalho apresentou detalhadamente as distribuições t multivariadas não-centrais relacionadas ao teste de Dunnett e desenvolveu, implementou e disponibilizou uma biblioteca no *software R* específica para obter as probabilidades das distribuições do teste, com dados balanceados ou não e em testes unilaterais e bilaterais, considerando o caso não-central. Além disso, a rotina também fornece o valor dos quantis para a estatística do teste nestes casos. Partindo das ideias propostas por Dunlap, Marx e Agamy (1981), buscou-se ampliar e melhorar a performance do algoritmo proposto por esses autores, utilizando métodos numéricos de quadratura gaussiana ao invés da regra de Simpson para resolver as integrais, desenvolvendo funções que fornecessem os valores desejados. Os algoritmos foram propostos e implementados com sucesso. Os valores das probabilidades e dos quantis obtidos pelas rotinas têm, no mínimo, a mesma precisão de valores comparados a partir de referências bibliográficas ou outras rotinas computacionais. Além disso, a precisão dos valores fornecidos pode aumentar conforme a necessidade do usuário, apenas aumentando o número de pontos para a quadratura. A biblioteca *nCDunnett* foi implementada e submetida ao CRAN-R e encontra-se disponibilizada para instalação em todas as versões R iguais ou superiores a 2.15.0, com livre acesso por todos os usuários do programa no mundo.

Palavras-chave: T multivariada não-central. Comparações múltiplas com um controle. *Software R*.

ABSTRACT

The Dunnett test is used to simultaneously compare the mean of the treatments to the mean of a control treatment, considering random and independent samples, derived from normal distribution variables. Obtaining the probabilities of multivariate t distribution and the statistical quantiles values is a limitation due to the possibility of applying the test in balanced and non-balanced situations, unilaterally or bilaterally, with infinite correlation possibilities between the comparisons. This work presented the detailed non-central multivariate t distribution related to the Dunnett test. It also developed, implemented and made available an R software library specific to obtaining the probabilities of the test's distribution, with balanced or unbalanced data in unilateral or bilateral tests, considering it a non-central case. In addition, the routine also provides the quantiles values for the statistics of the test in these cases. Based on the ideas proposed by Dunlap, Marx and Agamy (1981), we attempted to amplify and improve the performance of the algorithm proposed by these authors using Gaussian quadrature numeric methods instead of the Simpson rule to solve integrals, developing functions which provide the desired values. The algorithms were proposed and implemented successfully. The probability and quantiles values obtained by the routines have, at least, the same precision of values compared from bibliographical references or other computational routines. In addition, the precision of the provided values may increase according to the user's necessity, increasing only the number of quadrature points. The *nCDunnett* library was implemented and submitted to the CRAN-R and is available for installation in all R versions, equal or superior to 2.15.0, with free access for all users.

Key-words: Non-central multivariate T. Multiple comparisons with control. R software.

LISTA DE FIGURAS

Figura 1	Probabilidade de erro tipo I global em procedimentos de comparações múltiplas sem controle de multiplicidade para diferentes níveis nominais de significância (α) e número de hipóteses nulas testadas (m).....	28
Figura 2	Representação das funções peso $\lambda(x) = 1$, $\lambda(x) = 1 / \sqrt{1 - x^2}$ e $\lambda(x) = \sqrt{1 - x^2}$, ortogonais em relação a família de polinômios de Legendre, Chebyshev do 1º tipo e Chebyshev do 2º tipo, respectiva-mente, no intervalo $[-1,1]$	73
Figura 3	Representação das funções peso $\lambda(x) = e^{-x}$, $\lambda(x) = e^{-\frac{x^2}{2}}$ e $\lambda(x) = e^{-x^2}$, ortogonais em relação a família de polinômios de Laguerre no intervalo $[0,\infty)$, Hermite probabilísticos e Hermite físicos no intervalo $(-\infty,\infty)$	74
Figura 4	Representação da função $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ e sua integral no intervalo $(1,7, \infty)$	94
Figura 5	Representação da função $f(x) = \frac{1,7}{2\sqrt{2\pi}} e^{-\frac{(1,7x+1,7)^2}{8}}$ a ser integrada no intervalo $(-1, 1)$	95
Figura 6	Representação da função $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ e sua integral no intervalo $(-1, 1)$	96
Figura 7	Representação geométrica do método das secantes	100
Figura 8	Representação geométrica do método de Newton-Raphson.....	104
Figura 9	Representação gráfica da densidade da variável aleatória $U = \sqrt{\hat{\sigma}^2/\sigma^2}$, dada pela equação (5.1.2), para diferentes valores de graus de liberdade ν	111

Figura 10	Representação gráfica da densidade da variável aleatória $S = \hat{\sigma}/\sigma = \sqrt{U/\nu}$, dada pela equação (5.1.4), para diferentes valores de graus de liberdade ν	113
Figura 11	Representação gráfica da densidade da variável aleatória X com distribuição t de <i>Student</i> , dada pela equação (5.2.4), para diferentes valores de graus de liberdade ν	117
Figura 12	Representação gráfica da função $g(y; z, \delta, \lambda) = g(y; 1,35, (0,2; 0,3; 0,5), (0,2; 0,5; 0,5))$, integrando da expressão (6.1.1), em relação à função densidade da normal $\phi(y)$	159
Figura 13	Representação gráfica da função $g(y; z, \delta, \lambda) = g(y; 1,35, (0,2; 0,3; 0,5), (0,2; 0,5; 0,5))$, integrando da expressão (6.1.3), comparando-a com a função densidade da normal $\phi(y)$	162
Figura 14	Representação gráfica da função $g(y; s, d, \delta, \lambda, \nu)$, integrando da expressão (6.2.4), com os parâmetros $d = 1,35$, $\delta = (0,2; 0,3; 0,5)$, $\lambda = (0,2; 0,5; 0,5)$, $\nu = 120$ para os valores de $s = (1; 1,05; 1,1)$, comparando-a com a função densidade da normal $\phi(y)$	168
Figura 15	Representação gráfica da função $g(y; s, d, \delta, \lambda, \nu)$, integrando da expressão (6.2.6), com os parâmetros $d = 1,35$, $\delta = (0,2; 0,3; 0,5)$, $\lambda = (0,2; 0,5; 0,5)$, $\nu = 120$ para os valores de $s = (1; 1,05; 1,1)$, comparando-a com a função densidade da normal $\phi(y)$	174
Figura 16	Representação da probabilidade t multivariada, função $pNDUDF$ ($d = 1,35$, $\lambda = (0,2; 0,5; 0,5)$, ν , $\delta = 0$, $n = 200$), em função dos graus de liberdade ν , em que a linha horizontal tracejada refere-se a probabilidade da normal multivariada nessas mesmas condições	188

LISTA DE TABELAS

Tabela 1	Frequências de erros tipo I (letra V) e II (letra T) em testes de hipóteses múltiplas	29
Tabela 2	Características de algumas famílias de polinômios ortogonais clássicos	72
Tabela 3	Erro de aproximação das quadraturas cujas funções $f(x)$ são de grau maior do que $2n - 1$ ou são funções não polinomiais e foram aproximadas por polinômios ortogonais de grau n , dadas por Ralston e Rabinowitz (1965), sendo $f^{(2n)}(x)$ contínua em $[a, b]$.	75
Tabela 4	Fórmulas para montar a matriz de <i>Jacobi</i> das quadraturas gaussianas usuais e o valor da constante μ_0 utilizada para o cálculo dos pesos (w_i 's) de cada quadratura.....	86
Tabela 5	Probabilidade do teste de Dunnett em função do número de pontos da quadratura n , para infinitos graus de liberdade, $d = 1,35$ e $\lambda = (0,2; 0,5; 0,5)$, obtidas pelas funções $pNDUD(d, \lambda, \delta, n)$ para o teste unilateral e $pNDBD(d, \lambda, \delta, n)$ para o teste bilateral.....	189
Tabela 6	Probabilidade do teste de Dunnett em função do número de pontos da quadratura n e dos graus de liberdade v , para $\lambda = (0,2; 0,5; 0,5)$, $d = 1,35$, obtidas pelas funções $pNDUDF(d, \lambda, v, \delta, n)$ para o teste unilateral e $pNDBDF(d, \lambda, v, \delta, n)$ para o teste bilateral.....	191
Tabela 7	Probabilidades do teste de Dunnett bilateral obtidas pela função $pNDBD(d, \lambda, \delta, n)$ e $pmvt$ do pacote “ <i>mvtnorm</i> ” do sistema <i>R</i> e em algoritmos da literatura, para $v \rightarrow \infty$, $\delta = 0$, $d = 2$ e $n = 100$ pontos na quadratura	193

Tabela 8	Probabilidade do teste de Dunnett obtidas pelas funções $pNDUDF$ para o teste unilateral e $pNDBDF$ para o bilateral e pela função $pmvt$ do pacote “ <i>mvtnorm</i> ”, para o caso central e não-central, com $d = 1,35$, $r = 18$ e $n = 32$ pontos na quadratura, variando-se os valores dos graus de liberdade v	195
Tabela 9	Probabilidades do teste bilateral de Dunnett balanceado obtidas pela função $pNDBDF$ com $n = 100$ pontos na quadratura, pela função $mvstud$ do pacote “ <i>mvtnorm</i> ” e pela função $pmvt$ do pacote “ <i>mvtnorm</i> ”, para o quantil $ d = 2,0$ em função da correlação ρ , dos graus de liberdade v e do número de comparações r , para o caso central.....	197
Tabela 10	Quantis do teste bilateral de Dunnett balanceado, no caso central, com $\rho = 0,5$, $p = 1 - \alpha = 0,95$ e $n = 32$ pontos na quadratura, em função do número de comparações r e dos graus de liberdade v ...	199
Tabela 11	Quantis do teste bilateral de Dunnett balanceado, no caso central com $\rho = 0,5$, $\delta = 0$, $p = 1 - \alpha = 0,99$ e $n = 32$ pontos na quadratura, em função do número de comparações r e dos graus de liberdade v	201
Tabela 12	Quantis do teste bilateral de Dunnett balanceado, no caso central, com $p = 1 - \alpha = 0,95$ e $n = 32$ pontos na quadratura, em função do número de comparações r , dos graus de liberdade v e do valor da correlação ρ	202
Tabela 13	Quantis do teste bilateral de Dunnett balanceado, no caso central, com $p = 1 - \alpha = 0,99$ e $n = 32$ pontos na quadratura, em função do número de comparações r , dos graus de liberdade v e do valor da correlação ρ	204

SUMÁRIO

1	INTRODUÇÃO	15
2	PROCEDIMENTOS DE COMPARAÇÕES MÚLTIPLAS - PCM's.....	20
2.1	Introdução	20
2.2	Erro tipo I em PCM's	25
2.2.1	Taxa de erro por comparação - <i>TPC</i> (<i>per – comparison error rate</i>)	29
2.2.2	Taxa de erro por experimento - <i>TPE</i> (<i>per – experiment error rate</i>)	32
2.3	Poder dos PCM's.....	33
3	TESTE DE DUNNETT.....	36
3.1	Introdução	36
3.2	Modelo estatístico e pressuposições	40
3.3	Teste de Dunnett unilateral para dados balanceados.....	41
3.4	Teste de Dunnett unilateral para dados não-balanceados	48
3.5	Teste de Dunnett bilateral para dados balanceados	54
3.6	Teste de Dunnett bilateral para dados não balanceados.....	58
3.7	Aspectos computacionais relacionados às distribuições das estatísticas do teste de Dunnett disponíveis na literatura	62
4	TÉCNICAS MATEMÁTICAS.....	67
4.1	Integração numérica.....	67
4.1.1	Quadratura gaussiana	69
4.1.2	Algoritmo de <i>Golub – Welsch</i>	76
4.1.3	Quadraturas para integrais impróprias.....	85
4.1.4	Mudança dos intervalos de integração	88
4.1.5	Quadraturas gaussianas no <i>software R</i>.....	92
4.1.6	Exemplo de resolução de integral por quadratura gaussiana com o uso do <i>software R</i>	93
4.2	Métodos numéricos para a determinação de raízes de equações.	96
4.2.1	Método do falso positivo	97
4.2.2	Método da secante.....	99
4.2.3	Método de <i>Pégaso</i> ou <i>Pégasus</i>.....	101
4.2.4	Método de Newton-Raphson	102
4.3	Método do jacobiano de transformação de variáveis.....	105

5	DISTRIBUIÇÕES DE PROBABILIDADE	108
5.1	Distribuição da variável aleatória $S = \hat{\sigma}/\sigma$.....	109
5.2	Distribuição t de <i>Student</i>	112
5.3	Distribuição t não-central univariada	116
5.3.1	Aplicações da distribuição t não-central	121
5.4	Distribuições multivariadas	122
5.4.1	Distribuições elípticas simétricas.....	123
5.4.2	Distribuição t de Student multivariada	124
5.4.3	Distribuição do máximo e do máximo do módulo da normal multivariada.....	127
5.4.4	Distribuição do máximo e do máximo do módulo da t multivariada.....	134
6	METODOLOGIA.....	141
6.1	Distribuição do máximo e do máximo do módulo da normal multivariada não-central com parâmetro de não-centralidade $\delta = \mu$.....	141
6.2	Distribuição do máximo e do máximo do módulo da t multivariada não-central com parâmetro de não-centralidade $\delta = \mu$.....	149
6.3	Detalhes da implementação das funções	155
6.3.1	Função para a distribuição do teste de Dunnett unilateral para dados não balanceados com infinitos graus de liberdade e parâmetro de não-centralidade δ, denominada de $pNDUD$.....	157
6.3.2	Função para a distribuição do teste de Dunnett bilateral para dados não balanceados com infinitos graus de liberdade e parâmetro de não-centralidade δ, denominada de $pNDBD$.....	161
6.3.3	Função para a densidade do teste de Dunnett unilateral para dados não balanceados com infinitos graus de liberdade e parâmetro de não-centralidade δ, denominada $dNDUD$	163
6.3.4	Função para a densidade do teste de Dunnett bilateral para dados não balanceados com infinitos graus de liberdade e parâmetro de não- centralidade δ, denominada $dNDBD$.....	164
6.3.5	Função para a inversa da função de distribuição do teste de Dunnett unilateral para dados não balanceados com infinitos graus de liberdade e parâmetro de não-centralidade δ, denominada $qNDUD$	165

6.3.6	Função para a inversa da função de distribuição do teste de Dunnett bilateral para dados não balanceados com infinitos graus de liberdade e parâmetro de não-centralidade δ , denominada $qNDBD$	166
6.3.7	Função para a distribuição do teste de Dunnett unilateral para dados não balanceados com finitos graus de liberdade e parâmetro de não-centralidade δ , denominada $pNDUDF$	167
6.3.8	Função para a distribuição do teste de Dunnett bilateral para dados não balanceados com finitos graus de liberdade e parâmetro de não-centralidade δ , denominada $pNDBDF$	172
6.3.9	Função para a densidade do teste de Dunnett unilateral para dados não balanceados com finitos graus de liberdade e parâmetro de não-centralidade δ , denominada $dNDUDF$	177
6.3.10	Função para a densidade do teste de Dunnett bilateral para dados não balanceados com finitos graus de liberdade e parâmetro de não-centralidade δ , denominada $dNDBDF$	181
6.3.11	Função para a inversa da função de distribuição do teste de Dunnett unilateral para dados não balanceados com finitos graus de liberdade e parâmetro de não-centralidade δ , denominada $qNDUDF$	184
6.3.12	Função para a inversa da função de distribuição do teste de Dunnett bilateral para dados não balanceados com finitos graus de liberdade e parâmetro de não-centralidade δ , denominada $qNDBDF$	185
7	RESULTADOS	187
8	CONCLUSÕES	207
	REFERÊNCIAS	209
	APÊNDICES	214
	ANEXOS	231

1 INTRODUÇÃO

Em algumas áreas de pesquisa, como na experimentação farmacológica, agrícola, biológica e de alimentos, dentre outras, é comum os pesquisadores terem interesse em comparar simultaneamente a média dos tratamentos com a média de um tratamento padrão, denominado de controle. Um exemplo típico é o dos blocos aumentados. Neste tipo de experimento são avaliados muitos tratamentos e um tratamento controle. O objetivo é verificar quais destes tratamentos em teste superam em desempenho o tratamento controle. Nesses casos, quando as amostras são aleatórias e independentes, oriundas de variáveis com distribuições normais, um teste muito utilizado é o teste de Dunnett. Ele é referido na literatura como um método de comparações múltiplas com um controle - MCC (multiple comparisons with a control).

Uma vantagem do uso de um método de comparações múltiplas com um tratamento controle é a redução do número de comparações simultâneas executadas, pois apenas consideram-se os pares em que se comparam o tratamento padrão com os demais.¹

No entanto, a grande vantagem do uso do teste de Dunnett é que se trata de um procedimento de inferências simultâneas que mantém a probabilidade do erro tipo I no nível nominal α para todo o conjunto de comparações. Isto porque o teste considera um ajuste da multiplicidade de comparações possibilitando que as taxas de erro tipo I por experimento sejam controladas em um nível de significância exato α . A estatística deste teste também é utilizada para determinar os intervalos de confiança dos verdadeiros valores das diferenças

¹ Isto significa que num procedimento de comparações múltiplas em que se envolvem $r + 1$ tratamentos, das m comparações possíveis, com $m = r(r + 1)/2$, apenas r comparações referem-se à comparação de um tratamento controle com os demais.

entre a média de cada um dos tratamentos em teste e o tratamento controle, com um valor $1 - \alpha$ de coeficiente de confiança conjunta.

As estatísticas utilizadas no teste de Dunnett têm distribuição normal ou t multivariadas. O teste de Dunnett pode ser unilateral ou bilateral, aplicado em situações balanceadas² ou não-balanceadas.³ Todos estes casos levam a diferentes distribuições de probabilidade para a estatística do teste.

A distribuição da estatística para comparar médias com a testemunha considera v graus de liberdade associados à variância do erro (variância residual amostral σ^2), com distribuição independente à das médias. Os diversos contrastes considerados simultaneamente não são independentes, devido à presença do tratamento controle como um tratamento comum à todos eles. Assim, pode haver diferentes correlações entre os contrastes dos tratamentos em teste com o tratamento controle. Como as correlações possuem infinitos valores possíveis no intervalo $[0,1]$ e as tabelas existentes são específicas para correlação 0,5, o que limitaria o uso deste teste. A razão disso é a dificuldade de se obter as probabilidades da distribuição e os valores dos quantis da estatística do teste nesses casos.

Nos casos particulares de correlação 0,5, as tabelas contêm os quantis dos testes unilateral e bilateral, para $p = 1 - \alpha = 0,95$ e $p = 1 - \alpha = 0,99$ em função dos graus de liberdade. Elas só podem ser usadas para experimentos em que todos os tratamentos, inclusive o controle, possuem o mesmo tamanho amostral.

Isso pode impossibilitar a ampla aplicação do teste ou a sua aplicação de forma incorreta, por considerar valores críticos inapropriados. Uma alternativa

² Consideram-se balanceadas, para a aplicação do teste de Dunnett, aqueles experimentos em que todos os tratamentos em teste têm o mesmo tamanho amostral, podendo diferir ou não do tamanho da amostra do tratamento controle.

³ Consideram-se não balanceadas, para a aplicação do teste de Dunnett, aqueles experimentos em que os tratamentos podem ter qualquer tamanho amostral

nos casos em que a correlação não é 0,5 é obter as probabilidades ou os quantis via rotinas computacionais.

Nesta tese, o objetivo geral é apresentar detalhadamente a teoria das distribuições t multivariadas não-centrais relacionadas ao teste de Dunnett e desenvolver, implementar e disponibilizar um pacote no *software R* específico para obter a probabilidade da distribuição do teste, com dados balanceados ou não e em testes unilaterais e bilaterais. Além disso, a rotina fornece o valor dos quantis para a estatística do teste nesses casos. Partindo das ideias propostas por Dunlap, Marx e Agamy (1981), buscou-se ampliar e melhorar a performance do algoritmo, com o uso de métodos numéricos de quadratura gaussiana ao invés da regra de Simpson para resolver as integrais.

Optou-se por considerar as estatísticas dos testes para as distribuições não-centrais por serem mais abrangentes, recaindo no caso central quando o parâmetro de não-centralidade for o vetor nulo. Além disso, as distribuições não-centrais possibilitam a ampliação do uso das rotinas para a avaliação do poder do teste, sem a necessidade de uso de simulação. Os resultados do algoritmo proposto foram comparados com os valores disponíveis na literatura e com os obtidos em funções do *software R* relativas às distribuições normal e t multivariadas não-centrais.

Os objetivos específicos deste trabalho são:

- a) compreender as estatísticas do teste de Dunnett que possuem distribuições de probabilidades t multivariadas, obtendo matematicamente e apresentando as funções densidade e de distribuição não-centrais;

- b) escrever e disponibilizar um referencial teórico didático e claro sobre o teste de Dunnett em procedimentos unilaterais e bilaterais, para dados balanceados ou não, contribuindo para o entendimento do tema;
- c) implementar funções no *software R* que calcule a probabilidade e os quantis das distribuições do teste de Dunnett para os diversos casos, usando como diferencial a resolução das integrais múltiplas por quadraturas gaussianas;
- d) disponibilizar uma biblioteca no *software R* com todas as funções mencionadas anteriormente, denominada *nCDunnett*.

Este trabalho contém sete seções, dois anexos e dois apêndices. O referencial teórico está apresentado na segunda à quinta seção. A segunda seção faz uma breve revisão sobre conceitos e características básicas referentes a teoria dos procedimentos de comparações múltiplas. A terceira seção especifica e caracteriza o teste de Dunnett unilateral e bilateral, nos casos de dados balanceados e não balanceados, esclarecendo suas pressuposições e detalhando todas as passagens das demonstrações dos teoremas das estatísticas de cada teste. A quarta seção apresenta o referencial teórico matemático utilizado na resolução do problema e na construção dos algoritmos. A quinta seção faz uma revisão das distribuições de probabilidades que serão utilizadas para obter as estatísticas do teste de Dunnett nas diferentes situações.

A metodologia do trabalho é apresentada na sexta seção. Inicialmente é feita a obtenção das estatísticas do teste de Dunnett considerando um vetor de não-centralidade δ . São obtidas as funções densidade e de distribuição do máximo e do máximo do módulo da normal multivariada não-central e da t multivariada não-central. Após estas etapas, são expostos os detalhes para a

construção dos algoritmos propostos para resolver as funções densidade, de distribuição e obter os quantis para o teste de Dunnett nas diversas situações.

A sétima seção apresenta e comenta alguns resultados de probabilidades e quantis das distribuições do teste de Dunnett obtidos pelo algoritmo implementado. O ANEXO A contém um referencial teórico resumido de aspectos básicos de inferência estatística. O ANEXO B apresenta detalhes da determinação do coeficiente de correlação entre comparações de diferenças de médias do tratamento controle e dos tratamentos em teste. No APÊNDICE A está uma aplicação do teste de Dunnett e no APÊNDICE B estão as rotinas do R implementadas.

2 PROCEDIMENTOS DE COMPARAÇÕES MÚLTIPLAS - PCM'S

Nesta seção é apresentada uma revisão de alguns conceitos e características básicas que dão suporte aos diversos procedimentos de comparações múltiplas.

Inicialmente é discutido o que significa controle da multiplicidade de comparações e são apresentadas e caracterizadas as diversas formas de mensurar as taxas de erro tipo I num PCM e suas inter-relações. É abordada uma forma de classificar os PCM's dependendo de como se dá o controle da taxa de erro tipo I e são apresentados conceitos de poder para um teste de comparações múltiplas e suas indicações de uso.

2.1 Introdução

Na análise de variância dos diversos delineamentos experimentais o teste F encontra grande aplicação na comparação de dois ou mais tratamentos.

Na comparação das médias de dois tratamentos a partir de amostras independentes das populações normais A e B , respectivamente de tamanhos n_A e n_B , a escolha do teste depende de as variâncias associadas σ_A^2 e σ_B^2 serem consideradas diferentes ou não. Portanto, inicialmente deve-se aplicar um teste estatístico a fim de comparar as estimativas de variâncias. Este teste pode ser o teste F para homogeneidade de variâncias. A partir das amostras são calculadas as variâncias amostrais não-viesadas S_A^2 e S_B^2 . Por meio da comparação das estimativas das variâncias amostrais, testa-se a hipótese nula (H_0) de igualdade das variâncias populacionais das quais as amostras foram retiradas.

A distribuição F de Fisher-Snedecor corresponde a distribuição da variável resultante da razão entre esses dois estimadores das variâncias, sob o $H_0: \sigma_A^2 = \sigma_B^2$, cuja estatística é

$$F_c = \frac{S_A^2}{S_B^2}$$

com $v_A = n_A - 1$ e $v_B = n_B - 1$ graus de liberdade para o numerador e o denominador, respectivamente (FERREIRA, 2009).

Para um teste bilateral, com $H_1: \sigma_A^2 \neq \sigma_B^2$, considerando-se a maior das estimativas das variâncias no numerador, quanto maior for F_c em relação unidade maior é o indicativo de que as variâncias populacionais diferenciam. Os valores críticos da distribuição de probabilidade F varia de acordo com o nível de significância α estabelecido, denotado por $F_{\frac{\alpha}{2}}$. Quando compara-se $F_{\frac{\alpha}{2}}$ com a estatística calculada pelo quociente das variâncias F_c , pode-se obter dois resultados:

- a) $F_c \geq F_{\frac{\alpha}{2}}$: neste caso, rejeita-se a hipótese nula de que as variâncias das populações A e B são iguais e diz-se que o teste foi significativo⁴, no nível de significância α .
- b) $F_c < F_{\frac{\alpha}{2}}$: neste caso, não se rejeita a hipótese nula de que as variâncias das populações A e B são iguais e diz-se que o teste não foi significativo.

No caso a), em que concluiu-se que as amostras provêm de populações com variâncias diferentes, para verificar se os tratamentos possuem efeitos

⁴ Pode-se indicar o nível de probabilidade associado ao resultado da pesquisa F_c , indicando que a hipótese nula pode ser rejeitada àquele nível, denominado de valor- p .

iguais pode-se aplicar o teste t para diferença de médias populacionais, no caso $\mu_A - \mu_B$, cuja estatística é dada por

$$t_c = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}} .$$

Sob $H_0: \mu_A = \mu_B$, a estatística t_c tem distribuição *aproximada t* de *Student* com ν graus de liberdade dados pela fórmula de Satterthwaite⁵:

$$\nu = \frac{\left(\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}\right)^2}{\frac{\left(\frac{S_A^2}{n_A}\right)^2}{n_A - 1} + \frac{\left(\frac{S_B^2}{n_B}\right)^2}{n_B - 1}}$$

que faz um ajuste nos graus de liberdade para as variâncias diferentes (ANDRADE; OGLIARI, 2007). Para um teste bilateral, a hipótese nula de igualdade do efeito dos tratamentos é rejeitada quando $|t_c| > \left|t_{\left(\nu, \frac{\alpha}{2}\right)}\right|$.

No caso b), quando concluiu-se que as amostras provêm de populações com variâncias iguais ou aproximadamente iguais, aplica-se o teste t com a estatística dada por

$$t_c = \frac{\bar{X}_A - \bar{X}_B}{S_{(\bar{X}_A - \bar{X}_B)}}$$

⁵ No caso do valor de ν ser fracionário, recomenda-se utilizar o número inteiro arredondado para menos. Isto é feito para garantir o controle do nível de significância desejado (ANDRADE; OGLIARI, 2007).

sendo o estimador da variância comum das duas populações ($\sigma_A^2 = \sigma_B^2 = \sigma^2$), a média ponderada das variâncias amostrais, utilizando como peso os graus de liberdade e atribuindo mais peso às amostras maiores, ou seja:

$$S_{(\bar{X}_A - \bar{X}_B)} = \sqrt{\frac{(n_A - 1) S_A^2 + (n_B - 1) S_B^2}{n_A + n_B - 2} \left[\frac{1}{n_A} + \frac{1}{n_B} \right]} .$$

Sob a hipótese nula de igualdade de médias ou efeitos dos dois tratamentos, a estatística t_c tem distribuição t de Student com $v = n_A + n_B - 2$ graus de liberdade (ANDRADE; OGLIARI, 2007). Num teste bilateral, a hipótese nula de igualdade do efeito dos tratamentos é rejeitada quando $|t_c| > \left| t_{(n_A + n_B - 2, \frac{\alpha}{2})} \right|$.

Quando comparam-se mais de dois tratamentos a técnica mais utilizada é a análise de variância - ANOVA. No caso de análise de variância para modelo fixo, testa-se a hipótese nula (H_0) de que todos os tratamentos possuem o mesmo efeito sobre a variável resposta. Os quadrados médios dos tratamentos e dos resíduos, obtidos pelo quociente das somas dos quadrados⁶ pelos graus de

⁶ Num delineamento inteiramente casualizado balanceados para avaliação de $r + 1$ tratamentos, com n repetições cada, por exemplo, tomando o fator de correção dado por

$$C = \frac{\left(\sum_{i=1}^{r+1} \sum_{a=1}^n Y_{ia} \right)^2}{(r+1)n},$$

a soma de quadrados total é dada por

$$SQ_{TOTAL} = \sum_{i=1}^{r+1} \sum_{a=1}^n Y_{ia}^2 - C,$$

a soma de quadrados dos tratamentos é dada por

$$SQ_{TRATAMENTOS} = \sum_{i=1}^{r+1} \frac{(Y_{i.})^2}{n} - C,$$

liberdade correspondentes, constituem estimativas de variâncias. O quadrado médio do resíduo é uma medida que estima a variância residual, devido por exemplo às repetições do experimento, ao delineamento experimental, ao manejo experimental, e outros fatores não controláveis que geram variações nos resultados do experimento. O quadrado médio dos tratamentos é uma medida da variação residual e da variação devido ao efeito dos tratamentos, estimando portanto a soma da variância residual com a variância dos tratamentos.

A estatística F_c é obtida por meio do quociente entre o quadrado médio dos tratamentos e o quadrado médio residual. Se não houver efeito de tratamentos os dois quadrados médios estimam a mesma variância e o valor da estatística F_c é próxima da unidade, exceto por diferenças casuais. As variações detectadas são oriundas apenas do acaso, como ocorre naturalmente nos processos de amostragens de uma mesma população.

Caso a conclusão seja pela rejeição da hipótese nula, ou seja, num teste bilateral $F_c \geq F_{\frac{\alpha}{2}}$, aceita-se que existe pelo menos um contraste entre médias que possuem efeitos significativamente diferentes de zero. Como cada amostra provém de uma população que representa o efeito de um tratamento, caso a hipótese nula seja rejeitada, pode-se concluir que existe evidência suficiente para afirmar que os tratamentos apresentam efeitos diferentes na resposta da variável em estudo. Neste caso de comparação entre mais de dois tratamentos, quando o teste F for significativo, como complemento utilizam-se testes de comparações múltiplas para detectar quais dos tratamentos possuem efeitos diferentes.

O objetivo de uma análise de comparações múltiplas é realizar comparações das médias dos tratamentos tomados dois a dois ou das médias de

em que Y_i é o total do tratamento i , e a soma de quadrado dos resíduos é dada por

$$SQ_{RES} = SQ_{TOTAL} - SQ_{TRATAMENTOS} = \sum_{i=1}^{r+1} \sum_{a=1}^n Y_{ia}^2 - \sum_{i=1}^{r+1} \frac{(Y_i.)^2}{a}.$$

dois grupos de tratamentos. O tipo de teste escolhido depende das comparações que o pesquisador está interessado e da taxa de erro tipo I que ele deseja controlar. Pode-se controlar a probabilidade de rejeitar uma hipótese verdadeira em todas as possíveis combinações dos níveis dos tratamentos tomados dois a dois, que é a chamada taxa de erro tipo I por comparação. Outra opção é controlar a probabilidade de realizar pelo menos uma inferência errada por experimento, conhecida como taxa de erro tipo I por experimento.

O desafio para qualquer procedimento de comparações múltiplas é estabelecer critérios que os testes individualmente devem satisfazer para cada uma das hipóteses (IZBICKI, 2010), de forma que mensure e controle adequadamente o efeito da multiplicidade de comparações simultâneas, por meio de uma taxa de erro tipo I global mantida no nível nominal desejado. Os diferentes PCM's usam diferentes métodos de ajuste para a multiplicidade. Na abordagem clássica busca-se a minimização de probabilidades de erros de tomada de decisão, ou seja, equilibrar as exigências contraditórias de reduzir o risco de rejeitar uma hipótese nula verdadeira (controle da taxa de erro tipo I por experimento) mantendo a probabilidade de que um efeito experimental seja detectado (manter o teste poderoso, ou seja, controlar o erro tipo II). Geralmente a ênfase na questão das comparações múltiplas é fortemente inclinada para controlar a taxa de erro tipo I por experimento.

2.2 Erro tipo I em PCM's

Numa inferência com uma única hipótese nula (H_0) a estatística de teste é escolhida de forma que controle o risco de rejeitar a hipótese nula verdadeira mantendo a taxa de erro tipo I num valor fixo (α) para determinado tamanho de amostra. Estes procedimentos de inferências individuais asseguram a taxa de erro tipo I fixa para a comparação que está sendo realizada.

Numa sequência de testes de hipóteses, em que r é o número de tratamentos, tem-se $m = r(r - 1)/2$ hipóteses nulas quando se deseja testar todos os possíveis pares de médias. As múltiplas hipóteses são testadas simultaneamente e a inferência final deve ser válida para todas as m hipóteses de interesse.

Tomando-se, por exemplo, uma sequência de testes t de diferenças de médias, aplicados à cada hipótese. Se α é a probabilidade de erro tipo I em cada um desses m testes individualmente⁷ ter-se-á um nível de confiança de $1 - \alpha$ para cada teste. Usando estatísticas de teste independentes para os m testes simultaneamente e supondo nulidade global verdadeira⁸, o nível de confiança global é de $(1 - \alpha)^m$. Assim, a probabilidade de declarar que não existem efeitos significativos, ou não rejeitar H_0 , diminui com o aumento de m . probabilidade complementar de rejeitar de forma incorreta pelo menos uma hipótese nula, erro tipo I global das inferências simultâneas⁹, é de $1 - (1 - \alpha)^m$. Logo, o erro tipo I global é substancialmente maior do que o erro tipo I inicial α de cada comparação.

Por exemplo, para $\alpha = 5\%$ e $r = 3$, tem-se $m = 3$ hipóteses nulas. Supondo que as três sejam verdadeiras, o nível de confiança global dos testes simultaneamente é $(1 - 0,05)^3 = 0,8574 \cong 86\%$. Logo, supondo independência dos testes, tem-se um erro tipo I global de $1 - (1 - 0,05)^3 = 0,1426 \cong 14\%$, muito superior ao nível nominal de significância de cada teste individual $\alpha = 5\%$.

Na Figura 1 observa-se que a probabilidade de cometer um erro tipo I global das inferências simultaneamente numa sequência de m testes t de diferença de médias sem controle para a multiplicidade, supondo hipótese nula global verdadeira, cresce com o aumento do número de hipóteses nulas. Este

⁷ α é a taxa de erro tipo I por comparação.

⁸ A nulidade global verdadeira representa a situação em que todas as médias dos tratamentos são consideradas iguais, ou seja, $H_0 : \mu_1 = \mu_2 = \dots = \mu_r$.

⁹ O erro tipo I global das inferências simultâneas é a taxa de erro tipo I por experimento.

valor aproxima-se de 1 para m grande, mais rapidamente quanto maior for o nível de significância nominal α de cada teste t individual.

Portanto, com o aumento do número de comparações aumenta também a probabilidade de obter pelo menos um resultado significativo falso. Se houver um grande número de comparações e nenhum ajuste de multiplicidade de comparações simultâneas, a decisão simultânea vai levar quase certamente a se cometer um erro tipo I global e concluir por um aparente efeito significativo quando não há nenhum. Então, é preciso estar ciente desse problema e considerar o uso de um procedimento de comparação múltipla que ajuste isso.

Assim, Hochberg e Tamhane (1987) definem um procedimento de comparação múltipla como um teste estatístico projetado para mensurar e controlar adequadamente o efeito da multiplicidade por meio de uma taxa de erro tipo I global adequada. Os diferentes testes usam diferentes métodos de ajuste para a multiplicidade.

A fim de discutir o controle do efeito para a multiplicidade de pares de comparações simultâneas, será considerada a seguinte notação:

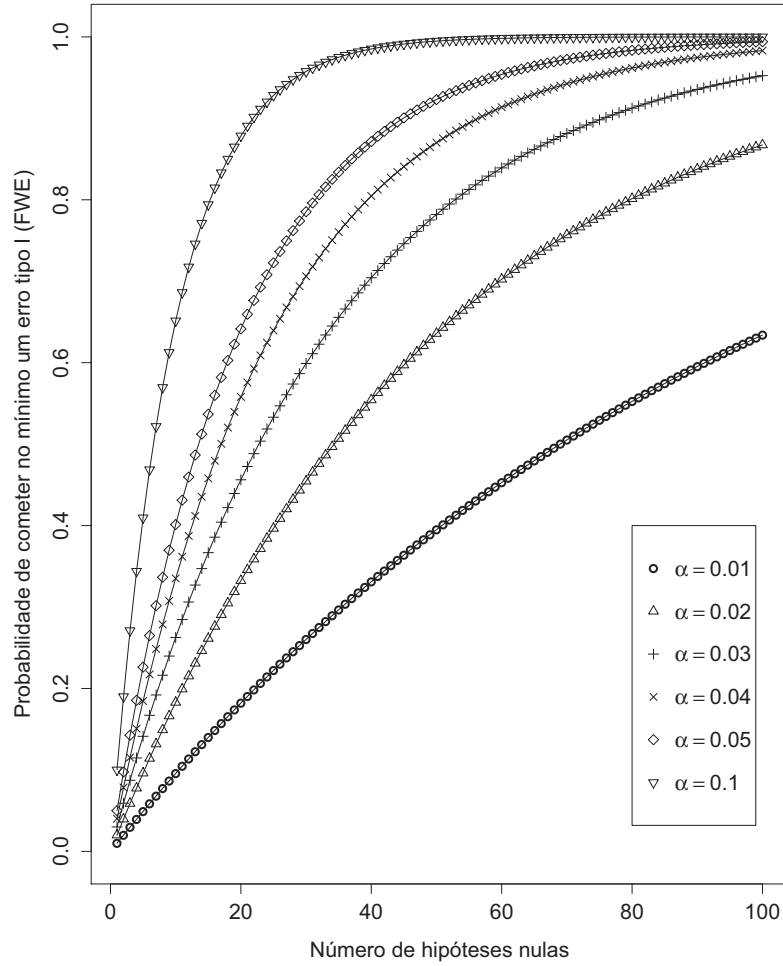


Figura 1 Probabilidade de erro tipo I global em procedimentos de comparações múltiplas sem controle de multiplicidade para diferentes níveis nominais de significância (α) e número de hipóteses nulas testadas (m)

- a) $M = \{1, \dots, m\}$ o conjunto indicador associado com as respectivas hipóteses nulas $H_{01}, H_{02}, \dots, H_{0m}$;

- b) $M_0 \subseteq M$,¹⁰ em que M_0 é o subconjunto indicador das hipóteses nulas verdadeiras, tal que $m_0 = |M_0|$ é a cardinalidade¹¹ do subconjunto M .

Na Tabela 1 (BENJAMINI; HOCHBERG, 1995; BRETZ; HOTHORN; WESTFALL, 2011), V denota o número de erros tipo I e R o número de hipóteses rejeitadas. A variável R é observável enquanto que as variáveis U , T , V e S são não observáveis. Os números m e m_0 são fixos, mas m_0 é desconhecido.

Tabela 1 Frequências de erros tipo I (letra V) e II (letra T) em testes de hipóteses múltiplas

Hipóteses Nulas	Aceitas	Rejeitadas	Total
Verdadeiras	U	V	m_0
Falsas	T	S	$m - m_0$
Total	W	R	m

Usando os m pares de comparações como uma família de inferências, a ideia é controlar a taxa de erro tipo I global no contexto de um PCM controlando a taxa para cada comparação individual.

Em testes de comparações múltiplas as taxas de erro tipo I comumente utilizadas são definidas como segue.

2.2.1 Taxa de erro por comparação - *TPC (per – comparison error rate)*

A taxa de erro tipo I por comparação, *TPC*, segundo Bretz, Hothorn e Westfall (2011) e Hochberg e Tamhane (1987) é definida como a proporção

¹⁰ Lê-se M_0 está contido em M e significa que M_0 é um subconjunto de M ou, de outra forma, que M contém M_0 .

¹¹ Número de elementos de M_0 .

esperada de erros tipo I, ou seja, de rejeições incorretas de hipóteses nulas, que na realidade são verdadeiras, dentre as m decisões, dada por

$$TPC = \frac{E(V)}{m}.$$

Se cada uma das m hipóteses nulas são testadas individualmente a um nível de significância α , a $TPC = \alpha \cdot m_0/m \leq \alpha$ é a adequada pois indica a probabilidade de erro tipo I em cada comparação (BRETZ; HOTHORN; WESTFALL, 2011), com m_0 o número de hipóteses nulas verdadeiras.

A TPC pode ser obtida por meio da simulação de experimentos sem efeito de tratamento submetidos ao PCM. Em cada experimento se determina a razão entre o número de inferências erradas ou erros tipo I (quantidade de vezes que o PCM rejeitou H_0) e o número total de inferências m . A média destes valores é a TPC (GIRARDI; CARGNELUTTI FILHO; STORCK, 2009).

Hinkelmann e Kempthorne (2007) afirmam que a TPC é utilizada em situações de comparações entre tratamentos qualitativos que possuem alguma relação entre si e, portanto, são de maior interesse. Estas comparações são muitas vezes a base para o experimento e são referidas como comparações a priori ou comparações pré-planejadas.

Como um exemplo, pode-se citar o experimento em que se deseja comparar a eficácia de dois pesticidas diferentes, A e B, ambos aplicados de duas formas diferentes: sob pulverização (A1 e B1) e em pó (A2 e B2). Um tratamento controle sem uso de pesticida, C, é incluído no experimento a fim de estabelecer a eficácia dos pesticidas em todos os aspectos. Assim, tem-se 5 tratamentos (C, A1, A2, B1 e B2), e cada tratamento é aplicado de forma aleatória para n parcelas uniformemente infestadas. O objetivo do experimento e a estrutura dos tratamentos sugerem as seguintes comparações:

- (i) controle vs. pesticidas;
- (ii) pesticida A vs. pesticida B;
- (iii) aplicação A1 vs. A2;
- (iv) aplicação B1 vs. B2;
- (v) pulverização vs. pó;
- (vi) A1 vs. B1;
- (vii) A2 vs B2.

Neste exemplo pode-se tomar diferentes conjuntos de comparações ou contrastes ortogonais, como as comparações (i), (ii), (iii) e (iv), ou então as comparações (i), (v), (vi) e (vii). De um modo geral, para r tratamentos existem muitos conjuntos de $r - 1$ contrastes ortogonais, porém apenas alguns podem apresentar resultados cuja interpretação seja de interesse prático. O ponto principal, neste caso, é que as comparações são pré-planejadas de acordo com a intenção do experimento, e não pelo resultado do experimento, e que o número de comparações é pequeno. Neste experimento não existe o interesse em inferências simultâneas das comparações.

Um teste clássico para comparar médias é o teste t , considerando as pressuposições de que as comparações foram escolhidas antes da execução do experimento, em número no máximo igual ao de graus de liberdade dos tratamentos, e que elas são ortogonais, ou seja, independentes. Porém, mesmo nestas condições, o teste t aplicado a duas ou mais comparações em um mesmo experimento não mantém o nível de significância nas comparações simultaneamente (PIMENTEL-GOMES, 2009). Com o aumento do número de comparações os múltiplos testes t sequenciais tornam-se liberais pois só controlam a taxa de erro tipo I por comparação e não por experimento (GIRARDI; CARGNELUTTI FILHO; STORCK, 2009).

2.2.2 Taxa de erro por experimento - *TPE* (*per – experiment error rate*)

Os autores Bretz; Hothorn e Westfall (2011), Hinkelmann e Kempthorne (2007) e Hochberg e Tamhane (1987) definem a taxa de erro por experimento como a probabilidade de cometer no mínimo uma inferência errada nas m inferências, dada por

$$TPE = P \{V > 0\};$$

Esta probabilidade também pode ser obtida por meio da simulação de experimentos sem efeito de tratamento submetidos ao PCM. Determina-se a razão entre o número de experimentos em que ocorreu pelo menos uma inferência errada pelo total de experimentos, quando um número muito grande de experimentos são realizados (ou simulados).

Segundo Bretz, Hothorn e Westfall (2011), quando se deseja que as m hipóteses sejam consideradas conjuntamente como uma família, onde um único erro tipo I já conduz a uma decisão errada, a taxa de erro por experimento é a indicada, pois representa a probabilidade de tomar no mínimo uma decisão incorreta. A *TPE* é a taxa de erro mais comum utilizada em ensaios múltiplos, especialmente quando o número de comparações é moderada ou quando é necessária uma forte evidência para a rejeição da hipótese nula. A *TPE* faz um ajuste para a multiplicidade de comparações considerando que é testada individualmente cada hipótese m a um nível de significância α/m . Assim, garante-se que $TPE = P \{V > 0\} \leq \alpha$.

Os testes clássicos de comparações de médias, como os testes t , Tukey, Scheffé, Scott-Knott, Bonferroni, Duncan, LSD e Student-Newman-Keuls (SNK), quando utilizados para mais de uma comparação simultânea diferem na forma de controle do erro tipo I. Ramalho, Ferreira e Oliveira (2005) afirmam

que os testes baseados na distribuição t de *Student*, como o de Duncan e o LSD, controlam a taxa de erro por comparação TPC e não controlam a taxa de erro por experimento TPE ; já os testes de Tukey e Scheffé controlam ambas as taxas de erro, preservando o nível nominal de significância (α).

Em estudo de Girardi, Cargnelutti Filho e Storck (2009) simulações relacionadas ao erro tipo I e ao poder confirmam que, em geral, o teste t controla a taxa de erro tipo I por comparação TPC e os testes de Tukey, Bonferroni e SNK controlam a taxa de erro tipo I por experimento TPE . Afirmam que o teste de Duncan não controla nenhuma destas taxa de erro e que ocorre um decréscimo na TPC dos testes Tukey, Bonferroni, Duncan e SNK e um aumento na TPE nos testes t e Duncan com o aumento do número de tratamentos.

O trabalho de Borges e Ferreira (2003) mostra que o teste de Scott-Knott controla as taxas de erro tipo I por comparação TPC , sob H_0 completa, e não controla as taxas de erro por experimento TPE para as distribuições normal, lognormal, exponencial e weibull; nesta situação, o teste de Scott-Knott é mais poderoso que os demais e é robusto à violação de normalidade.

2.3 Poder dos PCM's

Um requisito de qualquer teste estatístico é que tenha um alto poder e, assim, minimize a taxa de erro tipo II para um determinado erro tipo I estabelecido.

O poder de um teste pode ser definido de várias maneiras quando se passa de um teste de hipóteses simples para um teste de múltiplas hipóteses.

Conceitos de poder para medir o sucesso de um experimento são associados com a probabilidade de rejeitar uma hipótese nula H_{0i} , $i \in M$, quando de fato ela é falsa. Como os eventos individuais $\{H_{0i} \text{ ser rejeitado}, i \in$

$M\}$ podem ser combinados de maneiras diferentes, pode-se obter medidas diferentes de poder.

Usando a notação da Tabela 2, a seguir são apresentados alguns conceitos de poder dados por Bretz, Hothorn e Westfall (2011) e Dmitrienko, Tamhane e Bretz (2010) comumente utilizados em comparações múltiplas.

Poder individual (π_i^{ind}) é a probabilidade de rejeição correta de uma hipótese nula H_{0i} falsa, dado por

$$\pi_i^{ind} = P [\text{rejeitar } H_{0i}], i \in M_I = M \setminus M_0,$$

em que M_I é o conjunto de hipótese nulas falsas, ou seja, $M_I = M - M_0$, cuja cardinalidade é m_1 .

Poder médio (π^{md}) é o número médio esperado de rejeições corretas entre todas as hipóteses nulas falsas, isto é

$$\pi^{md} = \frac{E(S)}{m_1} = \frac{1}{m_1} \sum_{i \in M_I} \pi_i^{ind}.$$

Poder disjuntivo (π^{dis}) é a probabilidade de rejeitar no mínimo uma hipótese nula falsa, dada por

$$\pi^{dis} = P(S \geq 1),$$

em que S é o número de hipóteses nulas corretamente rejeitadas.

Poder conjuntivo (π^{con}) é a probabilidade de rejeitar todas as falsas hipóteses nula, dada por

$$\pi^{con} = P(S = m_1).$$

O poder disjuntivo e conjuntivo têm sido referidos como poder múltiplo (ou mínimo) e poder total (ou completo), respectivamente.

É importante escolher o conceito de poder adequado aos objetivos do estudo. O poder conjuntivo pode ser usado em estudos que visem detectar todos os efeitos existentes. Já o poder disjuntivo é recomendado em estudos que visem a detecção de pelo menos um efeito real, como em estudos envolvendo comparações múltiplas com um controle ou em estudos com vários pontos finais resposta, em que para demonstrar o efeito do tratamento é suficiente pelo menos um ponto resposta significativo (DMITRIENKO; TAMHANE; BRETZ, 2010). O poder individual é interessante em testes com muitas variáveis resposta secundárias enquanto que o poder médio é útil para comparar diferentes procedimentos de comparações múltiplas (BRETZ; HOTHORN; WESTFALL, 2011).

3 TESTE DE DUNNETT

Esta seção tem o propósito de apresentar as estatísticas do teste de Dunnett nas diversas situações de experimentos, esclarecendo suas pressuposições e mostrando as passagens das demonstrações dos teoremas a elas relacionadas. Esta revisão bibliográfica está baseada, principalmente, na teoria apresentada por Hsu (1999). O objetivo é disponibilizar um referencial teórico didático e claro para a comunidade científica sobre o teste de Dunnett, contribuindo para a compreensão do tema, apresentado em diversas bibliografias de forma superficial, sucinta e de difícil compreensão.

Os teoremas apresentados e demonstrados neste capítulo mostram como a confiança $100(1 - \alpha)\%$ é preservada nas r inferências simultâneas no caso do teste de Dunnett para cada situação experimental.

3.1 Introdução

O teste apresentado por Dunnett (1955), é utilizado para comparar um tratamento controle, também denominado de padrão ou testemunha com outros tratamentos de interesse em teste, que podem ser níveis de um fator de interesse.

São considerados $r + 1$ tratamentos envolvidos no teste: r tratamentos em teste e o tratamento controle denominado tratamento $r + 1$. Na literatura, testes deste tipo são referidos como MCC - Comparações Múltiplas com um Controle (HSU, 1999).

Chamando $\mu_1, \dots, \mu_r$ as médias dos r tratamentos em teste e μ_{r+1} a média do tratamento controle, no teste MCC o parâmetro de interesse é $\mu_i - \mu_{r+1}$, para $i = 1, \dots, r$, que é a diferença entre a média de cada tratamento em teste e a média do tratamento controle μ_{r+1} .

O teste de Dunnett mantém a probabilidade do erro tipo I em um valor específico α para todo o conjunto de comparações, chamado de taxa de erro tipo I por experimento (DUNLAP; MARX; AGAMY, 1981). A taxa de erro tipo I por experimento no *MCC* é o nível de significância conjunto associado com todos os r testes (MONTGOMERY, 2001). Esta probabilidade de erro tipo I é aplicada a todos os testes de hipóteses no geral, e não individualmente, como a taxa de erro por comparação.

O teste de Dunnett pode ser aplicado em diversas situações que conduzem a diferentes distribuições de probabilidade para a estatística do teste: testes bilaterais ou unilaterais; experimentos balanceados ou não balanceados. Para aplicação dos testes de Dunnett, experimentos balanceados são considerados aqueles experimentos em que todos os tratamentos têm o mesmo número de repetições ($n_i = n_{r+1} = n$ para todo i) ou todos os tratamentos em teste têm igual número de repetições, mesmo que diferente das repetições do tratamento controle ($n_i = n$ para todo i e $n \neq n_{r+1}$). São considerados experimentos não balanceados os que envolvem tratamentos com quaisquer número de repetições (n_i para $i = 1, \dots, r, r + 1$), sem nenhuma restrição.

A distribuição da estatística deste teste também é utilizada para determinar os intervalos de confiança dos verdadeiros valores das diferenças entre a média de cada um dos tratamentos em teste e o tratamento controle, com um valor $1 - \alpha$ de coeficiente de confiança conjunta.

No entanto, ao considerar simultaneamente as comparações, observa-se que elas não são independentes, já que todas têm em comum o tratamento controle.

A i -ésima correlação entre as diferenças do tratamento controle e dos tratamentos em teste, representada por ρ_i , $i = 1, 2, \dots, r$, é dada por:

$$\rho_i = \frac{\sigma_{r+1}^2}{\sigma_{r+1}^2 + \sigma_i^2},$$

em que n_i é o número de repetições do i -ésimo tratamento em teste e n_{r+1} é o número de repetições do tratamento controle, $\sigma_{r+1}^2 = \sigma_{\mu_{r+1}}^2 = \sigma^2/n_{r+1}$ é a variância da média do tratamento controle e $\sigma_i^2 = \sigma_{\mu_i}^2 = \sigma^2/n_i$ é a variância da média de cada tratamento em teste. Como é considerado o caso homocedástico, a variância comum é dada por σ^2 . A demonstração de como se chega a essa fórmula está no APÊNDICE B. Simplificando a expressão anterior, chega-se a

$$\rho_i = \frac{n_i}{n_i + n_{r+1}}.$$

Um caso particular ocorre quando $n_i = n_{r+1}$ para todo i , denominado de equicorrelacionado, em que se tem apenas um parâmetro $\rho = 0,5$.

A distribuição da estatística para comparar r médias com a testemunha, que será denominada de d , considera v graus de liberdade associados à variância do erro (variância residual amostral $\hat{\sigma}^2$), com distribuição independente à das médias e diferentes correlações ρ_i entre os contrastes dos tratamentos em teste i e o tratamento controle $r + 1$. Como as correlações possuem infinitos valores possíveis no intervalo $[0,1]$, a limitação para o uso do teste de Dunnett é a dificuldade em se obter os valores da estatística do teste.

Nos casos de correlação entre comparações iguais a 0,5, encontram-se disponíveis na literatura tabelas das estatísticas dos testes de Dunnett. Geralmente são tabelas que contém os quantis dos testes unilateral e bilateral, para $p = 1 - \alpha = 0,95$ e $p = 1 - \alpha = 0,99$ em função dos graus de liberdade. Porém, estas tabelas só podem ser usadas para experimentos em que todos os

tratamentos, inclusive o controle, possuem o mesmo tamanho amostral. Isso pode impossibilitar a ampla aplicação do teste ou levar ao uso incorreto destas estatísticas tabeladas, quando essa condição não for atendida.

Dunnett (1964) publicou novas tabelas acrescentando um fator de correção para o quantil tabelado caso a correlação estivesse no intervalo $[0,125; 0,5)$. Para cada valor da tabela foi determinado um fator que representa 1,5 vezes o aumento percentual do valor crítico para $\rho = 0,125$ em relação a $\rho = 0,5$. Multiplicando este fator pelo termo $[(1 - 2\rho)/(1 - \rho)]$ ou por $[1 - n_i/n_{r+1}]$ quando as variâncias são homogêneas mas o número de observações dos tratamento controle e dos tratamentos em teste são diferentes, obtém-se o aumento percentual necessário ao valor crítico tabelado para uma correlação $\rho \in [0,125; 0,5)$. No entanto, este é um método de ajuste aproximado para a estatística do teste. Além disso, não fornece valores para toda a amplitude de correlações possíveis no intervalo $[0,1]$. Para a adequada aplicação do teste de Dunnett é necessário que facilmente se obtenha o valor exato da estatística do teste, independente do valor da correlação.

Dunnett (1964) e outros autores (DEAN; VOSS, 1999; DUNLAP; MARX; AGAMY, 1981; MONTGOMERY, 2001) afirmaram que o teste de Dunnett é mais poderoso quando o tratamento controle tem um tamanho amostral (n_{r+1}) maior do que os tamanhos amostrais dos tratamentos em teste (n_i). Dunnett (1964), o que ocorre é que apesar do pequeno aumento do valor crítico tabelado (quantil) em relação ao caso equicorrelacionado, existe um ganho relativo devido ao maior número de observações do tratamento controle, o que resulta na diminuição do erro padrão entre a diferença dos tratamentos que aparece no denominador da estatística calculada.

3.2 Modelo estatístico e pressuposições

Tomando $Y_{i1}, \dots, Y_{in_i}$ uma amostra aleatória, de tamanho n_i , do i -ésimo tratamento. As pressuposições assumidas em todos os casos do teste de Dunnett são que:

- a) cada observação Y_{ia} , da amostra aleatória do i -ésimo tratamento é normal e independentemente distribuída para $a = 1, 2, \dots, n_i$;
- b) a amostra aleatória de cada tratamento é normal e independentemente distribuída para os diferentes valores de $i = 1, 2, \dots, r + 1$;
- c) as variância das amostras aleatórias são homogêneas.

O modelo estatístico para os dados é

$$Y_{ia} = \mu_i + \epsilon_{ia}, \quad i = 1, \dots, r + 1; \quad a = 1, \dots, n_i,$$

em que

μ_i é a média do i -ésimo tratamento;

$\epsilon_{11}, \dots, \epsilon_{(r+1)n_i}$ são os erros, aleatórios, independentes e identicamente distribuídos como uma normal com média 0 e variância σ^2 desconhecida.

As pressuposições do modelo são aditividade e independência dos efeitos e erros independentes, não correlacionados e com distribuição $N(0, \sigma^2)$.

Usando a notação

$$\hat{\mu}_i = \bar{Y}_i = \sum_{a=1}^{n_i} \frac{Y_{ia}}{n_i}$$

Para a média amostral do i -ésimo tratamento, e

$$\hat{\sigma}^2 = QME = \frac{\sum_{i=1}^{r+1} \sum_{a=1}^{n_i} (Y_{ia} - \bar{Y}_i)^2}{\nu} \quad (3.2.3)$$

para a variância residual amostral combinada com ν graus de liberdade.

Por exemplo, no caso de um delineamento inteiramente casualizado, sem perda de parcela, tem-se

$$\nu = \sum_{i=1}^{r+1} (n_i - 1) \text{ graus de liberdade.}$$

A variável aleatória

$$D_i = \frac{\hat{\mu}_i - \hat{\mu}_{r+1}}{\sigma \sqrt{\frac{1}{n_i} + \frac{1}{n_i + 1}}} \text{ para } i = 1, 2, \dots, r$$

é a estatística do teste de Dunnett unilateral, e $|D_i|$ a estatística para o teste bi-lateral. As distribuições do máximo e do máximo do módulo dessas r variáveis aleatórias são apresentadas na seção 5.

3.3 Teste de Dunnett unilateral para dados balanceados

Considerando que todos os tratamentos possuem o mesmo tamanho amostral, ou seja, $n_i = n$ para todo i , a correlação em cada contraste é constante e igual a $1/2$ ($\rho = 0,5$). A expressão para o erro padrão dos estimadores de $\hat{\mu}_i - \hat{\mu}_{r+1}$ é dada por

$$\hat{\sigma}\sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}} = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{1}{n}} = \hat{\sigma}\sqrt{\frac{2}{n}}.$$

De acordo com Hsu (1999), os testes unilaterais podem ser divididos em dois casos particulares. No primeiro, deseja-se fazer inferência sobre a média da variável resposta analisada em cada tratamento em teste que supere a média do tratamento controle, ou seja,

$$\mu_i - \mu_{r+1} > \hat{\mu}_i - \hat{\mu}_{r+1} - d\hat{\sigma}\sqrt{\frac{2}{n}}, \quad i = 1, \dots, r,$$

obter o limite inferior¹² de um intervalo unilateral de $100(1 - \alpha)\%$ de confiança, em que d é a solução da equação

$$\int_0^\infty \int_{-\infty}^{+\infty} [\Phi(z + \sqrt{2}sd)]^r \phi(z) dz f(s; \nu) ds = 1 - \alpha, \quad (3.3.1)$$

sendo

• $\phi(\cdot)$ e $\Phi(\cdot)$ a função densidade e de distribuição da normal padrão, respectivamente:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad \text{e} \quad \Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw;$$

• $f(s; \nu)$ a função densidade de $S = \hat{\sigma}/\sigma$, dada pela expressão:

$$f(s; \nu) = \frac{\nu^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2-1}} s^{\nu-1} e^{-\nu s^2/2}, \quad s \geq 0$$

¹² Se o lado direito da desigualdade for negativa, infere-se que o tratamento controle tem uma menor resposta do que o i -ésimo tratamento em teste.

No segundo caso, deseja-se fazer inferência sobre a média da variável resposta analisada em cada tratamento em teste que seja inferior ao tratamento controle. Assim, a inferência recai no limite superior¹³ do intervalo de confiança de $100(1 - \alpha)\%$, definido pela desigualdade em que d é a solução da equação (3.3.1).

$$\mu_i - \mu_{r+1} < \hat{\mu}_i - \hat{\mu}_{r+1} + d\hat{\sigma}\sqrt{\frac{2}{n}}, \quad i = 1, \dots, r,$$

Observa-se que a estatística d depende apenas do nível nominal α estabelecido para as inferências, do número r de tratamentos em teste e dos graus de liberdade ν da variável S . O valor crítico d é o quantil da distribuição do máximo da t multivariada, com correlação comum $\rho = 0,5$ e ν graus de liberdade, correspondente a uma probabilidade de erro tipo I igual a α (DEAN; VOSS, 1999).

As funções de distribuição da estatística d são apresentadas no quinto capítulo, especificamente nas expressões (5.4.6) para $\nu \rightarrow \infty$, caso em que a distribuição do máximo da t multivariada recai na distribuição do máximo da normal multivariada, e (5.4.14) nos demais casos. Naquelas expressões a correlação ρ_i está representada por λ_i^2 e foi deduzida para dados não-balanceados. Na situação analisada agora, a correlação é constante e igual a $0,5$, ou seja, $\lambda_i^2 = \rho_i = 0,5$ para todo i .

Teorema 3.3.1 *Definindo a estatística do teste unilateral para dados balanceados d como a solução da equação*

$$\int_0^\infty \int_{-\infty}^{+\infty} [\Phi(z + \sqrt{2}sd)]^r \phi(z) dz f(s; \nu) ds = 1 - \alpha,$$

¹³ Se o lado direito da desigualdade for negativa, infere-se que o tratamento controle tem uma maior resposta do que o i -ésimo tratamento em teste.

tem-se que

$$P \left\{ \mu_i - \mu_{r+1} > \hat{\mu}_i - \hat{\mu}_{r+1} - d\hat{\sigma}\sqrt{\frac{2}{n}}, \quad i = 1, \dots, r \right\} =$$

$$P \left\{ \mu_i - \mu_{r+1} < \hat{\mu}_i - \hat{\mu}_{r+1} + d\hat{\sigma}\sqrt{\frac{2}{n}}, \quad i = 1, \dots, r \right\} = 1 - \alpha.$$

Demonstração:

Partindo do pressuposto de que a probabilidade nas r inferências simultâneas é preservada e considerando $Z_1, \dots, Z_r$ variáveis aleatórias normais padrão independentes e identicamente distribuídas, pode-se escrever

$$P \left\{ \mu_i - \mu_{r+1} > \hat{\mu}_i - \hat{\mu}_{r+1} - d\hat{\sigma}\sqrt{\frac{2}{n}}, \quad i = 1, \dots, r \right\} = 1 - \alpha$$

$$P \left\{ (\hat{\mu}_i - \hat{\mu}_{r+1}) - (\mu_i - \mu_{r+1}) < d\frac{\hat{\sigma}}{\sqrt{n}}\sqrt{2}, \quad i = 1, \dots, r \right\} = 1 - \alpha$$

$$P \left\{ \frac{(\hat{\mu}_i - \mu_i) - (\hat{\mu}_{r+1} - \mu_{r+1})}{\frac{\hat{\sigma}}{\sqrt{n}}} < d\sqrt{2}, \quad i = 1, \dots, r \right\} = 1 - \alpha.$$

Multiplicando o membro esquerdo da desigualdade por σ/σ tem-se

$$P \left\{ \left(\frac{(\hat{\mu}_i - \mu_i)}{\sigma/\sqrt{n}} - \frac{(\hat{\mu}_{r+1} - \mu_{r+1})}{\sigma/\sqrt{n}} \right) \frac{\sigma}{\hat{\sigma}} < d\sqrt{2}, \quad i = 1, \dots, r \right\} = 1 - \alpha.$$

Tomando a expressão à esquerda da igualdade e condicionando ao conhecimento de $\hat{\sigma}^2$, com $s = \hat{\sigma}/\sigma$, chega-se a probabilidade condicional

$$P \left[\left\{ \left(\frac{(\hat{\mu}_i - \mu_i)}{\sigma/\sqrt{n}} - \frac{(\hat{\mu}_{r+1} - \mu_{r+1})}{\sigma/\sqrt{n}} \right) < sd\sqrt{2}, i = 1, \dots, r \right\} | \hat{\sigma} \right].$$

Multiplicando essa probabilidade pela função densidade de $S, f(s; \nu)$, tem-se a probabilidade conjunta

$$\left[P \left\{ \left(\frac{(\hat{\mu}_i - \mu_i)}{\sigma/\sqrt{n}} - \frac{(\hat{\mu}_{r+1} - \mu_{r+1})}{\sigma/\sqrt{n}} \right) < sd\sqrt{2}, i = 1, \dots, r \right\} | \hat{\sigma} \right] f(s; \nu).$$

Integrando essa expressão em relação a s , com $s \geq 0$, chega-se à função de probabilidade marginal

$$\int_0^\infty \left[P \left\{ \left(\frac{(\hat{\mu}_i - \mu_i)}{\sigma/\sqrt{n}} - \frac{(\hat{\mu}_{r+1} - \mu_{r+1})}{\sigma/\sqrt{n}} \right) < sd\sqrt{2}, i = 1, \dots, r \right\} | \hat{\sigma} \right] \times f(s; \nu) ds = 1 - \alpha, \quad (3.3.2)$$

Observa-se que essa expressão é uma esperança condicional e pode ser representada por

$$E \left[P \left\{ \left(\frac{(\hat{\mu}_i - \mu_i)}{\sigma/\sqrt{n}} - \frac{(\hat{\mu}_{r+1} - \mu_{r+1})}{\sigma/\sqrt{n}} \right) < sd\sqrt{2}, i = 1, \dots, r \right\} | \hat{\sigma} \right].$$

Essa definição de função de distribuição e esperança condicional pode ser melhor compreendida revisando alguns conceitos básicos de inferência estatística apresentados no APÊNDICE A.

De modo análogo, considerando-se as r variáveis aleatórias $Z_i = (\hat{\mu}_i - \mu_i) / (\sigma/\sqrt{n})$ e $Z = (\hat{\mu}_{r+1} - \mu_{r+1}) / (\sigma/\sqrt{n})$ um escore normal padrão e condicionando ao conhecimento de $\hat{\mu}_{r+1}$, pode-se rearranjar e escrever a probabilidade interna da expressão anterior como a probabilidade condicional

$$\left[P \left\{ Z_i < sd\sqrt{2} + z, i = 1, \dots, r \right\} | \hat{\mu}_{r+1} \right].$$

Multiplicando esta probabilidade pela função densidade da normal padrão Z , $\phi(z)$, tem-se a probabilidade conjunta

$$\left[P \left\{ Z_i < sd\sqrt{2} + z, \quad i = 1, \dots, r \right\} | \hat{\mu}_{r+1} \right] \phi(z),$$

que integrada em relação a z resulta na função de probabilidade marginal

$$\int_{-\infty}^{\infty} \left[P \left\{ Z_i < sd\sqrt{2} + z, \quad i = 1, \dots, r \right\} | \hat{\mu}_{r+1} \right] \phi(z) dz = 1 - \alpha. \quad (3.3.3)$$

A expressão anterior é, na verdade, a

$$E \left[P \left\{ Z_i < sd\sqrt{2} + Z, \quad i = 1, \dots, r \right\} | \hat{\mu}_{r+1} \right].$$

Mas $P(Z_i < sd\sqrt{2} + z)$ é a probabilidade acumulada de uma variável aleatória normal padrão avaliada no ponto $sd\sqrt{2} + z$. Como os Z_i 's, $i = 1, \dots, r$ são independentemente distribuídos, tem-se que

$$P(Z_i < sd\sqrt{2} + z) = \Phi(sd\sqrt{2} + z) \quad ,$$

ou seja

$$P(Z_1 < sd\sqrt{2} + z, Z_2 < sd\sqrt{2} + z, \dots, Z_r < sd\sqrt{2} + z) = [\Phi(sd\sqrt{2} + z)]^r.$$

Assim, a expressão (3.3.3) fica

$$\int_{-\infty}^{\infty} [\Phi(sd\sqrt{2} + z)]^r \phi(z) dz = 1 - \alpha. \quad (3.3.4)$$

Substituindo (3.3.4) em (3.3.2) chega-se a

$$\int_0^\infty \left[\int_{-\infty}^{+\infty} [\Phi(z + \sqrt{2}sd)]^r \phi(z) dz \right] f(s; \nu) ds = 1 - \alpha,$$

em que d é a solução da equação quando o valor de α for fixado, como se queria provar.

Prova-se, utilizando procedimento similar, para a expressão

$$P \left\{ \mu_i - \mu_{r+1} < \hat{\mu}_i - \hat{\mu}_{r+1} + d\hat{\sigma} \sqrt{\frac{2}{n}}, \quad i = 1, \dots, r \right\} = 1 - \alpha,$$

que o resultado é o mesmo.

No caso de experimentos balanceados em que os tratamentos em teste possuem igual tamanho amostral mas diferente em relação ao tamanho amostral do tratamento controle, ou seja, $n_i = n$ é o mesmo para todo $i = 1, \dots, r$ mas $n \neq n_{r+1}$, a correlação em cada contraste é constante e dada por $\rho = n/(n + n_{r+1})$.

Portanto, tem-se um caso especial do teste unilateral para dados balanceados com a correlação constante ρ qualquer.

A expressão para o erro padrão do estimador da diferença de médias entre o i -ésimo tratamento em teste e o tratamento controle neste caso é

$$\hat{\sigma} \sqrt{\frac{1}{n} + \frac{1}{n_{r+1}}}. \quad (3.3.5)$$

Portanto

$$P \left\{ \mu_i - \mu_{r+1} < \hat{\mu}_i - \hat{\mu}_{r+1} + d\hat{\sigma} \sqrt{\frac{1}{n} + \frac{1}{n_{r+1}}}, \quad i = 1, \dots, r \right\} = \int_0^\infty \int_{-\infty}^{+\infty} \left[\Phi \left(\frac{\sqrt{\rho}z + sd}{\sqrt{1-\rho}} \right) \right]^r \phi(z) dz f(s; \nu) ds = 1 - \alpha.$$

$$\text{com } \rho = \frac{n}{n + n_{r+1}}$$

3.4 Teste de Dunnett unilateral para dados não-balanceados

Neste caso, como todos os $r + 1$ tratamentos podem ter diferentes tamanhos amostrais, a correlação entre os contrastes varia e é dada por $\rho_i = n_i/(n_i + n_{r+1})$ para cada $i = 1, \dots, r$. Portanto, tem-se um vetor de correlações denominado de $\lambda_{r \times 1}$. A expressão para o erro padrão do estimador do contraste entre o i -ésimo tratamento em teste com o controle é dada por

$$\hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}.$$

A lógica deste teste de Dunnett é a mesma apresentada na seção anterior, considerando agora a correlação diferente em cada comparação. Então, o limite inferior simultâneo de confiança para a diferença entre a média de cada tratamento em teste μ_i e a média do tratamento controle μ_{r+1} é dada por

$$\mu_i - \mu_{r+1} > \hat{\mu}_i - \hat{\mu}_{r+1} - d\hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}, \quad i = 1, \dots, r,$$

em que $d = d_{\lambda, \alpha, \nu}$ é a solução da equação

$$\int_0^\infty \int_{-\infty}^{+\infty} \prod_{i=1}^r \left[\Phi \left(\frac{\sqrt{\rho_i} z + s d}{\sqrt{1 - \rho_i}} \right) \right] \phi(z) dz f(s; \nu) ds = 1 - \alpha, \quad (3.4.1)$$

com

$$\rho_i = \left(1 + \frac{n_{r+1}}{n_i} \right)^{-1} = \frac{n_i}{n_i + n_{r+1}}, \quad i = 1, \dots, r.$$

De forma análoga, o limite superior simultâneo de confiança para a diferença entre a média de cada tratamento em teste μ_i e a média do tratamento controle μ_{r+1} é dada por

$$\mu_i - \mu_{r+1} < \hat{\mu}_i - \hat{\mu}_{r+1} + d\hat{\sigma}\sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}, \quad i = 1, \dots, r.$$

Observa-se que agora o valor crítico d depende de α , r e ν e das razões entre os tamanhos amostrais $n_{r+1}/n_1, \dots, n_{r+1}/n_r$. O valor crítico d é o quantil da distribuição do máximo da t multivariada com correlação dada por λ e ν graus de liberdade, correspondente a uma probabilidade de erro tipo I igual a α (DEAN; VOSS, 1999). Portanto, não é possível tabular d de uma forma geral. Somente pode-se fazer uso de um programa de computador para calcular d dependendo do vetor referente aos tamanhos amostrais $\mathbf{n} = (n_1, \dots, n_{r+1})$ dos dados analisados que afetarão a correlação $\lambda = (\rho_1, \dots, r)$.

As funções de distribuição da estatística d são apresentadas no quinto capítulo, especificamente nas expressões (5.4.6) para $\nu = \infty$, caso em que a distribuição do máximo da t multivariada recai na distribuição do máximo da normal multivariada, e (5.4.14) nos demais casos. Naquelas expressões a correlação ρ_i está representada por λ_i^2 com $i = 1, \dots, r$.

Teorema 3.4.1 *Definindo a estatística do teste unilateral para dados não balanceados d como a solução da equação:*

$$\int_0^\infty \int_{-\infty}^{+\infty} \prod_{i=1}^r \left[\Phi \left(\frac{\sqrt{\rho_i}z + sd}{\sqrt{1 - \rho_i}} \right) \right] \phi(z) dz f(s; \nu) ds = 1 - \alpha,$$

com $\rho_i = \frac{n_i}{n_i + n_{r+1}}$ tem-se que

$$P \left\{ \mu_i - \mu_{r+1} > \hat{\mu}_i - \hat{\mu}_{r+1} - d\hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}, \quad i = 1, \dots, r \right\} =$$

$$P \left\{ \mu_i - \mu_{r+1} < \hat{\mu}_i - \hat{\mu}_{r+1} + d\hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}, \quad i = 1, \dots, r \right\} = 1 - \alpha.$$

Demonstração:

Partindo do pressuposto de que a probabilidade nas r inferências simultâneas é preservada e considerando $Z_1, \dots, Z_r$ variáveis aleatórias normal padrão independentes e identicamente distribuídas, então

$$P \left\{ \mu_i - \mu_{r+1} > \hat{\mu}_i - \hat{\mu}_{r+1} - d\hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}, \quad i = 1, \dots, r \right\} = 1 - \alpha$$

$$P \left\{ (\hat{\mu}_i - \mu_i) - (\hat{\mu}_{r+1} - \mu_{r+1}) < d\hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}, \quad i = 1, \dots, r \right\} = 1 - \alpha.$$

Multiplicando toda a desigualdade por $\sqrt{n_i}/\sigma$ tem-se

$$P \left\{ \frac{\sqrt{n_i}(\hat{\mu}_i - \mu_i)}{\sigma} - \frac{\sqrt{n_i}(\hat{\mu}_{r+1} - \mu_{r+1})}{\sigma} < d \frac{\hat{\sigma}}{\sigma} \sqrt{1 + \frac{n_i}{n_{r+1}}}, \quad i = 1, \dots, r \right\} = 1 - \alpha.$$

Multiplicando o termo $\frac{\sqrt{n_i}(\hat{\mu}_{r+1} - \mu_{r+1})}{\sigma}$ por $\sqrt{\frac{n_{r+1}}{n_{r+1}}}$ e rearranjando o resultado obtido tem-se

$$P\left\{\frac{(\hat{\mu}_i - \mu_i)}{\frac{\sigma}{\sqrt{n_i}}} - \sqrt{\frac{n_i}{n_{r+1}}} \frac{(\hat{\mu}_{r+1} - \mu_{r+1})}{\frac{\sigma}{\sqrt{n_{r+1}}}} < d \frac{\hat{\sigma}}{\sigma} \sqrt{1 + \frac{n_i}{n_{r+1}}}, \quad i = 1, \dots, r\right\} = 1 - \alpha.$$

Como

$$\begin{aligned} \rho_i &= \frac{n_i}{n_i + n_{r+1}} & \Rightarrow \frac{n_i + n_{r+1}}{n_i} &= \frac{1}{\rho_i} & \Rightarrow 1 + \frac{n_{r+1}}{n_i} &= \frac{1}{\rho_i} \\ \Rightarrow \frac{n_{r+1}}{n_i} &= \frac{1}{\rho_i} - 1 & \Rightarrow \frac{n_{r+1}}{n_i} &= \frac{1 - \rho_i}{\rho_i} & \Rightarrow \frac{\rho_i}{1 - \rho_i} &= \frac{n_i}{n_{r+1}} \end{aligned}$$

tem-se

$$\sqrt{\frac{n_i}{n_{r+1}}} = \frac{\sqrt{\rho_i}}{\sqrt{1 - \rho_i}}. \quad (3.4.2)$$

De modo similar

$$1 + \frac{\rho_i}{1 - \rho_i} = 1 + \frac{n_i}{n_{r+1}} \quad \Rightarrow \frac{1}{1 - \rho_i} = 1 + \frac{n_i}{n_{r+1}}$$

tem-se

$$\sqrt{1 + \frac{n_i}{n_{r+1}}} = \frac{1}{\sqrt{1 - \rho_i}}. \quad (3.4.3)$$

Usando os igualdades (3.4.2) e (3.4.3) e condicionando ao conhecimento de $\hat{\sigma}^2$, com $s = \hat{\sigma}/\sigma$, sendo $s \geq 0$, tem-se uma probabilidade condicional; essa probabilidade multiplicada pela função densidade de S, $f(s; v)$, resulta na

probabilidade conjunta de S e dos r contrastes; integrando-se essa probabilidade em relação a s , com $s \geq 0$, chega-se a função de probabilidade marginal

$$\int_0^\infty \left[P \left\{ \frac{(\hat{\mu}_i - \mu_i)}{\frac{\sigma}{\sqrt{n_i}}} - \frac{\sqrt{\rho_i}}{\sqrt{1 - \rho_i}} \frac{(\hat{\mu}_{r+1} - \mu_{r+1})}{\frac{\sigma}{\sqrt{n_{r+1}}}} < \frac{sd}{\sqrt{1 - \rho_i}}, \text{ com } i = 1, \dots, r \right\} | \hat{\sigma} \right] \times f(s; \nu) ds = 1 - \alpha. \quad (3.4.4)$$

Essa expressão é equivalente a esperança condicional

$$E \left[P \left\{ \frac{(\hat{\mu}_i - \mu_i)}{\frac{\sigma}{\sqrt{n_i}}} - \frac{\sqrt{\rho_i}}{\sqrt{1 - \rho_i}} \frac{(\hat{\mu}_{r+1} - \mu_{r+1})}{\frac{\sigma}{\sqrt{n_{r+1}}}} < \frac{sd}{\sqrt{1 - \rho_i}}, \quad i = 1, \dots, r \right\} \right].$$

De modo análogo, considerando $Z_i = (\hat{\mu}_i - \mu_i)/(\sigma/\sqrt{n_i})$ e $z = (\hat{\mu}_{r+1} - \mu_{r+1})/(\sigma/\sqrt{n_{r+1}})$ um escore normal padrão, condicionando ao conhecimento de $\hat{\mu}_{r+1}$, multiplicando pela função densidade de Z e integrando em relação a z , pode-se escrever a probabilidade interna da expressão (3.4.4) como a probabilidade marginal

$$\int_{-\infty}^\infty \left[P \left\{ Z_i < \frac{\sqrt{\rho_i}z + sd}{\sqrt{1 - \rho_i}}, \quad i = 1, \dots, r \right\} | \hat{\mu}_{r+1} \right] \phi(z) dz, \quad (3.4.5)$$

que é a esperança condicional

$$E \left[P \left\{ Z_i < \frac{\sqrt{\rho_i}Z + sd}{\sqrt{1 - \rho_i}}, \quad i = 1, \dots, r \right\} | \hat{\mu}_{r+1} \right].$$

Mas $P\left(Z_i < \frac{\sqrt{\rho_i}z + sd}{\sqrt{1-\rho_i}}\right)$ é a probabilidade acumulada de uma variável aleatória normal padrão avaliada no ponto $(\sqrt{\rho_i}z + sd) / \sqrt{1-\rho_i}$ para cada $i = 1, \dots, r$. Como os Z_i 's são independentemente distribuídos, a função de distribuição $\Phi\left(\frac{\sqrt{\rho_i}z + sd}{\sqrt{1-\rho_i}}\right)$ é dada por

$$P\left(Z_1 < \frac{\sqrt{\rho_1}z + sd}{\sqrt{1-\rho_1}}, \dots, Z_r < \frac{\sqrt{\rho_r}z + sd}{\sqrt{1-\rho_r}}\right) = \prod_{i=1}^r \left[\Phi\left(\frac{\sqrt{\rho_i}z + sd}{\sqrt{1-\rho_i}}\right) \right],$$

Fazendo com que (3.4.5) resulte em

$$\int_{-\infty}^{+\infty} \prod_{i=1}^r \left[\Phi\left(\frac{\sqrt{\rho_i}z + sd}{\sqrt{1-\rho_i}}\right) \right] \phi(z) dz. \quad (3.4.6)$$

Substituindo (3.4.6) em (3.4.4) chega-se a

$$\int_0^\infty \left[\int_{-\infty}^{+\infty} \prod_{i=1}^r \left[\Phi\left(\frac{\sqrt{\rho_i}z + sd}{\sqrt{1-\rho_i}}\right) \right] \phi(z) dz \right] f(s; \nu) ds = 1 - \alpha.$$

em que d é a solução da equação quando o valor de α for fixado, como se queria provar.

De forma similar, prova-se o mesmo para

$$P\left\{ \mu_i - \mu_{r+1} < \hat{\mu}_i - \hat{\mu}_{r+1} + d\hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}, \quad i = 1, \dots, r \right\} = 1 - \alpha.$$

3.5 Teste de Dunnett bilateral para dados balanceados

O teste de Dunnett bilateral é utilizado para determinar quais tratamentos em teste diferem do tratamento controle. No caso de experimentos balanceados, inicialmente são consideradas amostras de mesmo tamanho n para todos os $i = 1, \dots, r+1$ tratamentos.

Este teste fornece o seguinte intervalo simultâneo de $100(1 - \alpha)\%$ de confiança para a diferença entre a média de cada tratamento em teste μ_i e a média do tratamento controle μ_{r+1} :

$$\mu_i - \mu_{r+1} \in \hat{\mu}_i - \hat{\mu}_{r+1} \pm |d| \hat{\sigma} \sqrt{\frac{2}{n}}, \quad i = 1, \dots, r,$$

em que $|d|$ é a solução da equação:

$$\int_0^\infty \int_{-\infty}^{+\infty} [\Phi(z + \sqrt{2}|d|s) - \Phi(z - \sqrt{2}|d|s)]^r \phi(z) dz f(s; \nu) ds = 1 - \alpha. \quad (3.5.1)$$

Pode-se observar que $|d|$ depende apenas de α , r e ν . O valor crítico $|d|$ agora é o valor do quantil da distribuição do máximo do módulo da t -multivariada com correlação comum $\rho = 0,5$ e ν graus de liberdade, correspondente a uma probabilidade de erro tipo I igual a α (DEAN; VOSS, 1999).

As funções de distribuição da estatística $|d|$ são apresentadas no quinto capítulo, especificamente nas expressões (5.4.9) para $\nu = \infty$, caso em que a distribuição do máximo do módulo da t multivariada recai na distribuição do máximo do módulo da normal multivariada, e (5.4.16) nos demais casos. Naquelas expressões a correlação ρ_i está representada por λ_i^2 com $i = 1, \dots, r$.

Teorema 3.5.1 *Definindo a estatística do teste bilateral para dados balanceados $|d|$ como a solução da equação:*

$$\int_0^\infty \int_{-\infty}^{+\infty} [\Phi(z + \sqrt{2}|d|s) - \Phi(z - \sqrt{2}|d|s)]^r \phi(z) dz f(s; \nu) ds = 1 - \alpha,$$

tem-se que

$$P \left\{ \mu_i - \mu_{r+1} \in \hat{\mu}_i - \hat{\mu}_{r+1} \pm |d| \hat{\sigma} \sqrt{\frac{2}{n}}, \quad i = 1, \dots, r, \right\} = 1 - \alpha.$$

Demonstração:

Partindo do pressuposto de que a probabilidade nas r inferências simultâneas é preservada e considerando $Z_1, \dots, Z_r$ variáveis aleatórias normais padrão independentes e identicamente distribuídas, pode-se escrever

$$P \left\{ \mu_i - \mu_{r+1} \in \hat{\mu}_i - \hat{\mu}_{r+1} \pm |d| \hat{\sigma} \sqrt{\frac{2}{n}}, \quad i = 1, \dots, r, \right\} = 1 - \alpha$$

$$P \left\{ -|d| \hat{\sigma} \sqrt{\frac{2}{n}} < (\hat{\mu}_i - \mu_i) - (\hat{\mu}_{r+1} - \mu_{r+1}) < |d| \hat{\sigma} \sqrt{\frac{2}{n}}, \quad i = 1, \dots, r, \right\} = 1 - \alpha.$$

Dividindo a expressão por $\hat{\sigma}/\sqrt{n}$ e multiplicando o termo central da desigualdade por σ/σ tem-se:

$$P \left\{ -\sqrt{2}|d| < \frac{[(\hat{\mu}_i - \mu_i) - (\hat{\mu}_{r+1} - \mu_{r+1})] \sqrt{n}/\sigma}{\hat{\sigma}/\sigma} < \sqrt{2}|d|, \quad i = 1, \dots, r, \right\} = 1 - \alpha.$$

Condiccionando ao conhecimento de $\hat{\sigma}^2$, com $s = \hat{\sigma}/\sigma$, sendo $s \geq 0$, tem-se uma probabilidade condicional; essa probabilidade multiplicada pela função densidade de $S, f(s; \nu)$, resulta na probabilidade conjunta de S e dos r contrastes; integrando-se essa probabilidade em relação a s , com $s \geq 0$, chega-se a função de probabilidade marginal

$$\int_0^\infty \left[P \left\{ -\sqrt{2}|d|s < \frac{(\hat{\mu}_i - \mu_i) - (\hat{\mu}_{r+1} - \mu_{r+1})}{\frac{\sigma}{\sqrt{n}}} < \sqrt{2}|d|s, \quad i = 1, \dots, r, \right\} | \hat{\sigma} \right] \times f(s; \nu) ds = 1 - \alpha. \quad (3.5.2)$$

De modo análogo, considerando $Z_i = (\hat{\mu}_i - \mu_i) / (\sigma/\sqrt{n_i})$ e $z = (\hat{\mu}_{r+1} - \mu_{r+1}) / (\sigma/\sqrt{n_{r+1}})$ um escore normal padrão e condicionando ao conhecimento de $\hat{\mu}_{r+1}$, chega-se a probabilidade condicional

$$\left[P \left\{ z - \sqrt{2}|d|s < Z_i < z + \sqrt{2}|d|s, \quad i = 1, \dots, r \right\} | \hat{\mu}_{r+1} \right].$$

Multiplicando essa probabilidade pela função densidade de Z e integrando em relação a z , pode-se escrever a probabilidade interna da expressão (3.5.2) como a probabilidade marginal

$$\int_{-\infty}^{\infty} \left[P \left\{ z - \sqrt{2}|d|s < Z_i < z + \sqrt{2}|d|s, \quad i = 1, \dots, r, \right\} | \hat{\mu}_{r+1} \right] \phi(z) dz. \quad (3.5.3)$$

mas, pelo teorema fundamental do cálculo, tem-se que

$$P \left\{ z - \sqrt{2}|d|s < Z_i < z + \sqrt{2}|d|s \right\} = \Phi(z + \sqrt{2}|d|s) - \Phi(z - \sqrt{2}|d|s).$$

Como os $Z'_i s, i = 1, \dots, r$, são independentemente distribuídos, tem-se

$$P \left\{ z - \sqrt{2}|d|s < Z_i < z + \sqrt{2}|d|s, \quad i = 1, \dots, r, \right\} = [\Phi(z + \sqrt{2}|d|s) - \Phi(z - \sqrt{2}|d|s)]^r,$$

fazendo com que (3.5.3) resulte em:

$$\int_{-\infty}^{\infty} [\Phi(z + \sqrt{2}|d|s) - \Phi(z - \sqrt{2}|d|s)]^r \phi(z) dz. \quad (3.5.4)$$

Substituindo (3.5.4) em (3.5.2) chega-se a

$$\int_0^{\infty} \left[\int_{-\infty}^{\infty} [\Phi(z + \sqrt{2}|d|s) - \Phi(z - \sqrt{2}|d|s)]^r \phi(z) dz \right] f(s; \nu) ds = 1 - \alpha,$$

em que $|d|$ é a solução da equação quando o valor de α for fixado, como se queria provar.

No caso de experimentos balanceados em que $n_i = n$ é o mesmo para todo i mas $n = n_{r+1}$, a correlação em cada contraste é a mesma $\rho = n/(n + n_{r+1})$. Portanto, tem-se um caso especial do teste bilateral para dados balanceados com a correlação constante ρ qualquer. Nesta situação, mais geral, tem-se que

$$P \left\{ \mu_i - \mu_{r+1} \in \hat{\mu}_i - \hat{\mu}_{r+1} \pm |d| \hat{\sigma} \sqrt{\frac{1}{n} + \frac{1}{n_{r+1}}}, \quad i = 1, \dots, r \right\} = \int_0^{\infty} \int_{-\infty}^{+\infty} \left[\Phi \left(\frac{\sqrt{\rho}z + |d|s}{\sqrt{1-\rho}} \right) - \Phi \left(\frac{\sqrt{\rho}z - |d|s}{\sqrt{1-\rho}} \right) \right]^r \phi(z) dz f(s; \nu) ds = 1 - \alpha.$$

$$\text{com } \rho = \frac{n}{n+n_{r+1}}.$$

3.6 Teste de Dunnett bilateral para dados não balanceados

Neste caso são consideradas amostras de quaisquer tamanhos n_i para todos os $i = 1, \dots, r$ tratamentos. Assim, o teste bilateral de Dunnett dá o seguinte intervalo simultâneo de confiança para a diferença entre a média de cada tratamento em teste μ_i com a média do tratamento controle μ_{r+1} :

$$\mu_i - \mu_{r+1} \in \hat{\mu}_i - \hat{\mu}_{r+1} \pm |d| \hat{\sigma} \sqrt{\frac{1}{n_j} + \frac{1}{n_{r+1}}}, \quad i = 1, \dots, r,$$

em que $|d|$ é a solução da equação:

$$\int_0^\infty \int_{-\infty}^{+\infty} \prod_{i=1}^r \left[\Phi \left(\frac{\sqrt{\rho_i} z + |d| s}{\sqrt{1 - \rho_i}} \right) - \Phi \left(\frac{\sqrt{\rho_i} z - |d| s}{\sqrt{1 - \rho_i}} \right) \right] d\Phi(z) f(s; \nu) ds = 1 - \alpha, \quad (3.6.1)$$

com

$$\rho_i = \left(1 + \frac{n_{r+1}}{n_i} \right)^{-1} = \frac{n_i}{n_i + n_{r+1}}, \quad i = 1, \dots, r.$$

Observa-se que o quantil $|d|$ depende de α , r , ν e das correlações ρ_i , que dependem das razões entre os tamanhos amostrais $n_{r+1}/n_1, \dots, n_{r+1}/n_r$. Portanto, também não é possível tabular $|d|$ de uma forma geral. Somente pode-se fazer uso de um programa de computador para calcular $|d|$ dependendo do vetor referente aos tamanhos amostrais $\mathbf{n} = (n_1, \dots, n_{r+1})$ dos dados analisados que afetarão a correlação $\boldsymbol{\lambda} = (\rho_1, \dots, \rho_r)$.

As funções de distribuição da estatística $|d|$ são apresentadas no quinto capítulo, especificamente nas expressões (5.4.9) para $\nu = \infty$, caso em que a

distribuição do máximo do módulo da t multivariada recai na distribuição do máximo do módulo da normal multivariada, e (5.4.16) nos demais casos. Naquelas expressões a correlação ρ_i está representada por λ_i^2 com $i = 1, \dots, r$.

Teorema 3.6.1 *Definindo a estatística do teste bilateral para dados não balanceados $|d|$ como a solução da equação:*

$$\int_0^\infty \int_{-\infty}^{+\infty} \prod_{i=1}^r \left[\Phi \left(\frac{\sqrt{\rho_i} z + |d|s}{\sqrt{1 - \rho_i}} \right) - \Phi \left(\frac{\sqrt{\rho_i} z - |d|s}{\sqrt{1 - \rho_i}} \right) \right] d\Phi(z) f(s; \nu) ds = 1 - \alpha,$$

com $\rho_i = \frac{n_i}{n_i + n_{r+1}}$ tem-se que

$$P \left\{ \mu_i - \mu_{r+1} \in \hat{\mu}_i - \hat{\mu}_{r+1} \pm |d| \hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}, \quad i = 1, \dots, r \right\} = 1 - \alpha$$

Demonstração:

Partindo do pressuposto de que a probabilidade nas r inferências simultâneas é preservada e considerando $Z_1, \dots, Z_r$ variáveis aleatórias normais padrão independentes e identicamente distribuídas, então

$$P \left\{ \mu_i - \mu_{r+1} \in \hat{\mu}_i - \hat{\mu}_{r+1} \pm |d| \hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}, \quad i = 1, \dots, r \right\} = 1 - \alpha$$

ou seja,

$$P \left\{ -|d| \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}} < \frac{(\hat{\mu}_i - \mu_i) - (\hat{\mu}_{r+1} - \mu_{r+1})}{\hat{\sigma}} < |d| \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}, \right. \\ \left. \text{para } i = 1, \dots, r, \right\} = 1 - \alpha.$$

Multiplicando a desigualdade por $\left(\frac{\sqrt{n_i}\sigma}{\sigma}\right)$ tem-se

$$P \left\{ -|d| \sqrt{1 + \frac{n_i}{n_{r+1}}} < \frac{\frac{(\hat{\mu}_i - \mu_i)}{\frac{\sigma}{\sqrt{n_i}}}}{\frac{\hat{\sigma}}{\sigma}} - \frac{\frac{(\hat{\mu}_{r+1} - \mu_{r+1})}{\frac{\sigma}{\sqrt{n_i}}}}{\frac{\hat{\sigma}}{\sigma}} < |d| \sqrt{1 + \frac{n_i}{n_{r+1}}} \right\} = 1 - \alpha.$$

Multiplicando o termo $\left(\frac{\frac{(\hat{\mu}_{r+1} - \mu_{r+1})}{\frac{\sigma}{\sqrt{n_i}}}}{\frac{\hat{\sigma}}{\sigma}}\right)$ por $\sqrt{\frac{n_{r+1}}{n_{r+1}}}$ tem-se

$$P \left\{ -|d| \sqrt{1 + \frac{n_i}{n_{r+1}}} < \frac{\frac{(\hat{\mu}_i - \mu_i)}{\frac{\sigma}{\sqrt{n_i}}}}{\frac{\hat{\sigma}}{\sigma}} - \frac{\sqrt{\frac{n_i}{n_{r+1}}} \frac{(\hat{\mu}_{r+1} - \mu_{r+1})}{\frac{\sigma}{\sqrt{n_{r+1}}}}}{\frac{\hat{\sigma}}{\sigma}} < |d| \sqrt{1 + \frac{n_i}{n_{r+1}}} \right\} = 1 - \alpha.$$

Usando-se as expressões (3.4.2) e (3.4.3) e condicionando ao conhecimento de $\hat{\sigma}^2$, com $s = \hat{\sigma}/\sigma$, sendo $s \geq 0$, tem-se uma probabilidade condicional; essa probabilidade multiplicada pela função densidade de $S, f(s; \nu)$, resulta na probabilidade conjunta de S e dos r contrastes; integrando-se essa probabilidade em relação a s , com $s \geq 0$, chega-se a função de probabilidade marginal

$$\int_0^\infty \left[P \left\{ \frac{-|d|s}{\sqrt{1-\rho_i}} < \frac{(\hat{\mu}_i - \mu_i)}{\frac{\sigma}{\sqrt{n_i}}} - \frac{\sqrt{\rho_i}}{\sqrt{1-\rho_i}} \frac{(\hat{\mu}_{r+1} - \mu_{r+1})}{\frac{\sigma}{\sqrt{n_{r+1}}}} < \frac{|d|s}{\sqrt{1-\rho_i}}, \right. \right. \\ \left. \left. \text{para } i = 1, \dots, r \right\} | \hat{\sigma} \right] f(s; \nu) ds = 1 - \alpha. \quad (3.6.2)$$

De modo análogo, considerando $Z_i = (\hat{\mu}_i - \mu_i)/(\sigma/\sqrt{n_i})$ e $z = (\hat{\mu}_{r+1} - \mu_{r+1})/(\sigma/\sqrt{n_{r+1}})$ um escore normal padrão e condicionando ao conhecimento de $\hat{\mu}_{r+1}$, chega-se a probabilidade condicional

$$\left[P \left\{ \frac{\sqrt{\rho_i}z - |d|s}{\sqrt{1-\rho_i}} < Z_i < \frac{\sqrt{\rho_i}z + |d|s}{\sqrt{1-\rho_i}}, \quad i = 1, \dots, r \right\} | \hat{\mu}_{r+1} \right].$$

Multiplicando essa probabilidade pela função densidade de Z e integrando em relação a z , pode-se escrever a probabilidade interna da expressão (3.6.2) como a probabilidade marginal

$$\int_{-\infty}^{\infty} \left\{ \left[P \left(\frac{\sqrt{\rho_i}z - |d|s}{\sqrt{1-\rho_i}} < Z_i < \frac{\sqrt{\rho_i}z + |d|s}{\sqrt{1-\rho_i}} \right), \quad i = 1, \dots, r, \right] | \hat{\mu}_{r+1} \right\} \phi(z) dz. \quad (3.6.3)$$

Pelo teorema fundamental do cálculo tem-se

$$P \left(\frac{\sqrt{\rho_i}z - |d|s}{\sqrt{1-\rho_i}} < Z_i < \frac{\sqrt{\rho_i}z + |d|s}{\sqrt{1-\rho_i}} \right) = \Phi \left(\frac{\sqrt{\rho_i}z + |d|s}{\sqrt{1-\rho_i}} \right) - \Phi \left(\frac{\sqrt{\rho_i}z - |d|s}{\sqrt{1-\rho_i}} \right),$$

para cada $i = 1, \dots, r$. Pela independência dos Z_i 's, a parte interna da expressão (3.6.3) resulta em

$$\begin{aligned}
P \left[\left(\frac{\sqrt{\rho_i}z + |d|s}{\sqrt{1-\rho_i}} < Z_i < \frac{\sqrt{\rho_i}z + |d|s}{\sqrt{1-\rho_i}} \right), \quad i = 1, \dots, r \right] = \\
= \prod_{i=1}^r \left[\Phi \left(\frac{\sqrt{\rho_i}z + |d|s}{\sqrt{1-\rho_i}} \right) - \Phi \left(\frac{\sqrt{\rho_i}z - |d|s}{\sqrt{1-\rho_i}} \right) \right].
\end{aligned}$$

Assim, a expressão (3.6.3) corresponde a:

$$\int_{-\infty}^{+\infty} \prod_{i=1}^r \left[\Phi \left(\frac{\sqrt{\rho_i}z + |d|s}{\sqrt{1-\rho_i}} \right) - \Phi \left(\frac{\sqrt{\rho_i}z - |d|s}{\sqrt{1-\rho_i}} \right) \right] \phi(z) dz. \quad (3.6.4)$$

Substituindo a expressão (3.6.4) em (3.6.2) chega-se a

$$\int_0^\infty \left\{ \int_{-\infty}^\infty \prod_{i=1}^r \left[\Phi \left(\frac{\sqrt{\rho_i}z + |d|s}{\sqrt{1-\rho_i}} \right) - \Phi \left(\frac{\sqrt{\rho_i}z - |d|s}{\sqrt{1-\rho_i}} \right) \right] \phi(z) dz \right\} f(s; \nu) ds = 1 - \alpha,$$

em que $|d|$ é a solução da equação quando o valor de α é fixado, como se queria provar.

3.7 Aspectos computacionais relacionados às distribuições das estatísticas do teste de Dunnett disponíveis na literatura

As estatísticas utilizadas nos testes de Dunnett têm distribuição t multivariada e, em casos particulares, normal multivariada. A seguir são apresentados os principais estudos encontrados na literatura abordando

algoritmos específicos para o teste de Dunnett ou, de forma geral, para as distribuições normal e t multivariadas.

Dunlap, Marx e Agamy (1981) descreveram uma função em FORTRAN IV para obtenção da função de distribuição da estatística do teste bilateral de Dunnett para dados balanceados. Neste trabalho, as probabilidades acumuladas da normal padrão são resolvidas por aproximações fornecidas por Dunlap e Duffy (1975 apud DUNLAP; MAX; AGAMY, 1981) ou, para valores extremos, pela aproximação de Zelen e Severo (1965 apud DUNLAP; MARX; AGAMY, 1981). As demais integrais envolvidas no cálculo da distribuição de probabilidade de Dunnett foram resolvidas utilizando o método numérico de Newton-Côtes denominado de regra de Simpson com 20 a 30 pontos. Os autores comentam que a precisão de seus valores quando comparados aos valores das tabelas do teste de Dunnett publicadas em Hanhn e Hendrickson (1971 apud DUNLAP; MARX; AGAMY, 1981) era da ordem da quarta casa decimal. Porém, este algoritmo limita-se ao fornecimento da probabilidade. Não possibilita, por exemplo, a obtenção dos quantis da distribuição dada a probabilidade acumulada, invertendo a função de distribuição.

Também não aborda a obtenção da probabilidade no caso da estatística do teste envolver uma distribuição t multivariada não-central.

Schervish (1984) apresenta o algoritmo AS 195 para calcular probabilidades da normal multivariada dentro de uma faixa de erro estabelecida, utilizando para resolver as integrais envolvidas também os métodos numéricos de Newton - Côtes com sequências de subintervalos de 2 a 5 pontos. O autor baseou-se em estudos e algoritmos anteriores dados por Bohrer e Schervish (1981), Bohrer, Schervish e Sheft (1982) e Milton (1972 apud SCHERVISH, 1984). A vantagem desse algoritmo sobre os anteriores foi a melhora do tempo de execução com o aumento da precisão requerida para os valores. Porém, a restrição para o uso deste algoritmo é que ele calcula para no máximo 7

dimensões, em função do tempo de processamento. A matriz de correlação deve ser não singular e passível de ser escrita como uma matriz triangular inferior.

Dunnett publica em 1989 o algoritmo AS 251 que possibilita calcular integrais de probabilidade normal multivariada com estrutura de correlação produto entre as variáveis, com dimensão máxima de 50. Este algoritmo baseia-se no de Schervish (portanto, utilizando quadraturas de Newton-Côtes), exceto pela particularidade de assumir uma estrutura de correlação especial, o que reduz consideravelmente o tempo de cálculo e viabiliza a aplicação do método com um número maior de variáveis. A restrição da estrutura de correlação produto é aplicada em muitos problemas de comparações múltiplas (DUNNETT, 1955, 1989; HSU, 1999).

Para verificar as funções implementadas computacionalmente, específicas ou relacionadas ao teste de Dunnett, optou-se pelo *software R* (R CORE TEAM DEVELOPMENT, 2012) por ser um programa de livre distribuição e código fonte aberto. Foram encontradas funções que fornecem os quantis ou probabilidades das distribuições normal e t multivariadas. Uma possibilidade é a utilização da biblioteca *mvtnormp*, que utiliza o algoritmo AS 251 de Dunnett (1989). Porém, esta biblioteca deixou de ser disponibilizada no CRAN por falta de manutenção e só pode ser usado em versões mais antigas do R. O algoritmo resolve as integrais pelo método numérico de Simpson e permite no máximo 50 comparações entre tratamentos.

Wang e Kennedy (1992) propuseram uma expansão por série de Taylor da integral da normal multivariada para regiões retangulares. Eles compararam seus valores com os obtidos pelo algoritmo de Schervish (1984, 1985) obtendo maior precisão, porém, num maior tempo computacional. Concluem que o método é eficiente para integrais de até cinco dimensões. Acima disso o tempo computacional excessivo torna inviável o método.

Drezner (1992) apresenta uma generalização do seu método para calcular probabilidades bivariadas para o cálculo de probabilidades multivariadas. Utilizou o método de quadratura gaussiana proposto por Steen e Gelbard (1969) para resolver a integral m -dimensional. Steen e Gelbard (1969) estudaram as quadraturas gaussianas em integrais do tipo $\int_0^\infty e^{-x^2} f(x) dx$ e $\int_0^b e^{-x^2} f(x) dx$, com $b < \infty$ e propuseram uma quadratura especial. No artigo eles apresentam tabelas contendo os pontos e os seus respectivos pesos para resolver estas quadraturas no caso particular de $b = 1$. O método de Drezner difere do proposto por Schervish (1984) em relação à quadratura e região de integração utilizadas: o método de Schervish usa regras de Simpson's enquanto que o de Drezner usa quadratura gaussiana; o primeiro permite limites de integração finitos enquanto que o segundo requer intervalos infinitos.

Genz e Bretz (1999, 2002) apresentaram e vêm aperfeiçoando métodos para calcular probabilidades da distribuição t multivariada a partir de substituições que transformam uma integral inicialmente q -variada em um hipercubo $q - 1$ dimensional. Nesta situação são aplicados métodos padrões de integração numérica. Eles discutem e comparam o poder de testes de comparações múltiplas para dados com distribuição normal resolvidos utilizando a transformação proposta e resolvendo as integrais pelos métodos de Monte Carlo, de aceitação e rejeição e um algoritmo de regra lattice. Torsten Hothorn, Frank Bretz e Peter Westfal desenvolveram e aperfeiçoam o pacote “mvtnorm” do *software R* para o cálculo de probabilidades das distribuições t e normal multivariadas, cuja última atualização data de 2012.

Na biblioteca mvtnorm existem duas funções básicas diretamente aplicáveis ao teste de Dunnett: *pmvt* e *qmvt*. A *pmvt* é utilizada para a obtenção da função de distribuição da t multivariada e a *qmvt* para obtenção dos quantis. No cálculo das probabilidades acumuladas as rotinas utilizam métodos de aleatorização quase-Monte Carlo. A vantagem do procedimento é que permite

matrizes de correlação com estrutura arbitrária, o que não tem nenhuma implicação para o teste de Dunnett. Esta rotina também possibilita a consideração de parâmetros de não-centralidade nas distribuições. A função *qmv* calcula quantis equicoordenados invertendo a função *pmv* pelo método *uniroot* do R. Segundo os autores do pacote, o uso da função *uniroot* pode resultar em quantis com limitada acurácia.

Os algoritmos podem ainda ser utilizados para no máximo 1000 comparações, o que em circunstâncias práticas nas aplicações do teste de Dunnett dificilmente será atingido ou ultrapassado. A limitação deste método é que por empregar elementos de simulação, cada repetição da mesma implementação fornece resposta diferente a partir da quarta casa decimal, já que tem sua precisão estabelecido como 10^{-4} .

4 TÉCNICAS MATEMÁTICAS

Nesta seção são apresentadas as principais técnicas matemáticas utilizadas no desenvolvimento e implementação dos algoritmos para a avaliação numérica das integrais envolvidas nas funções densidade e distribuição dos testes de Dunnett no caso não-central.

As quadraturas Gaussianas mais comumente utilizadas para resolver integrais são as Gauss-Legendre, Gauss-Hermite e Gauss-Laguerre. Neste trabalho optou-se por transformações de variáveis e mudança das regiões de integração para o intervalo de $[-1,1]$ e aplicou-se a quadratura de Gauss-Legendre em praticamente todas as integrais. Em muitos casos as funções envolvem múltiplas integrais. Nesse caso, os métodos descritos devem ser aplicados em múltiplos estágios, tantos quantos forem o número de dimensões das integrais.

Para obter os quantis foi necessário determinar a inversa das funções de distribuição estudadas e, neste caso, métodos numéricos de determinação das raízes das equações são utilizados, principalmente o método das secante e o de Newton-Raphson.

Também é apresentado uma síntese do método do jacobiano para a mudança de variáveis em integrais múltiplas, que é utilizado para obter a distribuição de transformações de variáveis aleatórias contínuas, as funções marginais e de distribuição das distribuições de probabilidades apresentadas.

4.1 Integração numérica

Pelo teorema fundamental do cálculo, se a função $g(x)$ é contínua em $[a,b]$ e sua primitiva, $G(x)$, é conhecida ou facilmente determinável, então a integral definida desta função neste intervalo é dada por:

$$I = \int_a^b g(x)dx = G(b) - G(a),$$

em que $g(x) = dG(x)/dx = G'(x)$.

Se a primitiva é desconhecida ou de difícil obtenção, ou quando não é possível determinar a forma analítica da $g(x)$, a integral I pode ser calculada por métodos numéricos¹⁴. Integrar numericamente uma função $g(x)$ num dado intervalo $[a,b]$ consiste em integrar um polinômio $P_n(x)$ que aproxime satisfatoriamente a função integrando $g(x)$ e que seja de fácil manuseio, tornando a integração possível (SPERANDIO; MENDES; SILVA, 2003). Polinômios interpoladores, como os polinômios de Legendre e os polinômios de Hermite, são funções que possuem essas características.

A solução numérica de uma integral simples é comumente chamada de quadratura (BARROSO et al., 1987). Na integração numérica, a fórmula de quadratura aproxima a integral definida de uma função usando combinação linear dos valores da função, ou seja, uma soma ponderada dos valores da função em pontos específicos no domínio de integração (FRANCO, 2007). A localização dos pontos utilizados na integração, o grau e o tipo do polinômio interpolador são parâmetros que caracterizam o tipo de quadratura que está sendo aplicado.

Os métodos de integração numérica são classificados em dois grupos:

¹⁴ Às vezes só são conhecidos valores da função $g(x)$ em alguns pontos x_i s pré-definidos, inviabilizando o uso das técnicas de integração conhecidas no cálculo diferencial e integral; em outros casos a integral resulta em uma função que não pode ser expressa em termos de combinações finitas de outras funções algébricas, logarítmicas ou exponenciais (SPERANDIO; MENDES; SILVA, 2003).

- a) fórmulas de Newton-Côtes que empregam valores de $g(x)$ em pontos x_i 's cuja característica é estarem igualmente espaçados;
- b) fórmulas de quadratura gaussiana que baseiam-se em propriedades de polinômios ortogonais para determinar os pontos x_i 's e seus respectivos coeficientes w_i 's.

A diferença entre as fórmulas de Newton-Côtes e das quadraturas gaussianas é que as primeiras usam valores da função em pontos uniformemente espaçados pré-determinados para definir o polinômio interpolador $P_n(x)$ que será integrado, enquanto que as segundas escolhem os pontos x_i 's para calcular a aproximação de uma maneira ótima (BURDEN; FAIRES, 2003).

Quando são tomados livremente os pontos dentro do intervalo $[a, b]$, para calcular os valores da função $g(x)$ e assim determinar $P_n(x)$, pode-se definir um $P_n(x)$ melhor, que equilibre os erros positivos e negativos, fornecendo uma estimativa da integral mais precisa (CHAPRA; CANALE, 2010). Quadratura gaussiana é o nome dado para a classe de técnicas que implementam tal estratégia. Esta seção está centrada no estudo da integração numérica via quadraturas gaussianas.

4.1.1 Quadratura gaussiana

A quadratura gaussiana consiste em tomar um conjunto de n pontos distintos no intervalo $[a, b]$, $x_1 < x_2 < x_3 < \dots < x_n$, de modo que a aproximação

$$I = \int_a^b g(x)dx \approx \sum_{i=1}^n w_i g(x_i), \quad (4.1.1)$$

cujos coeficientes $w_1, w_2, \dots, w_n$ são determinados sob os pontos $x_1, x_2, \dots, x_n$, resulte no cálculo da integral com um erro mínimo, adaptado de Khuri (2003).

Os coeficientes são arbitrários e todos positivos, e os pontos (também chamados de nós) estão limitados a pertencer ao intervalo $[a,b]$, que é o intervalo de integração. A lógica por trás da quadratura gaussiana é a melhor escolha dos x_i 's a fim de maximizar o grau de exatidão da quadratura¹⁵.

Para medir a precisão desta aproximação, assume-se que a melhor escolha dos valores dos w_i 's e dos x_i 's é a que produz o resultado exato para a maior classe de polinômios, isto é, aquela escolha que dá o maior grau de precisão. Como o número total de quantidades desconhecidas são $2n$ (n pontos e n coeficientes), então, $2n$ condições devem ser especificadas. Se os coeficientes de um polinômio são considerados parâmetros, a classe de polinômios de grau no máximo $2n - 1$ é a que contém $2n$ parâmetros. Essa é, então, a maior classe de polinômios para os quais é razoável esperar que a aproximação dê um resultado exato.

Gil, Segura e Temme (2007) expõem o processo de determinar os x_i 's e seus correspondentes w_i 's para qualquer n baseado na teoria dos polinômios ortogonais, que leva a formulação do cálculo dos coeficientes (w_i 's) e dos pontos (x_i 's) como um problema de determinação de autovalores e autovetores de uma matriz, utilizando o algoritmo de Golub – Welsch.

Algumas extensões da aproximação (4.1.1) são obtidas tomando-se $f(x)$ como uma nova função dada por $f(x) = g(x)/\lambda(x)$. Assim:

¹⁵ Uma definição para grau de exatidão, ou grau de precisão, de uma quadratura pode ser obtida em Gil, Segura e Temme (2007). De forma sucinta, diz-se que uma quadratura tem grau de precisão m se ela fornece resultados exatos quando a função é um polinômio de grau menor ou igual a m , mas não é exata para todos os polinômios de grau maior do que m .

$$\int_a^b \lambda(x) f(x) dx \approx \sum_{i=1}^n w_i f(x_i), \quad (4.1.2)$$

em que $\lambda(x)$ é uma particular função positiva denominada de função peso.

A escolha dos x_i 's depende da forma de $\lambda(x)$. Os valores dos x_i 's são as raízes do polinômio de grau n , pertencente a sequência de polinômios ortogonais em relação a $\lambda(x)$ no intervalo $[a, b]$.

Segundo Khuri (2003), uma função peso, $\lambda(x)$, em $[a, b]$ tem a característica de ser não-negativa em qualquer intervalo aberto incluído em $[a, b]$, podendo $a = -\infty$ e/ou $b = \infty$, e

$$\int_a^b |x|^n \lambda(x) dx < \infty, \quad n = 0, 1, 2, \dots$$

Na Tabela 2 estão sintetizadas as principais famílias de polinômios ortogonais, suas respectivas função peso, intervalo e simbologia usualmente utilizada para representá-la. Cada função peso e seu polinômio ortogonal dá origem a um tipo de quadratura gaussiana, cujo nome está relacionado com o polinômio interpolador.

Burden e Faires (2003) comentam que a finalidade de uma função peso é atribuir graus variados de importância a aproximações em certas partes do intervalo¹⁶. Cada função peso tem um conjunto de polinômios que são ortogonais à ela no intervalo $[a, b]$.

¹⁶ O intervalo da função peso é o mesmo que o intervalo de integração numérica.

Tabela 2 Características de algumas famílias de polinômios ortogonais clássicos

Polinômio Ortogonal	Símbolo	Função Peso $\lambda(x)$	Intervalo $[a,b]$
Legendre	$P_n(x)$	1	$[-1,1]$
Jacobi	$P_n^{(\alpha,\beta)}(x)$	$(1-x)^\alpha(1+x)^\beta$, $\alpha, \beta > -1$	$[-1,1]$
Chebyshev do 1º tipo	$T_n(x)$	$\frac{1}{\sqrt{1-x^2}}$	$[-1,1]$
Chebyshev do 2º tipo	$U_n(x)$	$\sqrt{1-x^2}$	$[-1,1]$
Hermite físicos	$H_{nf}(x)$	e^{-x^2}	$(-\infty, \infty)$
Hermite probabilísticos	$H_{np}(x)$	$e^{-\frac{x^2}{2}}$	$(-\infty, \infty)$
Laguerre	$L_n^{(\alpha)}(x)$	$e^{-x}x^\alpha, \alpha > -1$	$[0, \infty)$

A Figura 2 mostra as funções peso $\lambda(x)$ ortogonais em relação a família de polinômios de Legendre, Chebyshev do 1º tipo e Chebyshev do 2º tipo no intervalo $[-1,1]$. A função peso $\lambda(x) = 1$ atribui a mesma ênfase para qualquer valor de x no intervalo $[-1,1]$. Já a função peso $\lambda(x) = 1/\sqrt{1-x^2}$ dá menos ênfase perto do centro do intervalo $(-1,1)$ e mais ênfase quando $|x|$ estiver próximo de 1. Ao contrário, a função peso $\lambda(x) = \sqrt{1-x^2}$ dá mais ênfase perto do centro do intervalo $(-1,1)$ e menos ênfase quando $|x|$ estiver próximo de 1.

A Figura 3 mostra as funções peso $\lambda(x)$ ortogonais em relação a família de polinômios de Laguerre no intervalo $[0, \infty)$ e Hermite probabilísticos e físicos no intervalo $(-\infty, \infty)$. Estas funções peso atribuem maior ênfase para valores de x próximos ao zero. Os polinômios de Hermite probabilísticos e físicos não são equivalentes, um é uma re-escala do outro, dados pela relação $H_{nf}(x) = 2^{n/2}H_{np}(\sqrt{2}x)$.

Pela integração gaussiana são escolhidos os w_i 's e os x_i 's tais que a aproximação seja exata quando a função a ser integrada $f(x)$ é um polinômio de grau menor ou igual a $2n - 1$. Nos casos de polinômios de grau $2n$ ou mais ou nos casos de $f(x)$ ser uma função não polinomial, tem-se um erro de aproximação¹⁷

$$E_n = \int_a^b \lambda(x) f(x) dx - \sum_{i=1}^n w_i f(x_i).$$

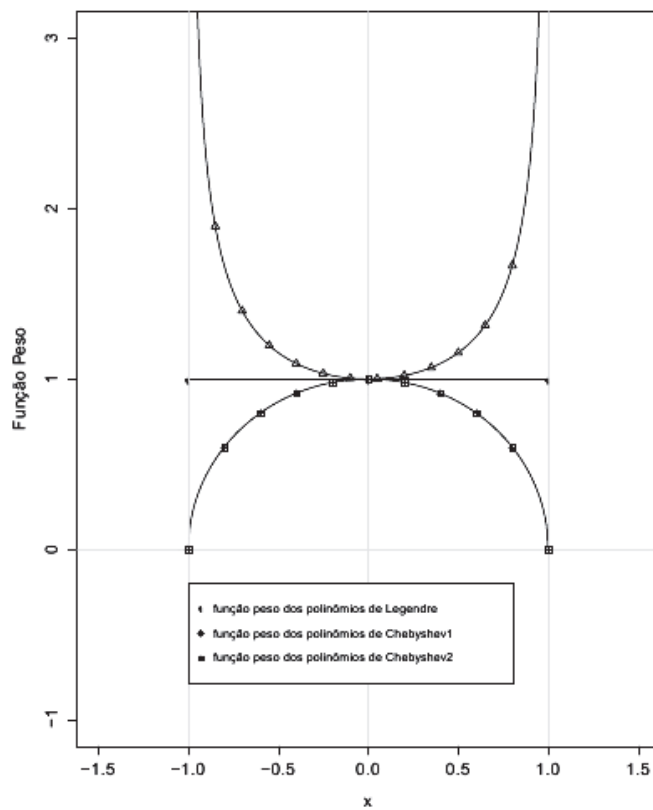


Figura 2 Representação das funções peso $\lambda(x) = 1$, $\lambda(x) = 1 / \sqrt{1 - x^2}$ e $\lambda(x) = \sqrt{1 - x^2}$, ortogonais em relação a família de polinômios de Legendre, Chebyshev do 1º tipo e Chebyshev do 2º tipo, respectivamente, no intervalo $[-1, 1]$

¹⁷ Pode ser chamado também de erro de truncamento, termo do resto ou fator de correção.

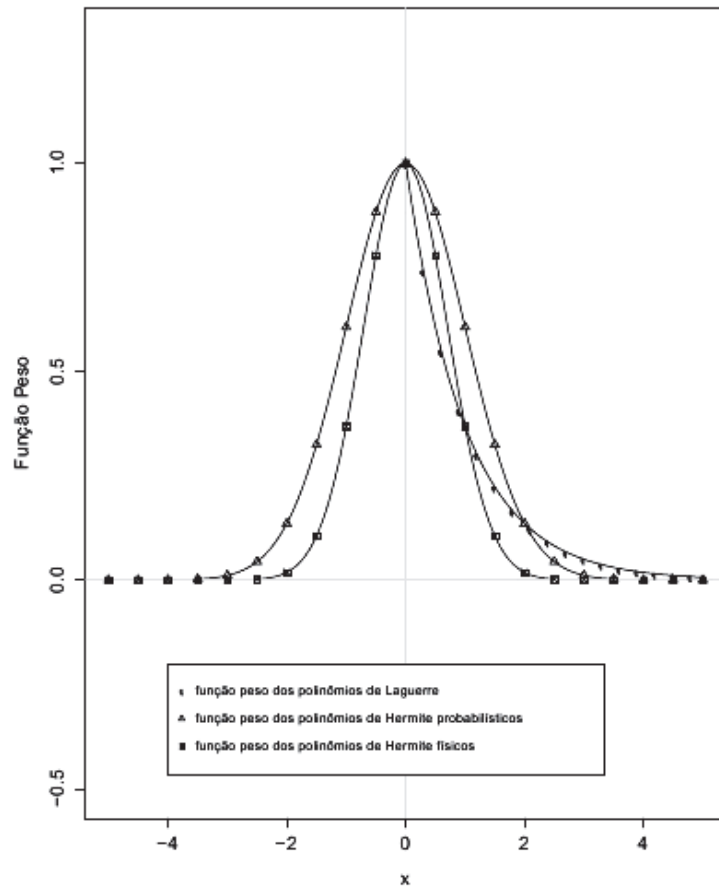


Figura 3 Representação das funções peso $\lambda(x) = e^{-x}$, $\lambda(x) = e^{-\frac{x^2}{2}}$ e $\lambda(x) = e^{-x^2}$, ortogonais em relação a família de polinômios de Laguerre no intervalo $[0, \infty)$, Hermite probabilísticos e Hermite físicos no intervalo $(-\infty, \infty)$

Na Tabela 3 estão apresentadas as fórmulas para obter o erro de aproximação das quadraturas mais usuais. Todas as expressões do erro de aproximação contém a derivada de ordem $2n$ da função $f(x)$, $f^{(2n)}(x)$, aplicada no ponto ξ . O inteiro n é o índice do último termo considerado no cálculo da integral. O valor de ξ a ser usado é um dos pontos x_i que maximiza o módulo da

função $f^{(2n)}(x)$, ou seja, é o valor de x_i tal que $\max_i |f^{(2n)}(x_i)|$. No entanto, obter a derivada de ordem $2n$ da função $f(x)$ muitas vezes é complexo, o que torna praticamente inviável determinar a precisão da aproximação (FRANCO, 2007). Na prática, fixa-se a precisão desejada ϵ no cálculo do resultado da integral e procede-se aos cálculos com valores de n crescentes, comparando-se cada resultado com o seu anterior. Quando o erro relativo entre dois resultados consecutivos for menor do que uma precisão ϵ pré-fixada, pára-se o processo.

Tabela 3 Erro de aproximação das quadraturas cujas funções $f(x)$ são de grau maior do que $2n - 1$ ou são funções não polinomiais e foram aproximadas por polinômios ortogonais de grau n , dadas por Ralston e Rabinowitz (1965), sendo $f^{(2n)}(x)$ contínua em $[a, b]$

Quadratura	$E_n = \int_a^b \lambda(x)f(x)dx - \sum_{i=1}^n w_i f(x_i)$	$a < \xi < b$
G-Legendre	$\frac{2^{2n+1} [(n!)]^4}{(2n+1)[(2n)!]^3} f^{(2n)}(\xi)$	$-1 < \xi < 1$
G-Jacobi	$\frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)\Gamma(n+\alpha+\beta+1)n!2^{(2n+\alpha+\beta+1)}}{(2n+\alpha+\beta+1)[\Gamma(2n+\alpha+\beta+1)]^2(2n)!} f^{(2n)}(\xi)$	$-1 < \xi < 1$
G-Chebyshev do 1° tipo	$\frac{2\pi}{2^{2n}(2n)!} f^{(2n)}(\xi)$	$-1 < \xi < 1$
G-Chebyshev do 2° tipo	$\frac{\pi}{2^{2n+1}(2n)!} f^{(2n)}(\xi)$	$-1 < \xi < 1$
G-Hermite físicos	$\frac{(n!)\sqrt{\pi}}{2^n(2n)!} f^{(2n)}(\xi)$	$-\infty < \xi < \infty$
G-Laguerre	$\frac{(n!)^2}{(2n)!} f^{(2n)}(\xi)$, para $\alpha = 0$	$0 < \xi < \infty$

O valor de ξ a ser usado é o valor x_i tal que $\max_i |f^{(2n)}(x_i)|$.

4.1.2 Algoritmo de *Golub – Welsch*

O algoritmo de *Golub–Welsch* é um método para calcular os nós (x_i 's) e os pesos (w_i 's) para as fórmulas das quadraturas gaussianas por meio da obtenção dos autovalores e autovetores de uma matriz.

Alguns conceitos necessários para entender o algoritmo de *Golub–Welsch* estão resumidos a seguir:

- a) dada uma função peso $\lambda(x)$ em $[a,b]$, chama-se produto interno ou produto escalar de dois polinômios $p(x)$ e $q(x)$ em relação a $\lambda(x)$ a definição

$$\langle p, q \rangle = \int_a^b p(x)q(x)\lambda(x)dx;$$

- b) a norma do polinômio $p(x)$ em relação a $\lambda(x)$ pode ser escrita como

$$\|p\|_\lambda = \sqrt{\langle p, p \rangle} = \left[\int_a^b [p(x)]^2 \lambda(x) dx \right]^{\frac{1}{2}};$$

- c) os polinômios $p(x)$ e $q(x)$ são ditos ortogonais em relação a $\lambda(x)$ se o produto interno entre eles for nulo, ou seja, $\langle p, q \rangle = 0$;
- d) dado um conjunto de polinômios de grau menor ou igual a n , representado por $\{p_n\}$, ele forma uma base ortogonal $\{p_k\}_{k=0}^n$ do espaço vetorial $\{\mathbb{P}_n\}$, se satisfaz a propriedade $\langle p_i, p_j \rangle = 0$ para $i, j = 0, \dots, n$ com $i \neq j$;
- e) quando, além da propriedade anterior, a base satisfaz a propriedade de $\langle p_i, p_i \rangle = 1$ para todo i , ela é chamada de ortonormal e é simbolizada por $\{\tilde{p}_k(x)\}_{k=0}^n$.

Segundo Gil, Segura e Temme (2007), o cálculo dos nós (x_i 's) parte da relação de recorrência dada por:

$$\begin{aligned}\alpha_1 \tilde{p}_1(x) + \beta_0 \tilde{p}_0(x) &= x \tilde{p}_0(x), \\ \alpha_{k+1} \tilde{p}_{k+1}(x) + \beta_k \tilde{p}_k(x) + \alpha_k \tilde{p}_{k-1}(x) &= x \tilde{p}_k(x) \quad k = 1, 2, \dots,\end{aligned}$$

em que $\{\tilde{p}_k(x)\}$ é um conjunto de polinômios ortonormais.

Tomando $x = x_i$ um dos nós da quadratura, isto é um dos zeros de $\tilde{p}_n(x)$, e escrevendo a relação para $k \leq n-1$, tem-se:

$$\left\{ \begin{array}{llll} \alpha_1 \tilde{p}_1(x_i) & + & \beta_0 \tilde{p}_0(x_i) & = & x_i \tilde{p}_0(x_i), \\ \alpha_2 \tilde{p}_2(x_i) & + & \beta_1 \tilde{p}_1(x_i) & + & \alpha_1 \tilde{p}_0(x_i) & = & x_i \tilde{p}_1(x_i), \\ \alpha_3 \tilde{p}_3(x_i) & + & \beta_2 \tilde{p}_2(x_i) & + & \alpha_2 \tilde{p}_1(x_i) & = & x_i \tilde{p}_2(x_i), \\ \vdots & & & & & & \\ \alpha_n \tilde{p}_n(x_i) & + & \beta_{n-1} \tilde{p}_{n-1}(x_i) & + & \alpha_{n-1} \tilde{p}_{n-2}(x_i) & = & x_i \tilde{p}_{n-1}(x_i) \end{array} \right.$$

Como x_i é raiz de $\tilde{p}_n$, na última equação $\tilde{p}_n(x_i) = 0$ e o termo $\alpha_n \tilde{p}_n(x_i)$ desaparece. Assim tem-se:

$$\left\{ \begin{array}{llll} \alpha_1 \tilde{p}_1(x_i) & + & \beta_0 \tilde{p}_0(x_i) & = & x_i \tilde{p}_0(x_i), \\ \alpha_2 \tilde{p}_2(x_i) & + & \beta_1 \tilde{p}_1(x_i) & + & \alpha_1 \tilde{p}_0(x_i) & = & x_i \tilde{p}_1(x_i), \\ \alpha_3 \tilde{p}_3(x_i) & + & \beta_2 \tilde{p}_2(x_i) & + & \alpha_2 \tilde{p}_1(x_i) & = & x_i \tilde{p}_2(x_i), \\ \vdots & & & & & & \\ \beta_{n-1} \tilde{p}_{n-1}(x_i) & + & \alpha_{n-1} \tilde{p}_{n-2}(x_i) & = & x_i \tilde{p}_{n-1}(x_i) \end{array} \right.$$

que pode ser escrita de forma matricial como:

$$\begin{pmatrix} \beta_0 & \alpha_1 & 0 & 0 & \cdots & 0 \\ \alpha_1 & \beta_1 & \alpha_2 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \beta_2 & \alpha_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \ddots & \alpha_{n-1} \\ 0 & 0 & \cdots & \cdots & \alpha_{n-1} & \beta_{n-1} \end{pmatrix} \begin{pmatrix} \tilde{p}_0(x_i) \\ \tilde{p}_1(x_i) \\ \vdots \\ \tilde{p}_{n-2}(x_i) \\ \tilde{p}_{n-1}(x_i) \end{pmatrix} = x_i \begin{pmatrix} \tilde{p}_0(x_i) \\ \tilde{p}_1(x_i) \\ \vdots \\ \tilde{p}_{n-2}(x_i) \\ \tilde{p}_{n-1}(x_i) \end{pmatrix}$$

ou

$$\mathbf{J}\hat{\mathbf{p}}(x_i) = x_i\hat{\mathbf{p}}(x_i)$$

em que $\mathbf{J}$ é a matriz de *Jacobi*, de ordem $n \times n$ com os coeficientes da relação de recorrência e $\hat{\mathbf{p}}(x_i) = [\tilde{p}_0(x_i), \tilde{p}_1(x_i), \dots, \tilde{p}_{n-1}(x_i)]^\top$.

Logo $\tilde{p}_n(x_i) = 0 \Leftrightarrow x_i$ é um autovalor da matriz $\mathbf{J}$ e o vetor $\hat{\mathbf{p}}(x_i)$ correspondente ao x_i é o respectivo autovetor.

Assim, os n zeros de $\tilde{p}_n$, $x_1, x_2, \dots, x_n$, que correspondem aos nós da quadratura gaussiana, podem ser obtidos calculando os autovalores da matriz $\mathbf{J}$.

Quando se parte de polinômios ortogonais não normalizados $\{p_k(x)\}_{k=0}^n$ para o cálculo dos nós (x_i 's), a relação de recorrência dos três termos - TTRR - é da forma (GIL; SEGURA; TEMME, 2007):

$$xp_k(x) = a_k p_{k+1}(x) + b_k p_k(x) + c_k p_{k-1}(x), \quad c_0 p_{-1}(x) = 0 \quad e \quad k \geq 0.$$

Procedendo da mesma forma que anteriormente, chega-se ao resultado de que os zeros de p_n são os autovalores da matriz

$$\mathbf{S} = \begin{pmatrix} b_0 & a_0 & 0 & 0 & \cdots & 0 \\ c_1 & b_1 & a_1 & 0 & \cdots & 0 \\ 0 & c_2 & b_2 & a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & a_{n-2} \\ 0 & 0 & 0 & \vdots & c_{n-1} & b_{n-1} \end{pmatrix}.$$

Para tornar mais simples a determinação destes autovalores, pode-se transformar a matriz $\mathbf{S}$ numa matriz semelhante $\mathbf{J}$ simétrica, aplicando uma transformação de similaridade e construindo a matriz de *Jacobi* para os polinômios ortonormais. O polinômio ortogonal $p_k(x)$ pode ser relacionado com o polinômio ortonormal $\tilde{p}_k(x)$ pela expressão $p_k(x) = \lambda_k \tilde{p}_k(x)$, com $\lambda_k = \|p_k\|$. Fazendo esta substituição na TTRR tem-se:

$$x\lambda_k \tilde{p}_k(x) = a_k \lambda_{k+1} \tilde{p}_{k+1}(x) + b_k \lambda_k \tilde{p}_k(x) + c_k \lambda_{k-1} \tilde{p}_{k-1}(x),$$

resultando em

$$x\tilde{p}_k(x) = a_k \frac{\lambda_{k+1}}{\lambda_k} \tilde{p}_{k+1}(x) + b_k \tilde{p}_k(x) + c_k \frac{\lambda_{k-1}}{\lambda_k} \tilde{p}_{k-1}(x).$$

Chamando $u_k = \frac{\lambda_{k+1}}{\lambda_k}$ tem-se $\tilde{a}_k = a_k \frac{\lambda_{k+1}}{\lambda_k} = a_k u_k$ e $\tilde{c}_k = \frac{c_k}{u_{k-1}}$ que, substituídos na expressão acima, dá:

$$x\tilde{p}_k(x) = \tilde{a}_k \tilde{p}_{k+1}(x) + b_k \tilde{p}_k(x) + \tilde{c}_k \tilde{p}_{k-1}(x), \quad k \geq 0.$$

A matriz de Jacobi para o conjunto de polinômios ortonormais $\{\tilde{p}_k(x)\}$ torna-se simétrica se e $\tilde{a}_k = \tilde{c}_{k+1}$ na matriz S , para $k=0,1,2,\dots, n-2$, ou seja, $a_k \frac{\lambda_{k+1}}{\lambda_k} = c_{k+1} \frac{\lambda_k}{\lambda_{k+1}}$ o que resulta em $\alpha_k = \sqrt{c_k a_{k+1}}$.¹⁸

$$x\tilde{p}_k(x) = \alpha_{k+1}\tilde{p}_{k+1}(x) + \beta_k\tilde{p}_k(x) + \alpha_k\tilde{p}_{k-1}(x),$$

em que $\alpha_0 p_{-1}(x) = 0$; $\beta_k = b_k$ $k \geq 0$; $\alpha_k = \sqrt{c_k a_{k+1}}$, $k \geq 1$.

A correspondente matriz de Jacobi da matriz S é dada por

$$J = \begin{pmatrix} b_0 & \sqrt{a_0 c_1} & 0 & 0 & \cdots & 0 \\ \sqrt{a_0 c_1} & b_1 & \sqrt{a_1 c_2} & 0 & \cdots & 0 \\ 0 & \sqrt{a_1 c_2} & b_2 & \sqrt{a_2 c_3} & \cdots & 0 \\ 0 & 0 & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \sqrt{a_{n-2} c_{n-1}} \\ 0 & 0 & 0 & \cdots & \sqrt{a_{n-2} c_{n-1}} & b_{n-1} \end{pmatrix}$$

¹⁸ Para $k = 0, 1, 2, \dots, n-2$, tem-se

$$\tilde{a}_k = a_k u_k \Rightarrow u_k = \frac{\tilde{a}_k}{a_k}$$

e

$$\tilde{c}_{k+1} = \frac{c_{k+1}}{u_k} \Rightarrow u_k = \frac{c_{k+1}}{\tilde{c}_{k+1}},$$

ou seja,

$$\Rightarrow u_k = u_k \Rightarrow \frac{\tilde{a}_k}{a_k} = \frac{c_{k+1}}{\tilde{c}_{k+1}} \Rightarrow \tilde{a}_k \tilde{c}_{k+1} = a_k c_{k+1}.$$

Como deve-se ter:

$$\tilde{a}_k = \tilde{c}_{k+1} \Rightarrow \tilde{a}_k^2 = a_k c_{k+1} \Rightarrow \tilde{a}_k = \sqrt{a_k c_{k+1}}.$$

Assim, $\tilde{a}_k = \tilde{c}_{k+1} = \sqrt{a_k c_{k+1}}$, para $k = 0, 1, \dots, n-2$, que pode ser chamada de $\alpha_k = \sqrt{c_k a_{k+1}}$, para $k = 1, \dots, n-1$.

Esta matriz tem os mesmos autovalores do que a matriz S (já que elas estão relacionados por uma transformação de equivalência), e agora pode-se aplicar o algoritmo para calcular os pesos w_i 's para as quadraturas usando a matriz J .

De forma similar, os pesos (w_i 's) da quadratura gaussiana podem ser calculados a partir dos autovetores correspondentes a esses autovalores. Sejam $\tilde{p}_j(x)$ e $\tilde{p}_k(x)$ com $j, k = 0, 1, \dots, n-1$, polinômios de grau menor ou igual a $(2n-2)$. A integral

$$\langle \tilde{p}_j(x), \tilde{p}_k(x) \rangle = \int_a^b \tilde{p}_j(x) \tilde{p}_k(x) \lambda(x) dx$$

pode ser calculada exatamente usando uma quadratura gaussiana com n pontos. Então,

$$\delta_{jk} = \langle \tilde{p}_j(x), \tilde{p}_k(x) \rangle = \sum_{i=1}^n w_i \tilde{p}_j(x_i) \tilde{p}_k(x_i),$$

em que $\delta_{jk} = 1$ se $j = k$, pois resulta na norma de $\tilde{p}_j(x)$, e $\delta_{jk} = 0$ se $j \neq k$, pois neste caso os polinômios são ortogonais. Essa expressão pode ser escrita na forma matricial como

$$\mathbf{P}^T \mathbf{W} \mathbf{P} = \mathbf{I}$$

em que a matriz

a)

$$\mathbf{W} = \begin{pmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & w_n \end{pmatrix}$$

Tem na sua diagonal principal os pesos w_i 's a serem determinados, correspondentes a cada um dos autovalores x_i 's da matriz de Jacobi $\mathbf{J}$;

b)

$$\mathbf{P} = \begin{pmatrix} \tilde{p}_0(x_1) & \cdots & \tilde{p}_{n-1}(x_1) \\ \tilde{p}_0(x_2) & \cdots & \tilde{p}_{n-1}(x_2) \\ \vdots & & \vdots \\ \tilde{p}_0(x_n) & \cdots & \tilde{p}_{n-1}(x_n) \end{pmatrix}$$

tem nas suas colunas cada um dos polinômios ortonormais $\tilde{p}_j(x)$ ($j = 0, \dots, n-1$) sendo que em cada linha estão aplicados em um x_i ($i = 1, \dots, n$).

Assim:

$$\begin{pmatrix} \tilde{p}_0(x_1) & \cdots & \tilde{p}_0(x_n) \\ \tilde{p}_1(x_1) & \cdots & \tilde{p}_1(x_n) \\ \vdots & \ddots & \vdots \\ \tilde{p}_{n-1}(x_1) & \cdots & \tilde{p}_{n-1}(x_n) \end{pmatrix} \cdot \begin{pmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & w_n \end{pmatrix} \cdot \begin{pmatrix} \tilde{p}_0(x_1) & \cdots & \tilde{p}_{n-1}(x_1) \\ \tilde{p}_0(x_2) & \cdots & \tilde{p}_{n-1}(x_2) \\ \vdots & \ddots & \vdots \\ \tilde{p}_0(x_n) & \cdots & \tilde{p}_{n-1}(x_n) \end{pmatrix} =$$

$$\begin{pmatrix} \tilde{p}_0(x_1)w_1 & 0 & \cdots & 0 \\ 0 & \tilde{p}_1(x_2)w_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \tilde{p}_{n-1}(x_n)w_n \end{pmatrix} \cdot \begin{pmatrix} \tilde{p}_0(x_1) & \cdots & \tilde{p}_{n-1}(x_1) \\ \tilde{p}_0(x_2) & \cdots & \tilde{p}_{n-1}(x_2) \\ \vdots & \ddots & \vdots \\ \tilde{p}_0(x_n) & \cdots & \tilde{p}_{n-1}(x_n) \end{pmatrix} =$$

$$\begin{pmatrix} \tilde{p}_0(x_1)w_1\tilde{p}_0(x_1) & 0 & \cdots & 0 \\ 0 & \tilde{p}_1(x_2)w_2\tilde{p}_1(x_2) & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \tilde{p}_{n-1}(x_n)w_n\tilde{p}_{n-1}(x_n) \end{pmatrix} =$$

$$\begin{pmatrix} [\tilde{p}_0(x_1)]^2w_1 & 0 & \cdots & 0 \\ 0 & [\tilde{p}_1(x_2)]^2w_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & [\tilde{p}_{n-1}(x_n)]^2w_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Para que a igualdade matricial acima se verifique é necessário que $\mathbf{P}$ seja invertível. Assim, pré-multiplicando esta expressão por $(\mathbf{P}^\top)^{-1}$ e pós-multiplicando por $\mathbf{P}^{-1}$ tem-se que,

$$\begin{aligned} (\mathbf{P}^\top)^{-1} \mathbf{P}^\top \mathbf{W} \mathbf{P} \mathbf{P}^{-1} &= (P^\top)^{-1} I P^{-1}, \\ \Rightarrow \mathbf{W} &= (P^\top)^{-1} P^{-1}. \end{aligned}$$

Por propriedades de matrizes, tem-se que,

$$\begin{aligned} \mathbf{W} &= (P P^\top)^{-1} \\ \Rightarrow \mathbf{W}^{-1} &= \mathbf{P} \mathbf{P}^\top. \end{aligned}$$

Então, chega-se a

$$\frac{1}{w_i} = \sum_{k=0}^{n-1} (\tilde{p}_k(x_i))^2 = \|\hat{\mathbf{p}}(x_i)\|_E^2,$$

ou seja,

$$w_i = \frac{1}{\|\hat{\mathbf{p}}(x_i)\|_E^2}, \quad (4.1.3)$$

em que $\|\cdot\|_E$ é a norma Euclidiana usual

Por outro lado, dado $\hat{\phi}^{(i)} = (\phi_1^{(i)}, \dots, \phi_n^{(i)})^\top$ um autovetor da matriz $\mathbf{J}$ associado ao autovalor x_i , então existe uma constante C tal que $\hat{\phi}^{(i)} = C \hat{\mathbf{p}}(x_i) = C(\tilde{p}_0(x_i), \dots, \tilde{p}_{n-1}(x_i))^T$. Isto acontece porque o autovetor associado a cada autovalor é único, a menos de uma constante multiplicativa. Ou seja

$$\hat{\phi}^{(i)}(x_i) = \begin{pmatrix} \phi_1^{(i)} \\ \phi_2^{(i)} \\ \vdots \\ \phi_n^{(i)} \end{pmatrix} = C \begin{pmatrix} \tilde{p}_0(x_i) \\ \tilde{p}_1(x_i) \\ \vdots \\ \tilde{p}_{n-1}(x_i) \end{pmatrix}$$

O valor de C pode ser obtido considerando

$$\begin{aligned}
 1 &= \langle \tilde{p}_0(x), \tilde{p}_0(x) \rangle \\
 \Rightarrow 1 &= [\tilde{p}_0(x)]^2 \int_a^b \lambda(x) dx \\
 \Rightarrow 1 &= [\tilde{p}_0(x)]^2 \mu_0, \quad \text{com } \mu_0 = \int_a^b \lambda(x) dx \\
 \Rightarrow [\tilde{p}_0(x)]^2 &= \frac{1}{\mu_0} \\
 \Rightarrow \sqrt{\mu_0} &= \frac{1}{\tilde{p}_0(x)}.
 \end{aligned}$$

Então

$$\begin{pmatrix} \phi_1^{(i)} \\ \phi_2^{(i)} \\ \vdots \\ \phi_n^{(i)} \end{pmatrix} = C \begin{pmatrix} \tilde{p}_0(x_i) \\ \tilde{p}_1(x_i) \\ \vdots \\ \tilde{p}_{n-1}(x_i) \end{pmatrix} \Leftrightarrow C = \phi_1^{(i)} \frac{1}{\tilde{p}_0(x_i)} \Rightarrow C = \sqrt{\mu_0} \phi_1^{(i)}$$

obtendo-se $\hat{\phi}^{(i)}(x_i) = \sqrt{\mu_0} \phi_1^{(i)} \tilde{\mathbf{p}}(x_i)$.

Substituindo $\tilde{\mathbf{p}}(x_i) = \hat{\phi}^{(i)}(x_i) / (\sqrt{\mu_0} \phi_1^{(i)})$ na expressão (4.1.3), obtém-se

$$w_i = \frac{1}{\|\tilde{\mathbf{p}}(x_i)\|_E^2} = \frac{1}{\left[\left\| \frac{\hat{\phi}^{(i)}(x_i)}{\sqrt{\mu_0} \phi_1^{(i)}} \right\| \right]^2} = \mu_0 \frac{[\phi_1^{(i)}]^2}{\|\hat{\phi}^{(i)}(x_i)\|^2},$$

sendo:

- a) $\phi_1^{(i)}$ o primeiro elemento do autovetor $\hat{\phi}^{(i)}$;
- b) $||\hat{\phi}^{(i)}(x_i)||^2$ a norma quadrática de $\hat{\phi}^{(i)}(x_i)$, que vale 1 pois o vetor $\hat{\phi}^{(i)}(x_i)$ é ortonormal.

Assim, os pesos w_i 's para as quadraturas gaussianas são obtidos a partir dos respectivos x_i 's e seus autovetores normalizados $\hat{\phi}^{(i)}(x_i)$, pela equação

$$w_i = \mu_0 [\phi_1^{(i)}]^2. \quad (4.1.4)$$

A Tabela 4 resume as fórmulas para montar a matriz de *Jacobi* das quadraturas gaussianas usuais e o valor da constante μ_0 utilizada para o cálculo dos pesos (w_i 's) de cada quadratura na expressão (4.1.4).

4.1.3 Quadraturas para integrais impróprias

Nas quadraturas de Gauss-Laguerre $\int_0^\infty e^{-x} g(x) dx$ e de Gauss-Hermite $\int_{-\infty}^\infty e^{-x^2} g(x) dx$ tem-se uma integral imprópria do 1º tipo, ou seja, uma integral em que $g(x)$ é Riemann integrável em $[a, b]$, com $b > a$, em que o(s) ponto(s) de integração a e/ou b , é(são) infinito(s) (KHURI, 2003).

Uma integral imprópria deste tipo é definida como o limite de uma integral definida quando um dos pontos finais do intervalo de integração tende ao infinito (ou ambos os pontos finais do intervalo), ou seja:

$$I = \int_a^\infty g(x) dx = \lim_{b \rightarrow \infty} \int_a^b g(x) dx.$$

Tabela 4 Fórmulas para montar a matriz de *Jacobi* das quadraturas gaussianas usuais e o valor da constante μ_0 utilizada para o cálculo dos pesos (w_i 's) de cada quadratura

Quadratura	$\beta_k (k = 0, \dots, n-1)$	$\alpha_k (k = 1, \dots, n-2)$	$\mu_0 = \int_a^b \lambda(x) dx$
G-Legendre	0	$\frac{k}{\sqrt{4k^2-1}}$	2
G-Chebyshev do 1° tipo	0	$\alpha_1 = \sqrt{\frac{1}{2}},$ $\frac{1}{2}$ para $k \geq 2$	π
G-Chebyshev do 2° tipo	0	$\frac{1}{2}$	$\frac{\pi}{2}$
G-Hermite físicos	0	$\sqrt{\frac{k}{2}}$	$\sqrt{\pi}$
G-Laguerre	$\beta_{k-1} = 2k - 1 + \alpha$ para $k \geq 1$	$\sqrt{k(k+\alpha)}$	$\Gamma(\alpha+1)$
G-Jacobi	$\beta_0 = \frac{\beta-\alpha}{\alpha+\beta+2}$ $\beta_k = \frac{\beta^2-\alpha^2}{(2k+\alpha+\beta)(2k+\alpha+\beta+2)}$ para $k \geq 1$	$(*) \alpha_k$	$2^{\alpha+\beta+1} \frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+1)}$
$(*) \alpha_k = \frac{2}{(2k+\alpha+\beta)} \sqrt{\frac{k(k+\alpha)(k+\beta)(k+\alpha+\beta)}{(2k+\alpha+\beta+1)(2k+\alpha+\beta-1)}}$			

Se esta integral for convergente¹⁹, ou seja, se este limite for finito, pode-se avaliá-la por métodos numéricos, pois o valor da integral é o seu respectivo limite. É possível utilizar os métodos de quadratura gaussiana, pois eles não utilizam o valor do integrando nos pontos finais do limite de integração e utilizam um número finitos de pontos no cálculo da aproximação da integral²⁰.

Para integrais $\int_a^b g(x) dx$ imprópria do 2° tipo, em que o intervalo $[a, b]$ é finito mas a função $g(x)$ a ser integrada assume um valor infinito em alguns pontos²¹ dentro do intervalo $[a, b]$, por exemplo para $x = c$, segundo Khuri (2003) se $g(x) \rightarrow \infty$ quando $x \rightarrow c$, com $a < c < b$, então:

¹⁹ A convergência da integral pode ser determinada pelo uso de critérios de convergência de modo similar ao estudo de convergência de sequências (KHURI, 2003).

²⁰ Considerando que a função a ser integrada tem derivadas de todas as ordens contínuas (SPERANDIO; MENDES; SILVA, 2003).

²¹ Um número finito de pontos.

$$\int_a^b g(x)dx = \int_a^c g(x)dx + \int_c^b g(x)dx,$$

Desde que ambas as integrais convirjam. O ponto $c \in [a, b]$ é chamado de um valor singular de $g(x)$ ²².

Nestes casos, pode-se aproximar cada integral usando uma quadratura de Gauss-Legendre ou Gauss-Chebyshev, dependendo da forma da função $f(x) = \lambda(x)g(x)$. Aqui vale a mesma observação feita para o uso das quadraturas gaussianas nas integrais impróprias do 1º tipo, além de que o integrando é indefinido para um número finito de pontos e, portanto, o valor da área sob a curva não muda.

Steen e Gelbard (1969) estudaram as quadraturas gaussianas em integrais do tipo $\int_0^\infty e^{-x^2} f(x)dx$ e $\int_0^b e^{-x^2} f(x)dx$, com $b < \infty$. Quando a função $f(x)$ não é definida para $x < 0$ existe dificuldade para avaliar as integrais deste tipo. Tentativas de aplicar a quadratura Gauss-Laguerre em integrais do tipo $\int_0^\infty e^{-x^2} f(x)dx$ usando a transformação de variável $u = x^2 \Rightarrow x = \pm \sqrt{u}$ falham pois geram integrandos com ponto singular na origem, ou seja

$$\int_0^\infty e^{-x^2} f(x)dx = \int_0^\infty e^{-u} f(\sqrt{u}) \frac{du}{2\sqrt{u}}.$$

Aplicar uma quadratura Gauss-Hermite considerando uma função g que seja extensão da função f na origem também pode falhar. Os autores afirmam que se a função f é par a quadratura fornece um resultado exato e deve ser preferida. Porém, se a função f é ímpar a aplicação de sucessivas quadraturas Gauss-Hermite (com número de pontos crescente da quadratura) fornecem resultados que tendem a oscilar em torno da solução correta. Por exemplo, para a

²² Segundo Sperandio, Mendes e Silva (2003) uma alternativa para levantar a singularidade da função integrando é fazer uma troca de variável para obter uma integral não singular.

função par $f(x) = x^2$ aplicando uma quadratura Gauss-Hermite para a função $g(x) = |x|^2$ como mostrado a seguir:

$$\int_0^\infty e^{-x^2} x^2 dx = \frac{1}{2} \int_{-\infty}^\infty e^{-x^2} |x|^2 dx \approx \frac{1}{2} \sum_{i=1}^n w_i |x_i|^2,$$

obtem-se para $n \geq 3$ o valor de 0,4431135 exato. Já no caso da função ímpar $f(x) = x^3$ aplicando uma quadratura Gauss-Hermite para a função $g(x) = |x|^3$ da seguinte forma

$$\int_0^\infty e^{-x^2} x^3 dx = \frac{1}{2} \int_{-\infty}^\infty e^{-x^2} |x|^3 dx \approx \frac{1}{2} \sum_{i=1}^n w_i |x_i|^3,$$

para $n = 3$ resulta em 0,27135045; para $n = 4$ resulta em 0,24097925; para $n = 5$ resulta em 0,2556027; para $n = 6$ resulta em 0,24664255; para $n = 7$ resulta em 0,252573; e assim sucessivamente.

Assim, Steen e Gelbard (1969) propõem uma quadratura especial e apresentam tabelas contendo os pontos e os seus respectivos pesos para resolver estas quadraturas no caso particular de $b = 1$. Eles checam a precisão destes valores comparando os resultados das integrais conhecidas como $\int_0^\infty x e^{-x^2} dx = 0.5 (1 - e^{-b^2})$. Desta forma são obtidos valores com precisão na ordem da 15 casa decimal.

4.1.4 Mudança dos intervalos de integração

a) Intervalos finitos quaisquer

Nas quadraturas Gauss-Jacobi, Gauss-Legendre e Gauss-Chebyshev, quando os limites de integração em $\int_a^b \lambda(x)f(x)dx$ não são $a = -1$ e $b = 1$, pode-se converter a integral para estes limites fazendo uma mudança de variável. Este processo consiste em determinar a equação da reta que passa pelos pontos $(a, -1)$ e $(b, 1)$. Diretamente tem-se que (RALSTON; RABINOWITZ, 1965):

$$x = \frac{(b-a)z + (a+b)}{2} \text{ com } a \leq x \leq b \Rightarrow dx = \frac{(b-a)}{2} dz.$$

Então, para quaisquer valores de a e b tem-se que

$$\int_a^b \lambda(x)f(x)dx = \frac{(b-a)}{2} \int_{-1}^1 \lambda\left(\frac{(b-a)z + (a+b)}{2}\right) f\left(\frac{(b-a)z + (a+b)}{2}\right) dz.$$

A partir daí a integral pode ser resolvida por uma quadratura Gauss-Jacobi, Legendre ou Chebyshev, dependendo da forma da função peso $\lambda(x)$.

Se for resolvido por uma quadratura Gauss-Legendre, em que $\lambda(x) = 1$, pela expressão (4.1.2) a aproximação fica

$$\int_a^b \lambda(x)f(x)dx \approx \frac{(b-a)}{2} \sum_{i=1}^n w_i f\left(\frac{(b-a)z_i + (a+b)}{2}\right), \quad (4.1.5)$$

em que os z'_i s são as raízes dos polinômios de Legendre de grau n .

b) Intervalos infinitos ou semi-infinitos

Nas quadraturas Gauss-Laguerre e Gauss-Hermite caso o integrando não apresente a função $\lambda(x) = e^{-x}$ ou $\lambda(x) = e^{-x^2}$, respectivamente, pode-se proceder ao artifício de multiplicar o integrando pelos termos $(e^{-x}e^x)$ ou $(e^{-x^2}e^{x^2})$ obtendo-se:

$$\int_0^\infty g(x)dx = \int_0^\infty e^{-x}e^x g(x)dx = \int_0^\infty e^{-x} f(x)dx \approx \sum_{i=1}^n w_i f(x_i).$$

ou

$$\int_{-\infty}^{\infty} g(x) dx = \int_{-\infty}^{\infty} e^{-x^2} e^{x^2} g(x) dx = \int_{-\infty}^{\infty} e^{-x^2} f(x) dx \approx \sum_{i=1}^n w_i f(x_i).$$

No caso da quadratura de Gauss-Laguerre se o intervalo de integração não coincide com $[0, \infty]$, sendo $[a, \infty]$ por exemplo, pode-se fazer uma mudança de variável simples do tipo $x = z + a$. Então tem-se:

$$\int_a^{\infty} e^{-x} f(x) dx = e^{-a} \int_0^{\infty} e^{-z} f(z+a) dz \approx e^{-a} \sum_{i=1}^n w_i f(z_i + a),$$

em que os z_i 's são as raízes do polinômio de Laguerre de grau n .

Além disso, nestes casos, tem-se a opção de não usar a quadratura Gauss-Laguerre e aproximar o intervalo semi-infinito de integração para um intervalo finito fazendo uma mudança de variável. Alguns exemplos:

$$\text{a) } \int_a^{\infty} g(x) dx = \int_0^1 g\left(a + \frac{t}{1-t}\right) \frac{1}{(1-t)^2} dt;$$

$$\text{b) } \int_{-\infty}^b g(x) dx = \int_{-1}^0 g\left(b - \frac{t}{1+t}\right) \frac{-1}{(1+t)^2} dt;$$

$$\text{c) } \int_a^{\infty} g(x) dx = \int_0^1 g\left(a + \frac{1-t}{t}\right) \frac{dt}{t^2};$$

$$\text{d) } \int_{-\infty}^b g(x) dx = \int_0^1 g\left(b - \frac{1-t}{t}\right) \frac{dt}{t^2};$$

- para $a = 0$ pode-se ainda usar a transformação

$$\int_0^{\infty} g(x) dx = \int_1^0 g\left(\frac{1}{t} - 1\right) \frac{-dt}{t^2} = \int_0^1 g\left(\frac{1}{t} - 1\right) \frac{dt}{t^2}$$

ou então a transformação

$$\int_0^\infty g(x)dx = \int_1^0 g(-\ln(t)) \frac{-dt}{t} = \int_1^0 g(-\ln(t)) \frac{dt}{t}$$

- se $a > 0$ e $b = \infty$ ou $a = -\infty$ e $b < 0$, ou seja $ab > 0$, e a função a ser integrada decresce para zero na mesma velocidade em que $1/x^2$ se aproxima de zero quando x se aproxima do infinito, uma alternativa mostrada por Chapra e Canale (2010) é proceder uma transformação do tipo

$$\int_a^b g(x)dx = \int_{1/b}^{1/a} \frac{1}{t^2} g\left(\frac{1}{t}\right) dt.$$

- Um caso particular desta situação é quando $[a,b] = [1,\infty)$ e deseja-se transformar a variável para o intervalo $[-1,1]$. Uma forma é inicialmente proceder uma transformação do tipo

$$\int_1^\infty g(x)dx = \int_0^{e^{-1}} \frac{g(-\ln(t))}{t} dt, \quad (4.1.6)$$

e depois com $h(t) = \frac{g(-\ln(t))}{t}$ usar nova transformação do tipo

$$\int_0^{e^{-1}} h(t)dt = \frac{e^{-1}}{2} \int_{-1}^1 h\left(e^{-1} \frac{z+1}{2}\right) dz, \quad (4.1.7)$$

que é uma transformação de um intervalo $[a,b]$ para o intervalo $[-1,1]$.

Nos exemplos acima, qualquer das integrais com intervalos finitos podem sofrer outra mudança de variável, passando para o intervalo de integração $[-1,1]$ e serem resolvidas via quadratura Gauss-Legendre ou Gauss-Chebyshev.

Uma opção à quadratura Gauss-Hermite quando a integral tem intervalo infinito também é transformá-la numa integral com intervalo $[-1,1]$ e avaliá-la por uma quadratura de Gauss-Legendre ou Gauss-Chebyshev. Neste caso, pode-se usar uma mudança de variável do tipo:

$$x = \frac{t}{1-t^2} \quad \text{com } -\infty \leq x \leq \infty, \Rightarrow dx = \frac{1+t^2}{(1-t^2)^2} dt \quad \text{com } -1 \leq t \leq 1,$$

resultando em

$$\int_{-\infty}^{\infty} g(x) dx = \int_{-1}^1 g\left(\frac{t}{1-t^2}\right) \frac{1+t^2}{(1-t^2)^2} dt. \quad (4.1.8)$$

Uma transformação alternativa é inicialmente fazer

$$\int_{-\infty}^{\infty} g(x) dx = \int_0^1 \left[g\left(\frac{1-t}{t}\right) + g\left(-\frac{1-t}{t}\right) \right] \frac{dt}{t^2},$$

seguida de outra mudança de variável passando para o intervalo de integração $[-1,1]$.

4.1.5 Quadraturas gaussianas no *software R*

Na biblioteca “*statmod*” encontra-se a função “*gauss.quad*” que possibilita a obtenção dos nós (x_i ’s) e pesos (w_i ’s) da quadratura desejada fornecendo os seguintes argumentos:

- a) o número de pontos que se deseja realizar a quadratura (n);
- b) o tipo de quadratura (*kind* =), que pode ser: “*legendre*”; “*chebyshev1*”; “*chebyshev2*”; “*hermite*” que neste programa é com os polinômios de Hermite físicos, ou seja, com $\lambda(x) = e^{-x^2}$; “*jacobi*” que neste caso é necessário

especificar o valor de α e de β : e “*laguerre*” cujo padrão é $\alpha = 0$ mas pode ser especificado outro valor de α , contanto que seja maior do que -1 ;

c) valor de α que por padrão é zero;

d) valor de β que por padrão é zero;

4.1.6 Exemplo de resolução de integral por quadratura gaussiana com o uso do software R

Calcular a probabilidade $P(Z > 1,7) = \int_{1,7}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$.

A função a ser integrada $g(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ é a densidade da normal padrão e está representada na Figura 4. A área hachurada é a que se deseja quantificar.

Pela simetria da distribuição de probabilidade normal padrão tem-se que

$$P = \int_{1,7}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 0.5 - \int_0^{1,7} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz.$$

Fazendo uma mudança de variável de z para x , com $z = \frac{1,7x+1,7}{2}$, o intervalo de integração passa de $[0, 1,7]$ para $[-1,1]$ resultando em:

$$P = \int_0^{1,7} g(z) dz = \frac{1,7}{2} \int_{-1}^1 f\left(\frac{1,7x+1,7}{2}\right) dx.$$

A função a ser integrada agora é $f(x) = \frac{1,7}{2\sqrt{2\pi}} e^{-\frac{(1,7x+1,7)^2}{8}}$ no intervalo $[-1,1]$ e está representada na Figura 5. Resolvendo essa integral por uma quadratura de Gauss-Legendre, com $\lambda(x) = 1$, para $n = 6$ pontos na quadratura, a Figura 6 traz uma representação gráfica dessa aproximação. Nesta representação, a quadratura consiste em aproximar a área sob a curva por áreas de n regiões retangulares, cujas base são os pesos w_j de cada parcela do somatório e a altura de cada retângulo é a imagem da função $f(x)$ aplicada no ponto $x = t_j$.

Analiticamente, tem-se

$$\int_{-1}^1 \lambda(x) f(x) dx \approx \frac{1,7}{2\sqrt{2\pi}} \sum_{j=1}^6 w_j e^{-\frac{(1,7t_j+1,7)^2}{8}},$$

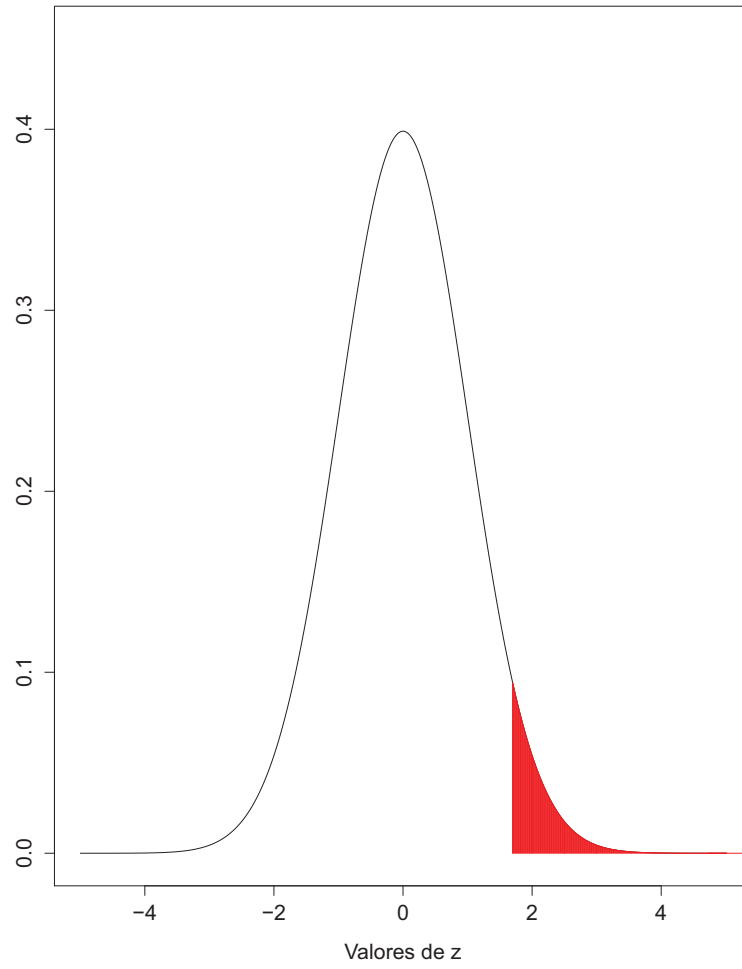


Figura 4 Representação da função $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ e sua integral no intervalo $(1,7, \infty)$

ou seja

$$P \approx \frac{1,7}{2\sqrt{2\pi}} \left[w_1 e^{-\frac{(1,7t_1+1,7)^2}{8}} + w_2 e^{-\frac{(1,7t_2+1,7)^2}{8}} + w_3 e^{-\frac{(1,7t_3+1,7)^2}{8}} \right. \\ \left. + w_4 e^{-\frac{(1,7t_4+1,7)^2}{8}} + w_5 e^{-\frac{(1,7t_5+1,7)^2}{8}} + w_6 e^{-\frac{(1,7t_6+1,7)^2}{8}} \right].$$

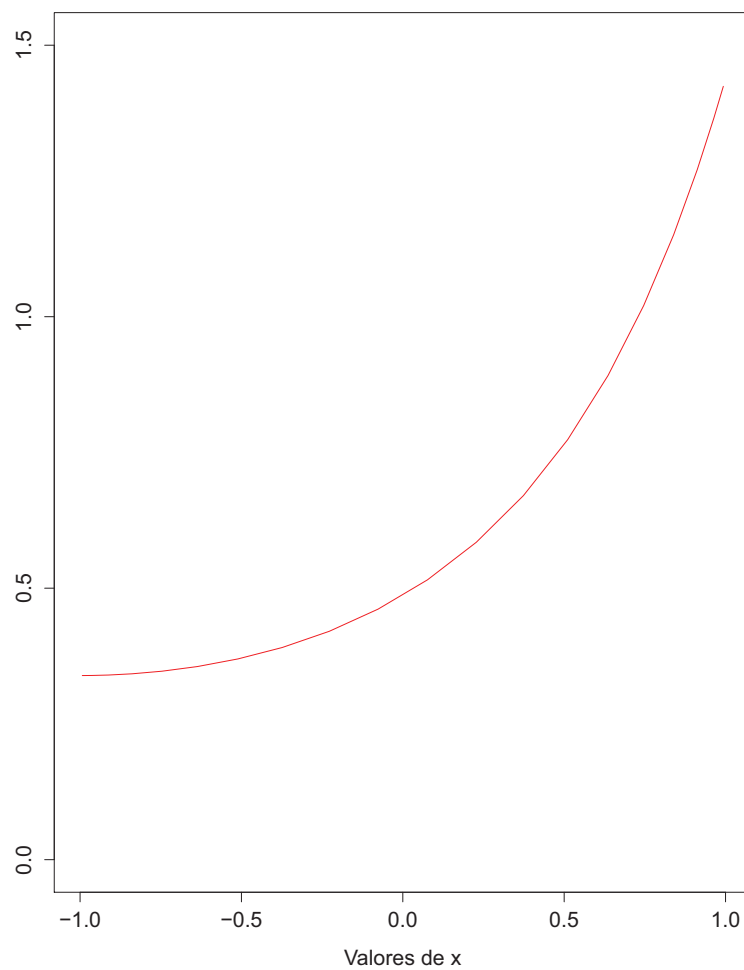


Figura 5 Representação da função $f(x) = \frac{1,7}{2\sqrt{2\pi}}e^{-\frac{(1,7x+1,7)^2}{8}}$ a ser integrada no intervalo $(-1, 1)$.

Realizando as operações necessárias verifica-se que o valor da integral é 0,04456546.

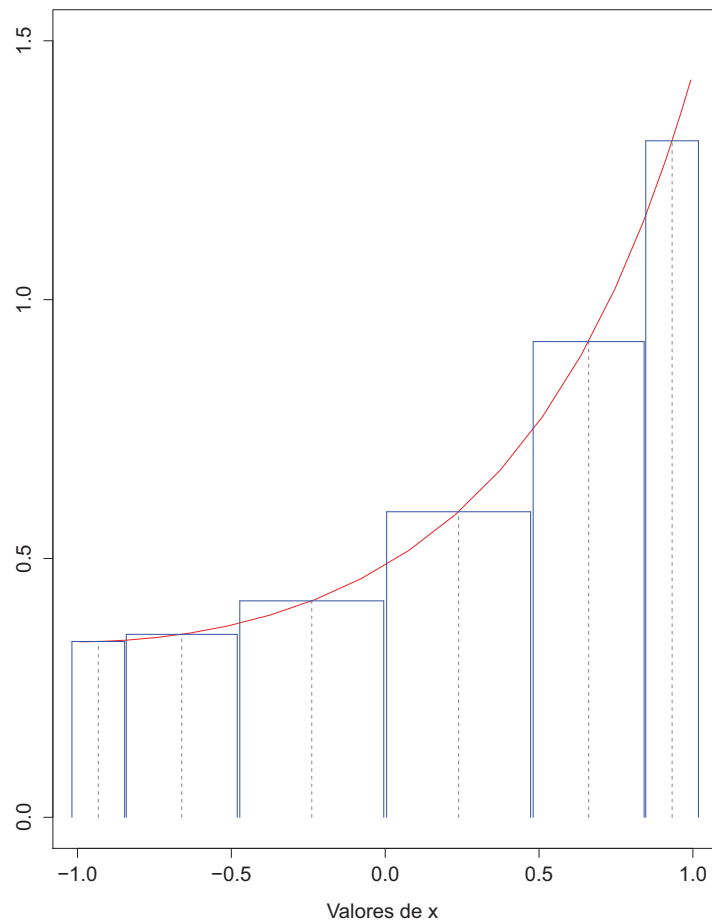


Figura 6 Representação da função $f(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$ e sua integral no intervalo $(-1, 1)$.

4.2 Métodos numéricos para a determinação de raízes de equações

Nas diversas funções densidade ou de distribuição das estatísticas estudadas, quando se deseja encontrar o valor do quantil ou o valor do parâmetro de não-centralidade para determinada probabilidade acumulada p , será necessário resolver equações da forma $f(x) - p = 0$. Pela complexidade das expressões, geralmente estas equações não possuem soluções analíticas sendo

necessário utilizar um método numérico iterativo para encontrar as soluções ou raízes destas equações. A seguir são descritos os métodos iterativos numéricos usuais mais utilizados.

4.2.1 Método do falso positivo

Este método é útil quando não se tem informação a respeito da localização da raiz x_ξ procurada. Neste caso, pelo Teorema de Bolzano²³, basta procurar dois pontos em que a função tenha sinais opostos e aplicar o método, que possui convergência linear, mesmo que algumas vezes lenta (BARROSO et al., 1987; RALSTON; RABINOWITZ, 1965).

Para obter a aproximação da raiz x_ξ desejada por este método, os seguintes passos devem ser executados de forma iterativa:

1. toma-se um intervalo $[x_1, x_2]$ e seus respectivos valores $f(x_1)$ e $f(x_2)$, tal que $f(x_1) \times f(x_2) < 0$;
2. calcula-se o valor de x_n , que é a interseção da reta que passa pelos pontos $(x_1, f(x_1))$ e $(x_2, f(x_2))$ e o eixo das abscissas, pela expressão

$$x_n = \frac{f(x_2)x_1 - f(x_1)x_2}{f(x_2) - f(x_1)};$$

3. avalia-se o valor da função neste ponto. Se o valor de $f(x_n)$ aproxima-se de zero com a precisão desejada, x_n é a raiz x_ξ procurada. Caso contrário determina-se um novo intervalo dado por:

- a) $[x_1, x_n]$ se $f(x_1)$ e $f(x_n)$ tenham sinais opostos, isto é, se $f(x_1) \times f(x_n) < 0$;

²³ O Teorema de Bolzano é um caso particular do Teorema do valor intermediário em que o valor da função no ponto é nula. Diz que: “dada uma função real $f(x)$, contínua em um intervalo fechado $[a, b]$, com $f(a)$ e $f(b)$ de sinais contrários, existe pelo menos um ponto $c \in (a, b)$ tal que $f(c) = 0$ ”.

b) $[x_n, x_2]$ se $f(x_1)$ e $f(x_n)$ tenham sinais iguais;

4. repete-se o processo até que o valor de x_n convirja para a raiz da função x_ξ , com a precisão $\epsilon = x_n - x_\xi$ desejada.

O método falha se o mesmo intervalo é obtido duas vezes seguidas. Neste caso pode-se proceder a um mecanismo de proteção calculando-se x_n de uma das seguintes formas

$$x_n = \frac{\frac{1}{2}f(x_2)x_1 - f(x_1)x_2}{\frac{1}{2}f(x_2) - f(x_1)} \text{ ou } x_n = \frac{f(x_2)x_1 - \frac{1}{2}f(x_1)x_2}{f(x_2) - \frac{1}{2}f(x_1)},$$

dando um peso maior a um dos extremos do intervalo para forçar que o próximo x_k ocorra neste lado da função e evitando intervalos seguidos iguais. Esta é uma forma de contornar a falha do método e garantir convergência. Outras formas podem ser utilizadas, como a que deu origem ao método de *Pégaso* ou *Pégasus* descrito a seguir.

O método do falso positivo geralmente é não estacionário, pois em cada iteração n o ponto x_i tomado do passo anterior, pode variar entre $i = n - 1$ ou $n - 2$ dependendo de qual valor atender a condição $f(x_n) \times f(x_i) < 0$. Ralston e Rabinowitz (1965) afirma que para funções convexas entre x_1 e x_2 o método do falso positivo é estacionário, ou seja, sempre o ponto x_i a ser tomado em todas as iterações é fixo com $i = 1$ ou 2 . Por exemplo, se $x_i = x_1$ tem-se

$$x_n = x_1 \frac{f(x_{n-1})}{f(x_{n-1}) - f(x_1)} - x_{n-1} \frac{f(x_1)}{f(x_1) - f(x_{n-1})}, \quad n = 3, 4, \dots,$$

e o novo intervalo será sempre $[x_1, x_n]$.

O método da secante é o método do falso positivo para funções convexas em que, para melhorar a velocidade de convergência do método, abriu-se mão da exigência de sinais opostos da função nos dois pontos usados para

gerar o próximo (KHURI, 2003; LIMA, 2007)²⁴. No método da secante fixou-se que sempre seria usado x_i e x_{i-1} para gerar x_{i+1} . Porém, com isso, foi perdida a garantia da convergência.

Nesse método a iteração só irá convergir se os valores iniciais tomados x_i e x_{i-1} forem suficientemente próximos do valor procurado.

4.2.2 Método da secante

O método da secante é um caso particular do método do falso positivo. Seu algoritmo faz uso de uma sucessão de raízes de linhas secantes à $f(x)$ que se aproximam cada vez mais da raiz da função $f(x)$.

Tomam-se dois pontos $A = [x_{n-1}, f(x_{n-1})]$ e $B = [x_n, f(x_n)]$ pertencentes à função $f(x)$, pelos quais constrói-se uma reta s_1 passando por eles, como ilustrado na Figura 7. Essa reta é uma secante ou corda do gráfico da função f , cuja equação é dada por

$$\begin{vmatrix} x_{s_1} & y_{s_1} & 1 \\ x_n & f(x_n) & 1 \\ x_{n-1} & f(x_{n-1}) & 1 \end{vmatrix} = 0$$

²⁴ (KHURI, 2003; LIMA, 2007) Uma função $f: D \Rightarrow \mathbb{R}$ é convexa se $f[tx_1 + (1-t)x_2] \leq tf(x_1) + (1-t)f(x_2)$ para todo x_1 e $x_2 \in D$ e para qualquer $t \in [0,1]$. Geometricamente, significa que o seu gráfico se situa abaixo de qualquer de suas secantes, ou seja, tomando-se A e B dois pontos quaisquer pertencentes à curva da função $f(x)$ a região do gráfico entre A e B está abaixo da secante AB. (LIMA, 2007) Se a função f é convexa então f' é

$$f'(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f'(x_2)$$

que significa que o segmento de reta entre dois pontos x_1 e x_2 está contida entre os valores extremos de sua derivada. Assim, a derivada é monótona não-decrescente em todo o domínio de f' , o que geometricamente significa que o gráfico da função está situado acima de qualquer de suas tangentes. Neste caso, $f''(x) \geq 0$ para todo $x \in D$.

$$\begin{aligned}
x_{s_1}f(x_n) + x_nf(x_{n-1}) + y_{s_1}x_{n-1} - x_{n-1}f(x_n) - x_{s_1}f(x_{n-1}) - y_{s_1}x_n &= 0 \\
y_{s_1}[x_{n-1} - x_n] - f(x_n)[x_{n-1} - x_{s_1}] + f(x_{n-1})[x_n - x_{s_1}] &= 0 \\
y_{s_1}[x_{n-1} - x_n] - f(x_n)[x_{n-1} + (-x_n + x_n) - x_{s_1}] + f(x_{n-1})[x_n - x_{s_1}] &= 0 \\
y_{s_1}[x_{n-1} - x_n] - f(x_n)[x_{n-1} - x_n] - f(x_n)[x_n - x_{s_1}] + f(x_{n-1})[x_n - x_{s_1}] &= 0 \\
(y_{s_1} - f(x_n))[x_{n-1} - x_n] - (f(x_n) - f(x_{n-1}))[x_n - x_{s_1}] &= 0
\end{aligned}$$

ou seja

$$y_{s_1}(x_{s_1}) = \frac{f(x_n) - f(x_{n-1})}{(x_{n-1} - x_n)} \cdot (x_n - x_{s_1}) + f(x_n)$$

Se x_{n+1} é a raiz da reta s_1 , então $y_{s_1}(x_{n+1}) = 0$. Resolvendo essa equação obtém-se a relação de recorrência para o método das secantes dada por:

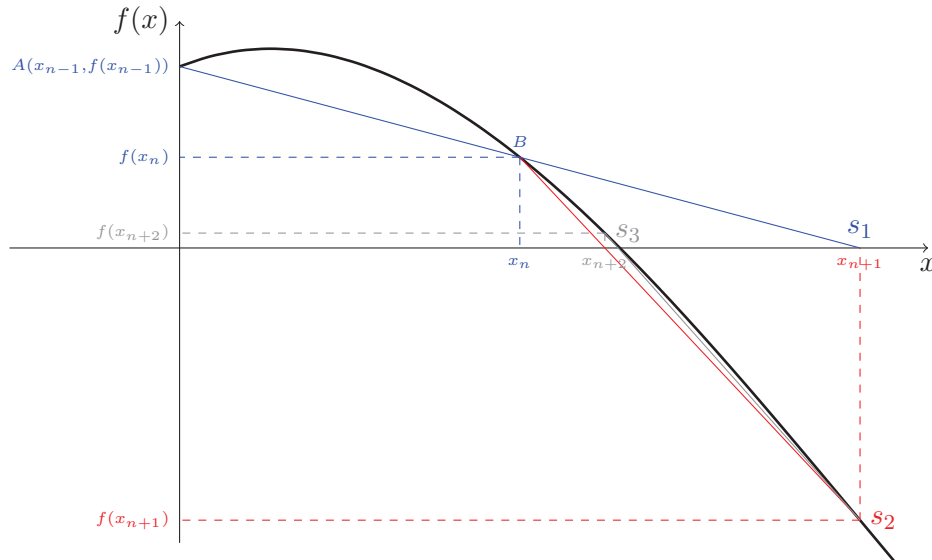


Figura 7 Representação geométrica do método das secantes

$$x_{n+1} = x_n - f(x_n) \cdot \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}, \quad n = 1, 2, \dots \quad (4.2.1)$$

Este método depende de dois valores iniciais $x_{n-1} = x_0$ e $x_n = x_1$. Em seguida, pela fórmula (4.2.1) obtém-se um novo valor x_2 , que será utilizado no lugar de x_0 para repetir o processo.

Determina-se assim x_3 , que substituirá x_1 continuando o processo até chegar a um nível de precisão desejado, em que a diferença entre x_{n+1} e x_n é suficientemente pequena.

Em análise numérica a velocidade na qual uma sucessão converge ao seu limite é chamado de ordem de convergência. Segundo Ralston e Rabinowitz (1965), as iterações do método da secante convergem para uma raiz de $f(x)$ se os valores iniciais x_0 e x_1 estiverem próximas da raiz, com ordem de convergência de $(1 + \sqrt{5})/2 \approx 1,618$, que é a razão áurea. Além disso, a função deve ser duas vezes continuamente diferenciável e a raiz em questão deve ser simples (raiz de multiplicidade 1) no intervalo $[x_0, x_1]$ para assegurar a convergência do método.

Uma vantagem do método da secante sobre outros métodos como o de Newton-Raphson, é que ele só requer a avaliação da função e não de suas derivadas, que podem ser expressões difíceis de serem obtidas.

4.2.3 Método de Pégaso ou Pégasus

É um método também baseado no método do falso positivo modificado para acelerar a convergência. Segundo Barroso et al. (1987), dada uma função $f(x)$ contínua no intervalo $[x_0, x_1]$ e tal que $f(x_0) \times f(x_1) < 0$, a raiz existente neste intervalo pode ser obtida pela fórmula de recorrência

$$x_{n+1} = x_n - f(x_n) \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}, \quad n = 1, 2, \dots,$$

onde as aproximações da iteração seguinte são escolhidas de forma que se

$$\begin{cases} f(x_{n+1})f(x_n) < 0, & \text{então } [x_{n-1}, f(x_{n-1})] \text{ é trocado por } [x_n, f(x_n)]; \\ f(x_{n+1})f(x_n) > 0, & \text{então } [x_{n-1}, f(x_{n-1})] \text{ é trocado por } \left[x_{n-1}, \frac{f(x_{n-1})f(x_n)}{[f(x_n) + f(x_{n+1})]} \right]. \end{cases}$$

Em ambos os casos $[x_n, f(x_n)]$ é trocado por $[x_{n+1}, f(x_{n+1})]$, o que garante que os valores da função usados a cada iteração tenham sempre sinais opostos. Barroso et al. (1987) confirmam que a ideia do método de Pégaso é reduzir $f(x_{n-1})$ por um fator $[f(x_n)/(f(x_n) + f(x_{n+1}))]$ para evitar a retenção de um ponto e obter uma convergência mais rápida.

4.2.4 Método de Newton-Raphson

O método de Newton-Raphson²⁵ é outro método de análise numérica cujo objetivo é estimar as raízes de uma função. Segundo Barroso et al. (1987), dada uma função $f(x)$ contínua no intervalo $[x_0, x_1]$, com as derivadas $f'(x)$ e $f''(x)$ também contínuas e $f'(x) \neq 0$, tal que exista uma única raiz x_ξ neste intervalo, tomando-se uma aproximação inicial para x_ξ e fazendo uma expansão em Série de Taylor de primeira ordem para $f(x) = 0$ obtém-se uma fórmula de recorrência para o valor aproximado da raiz x_ξ .

Segundo Khuri (2003) e Lima (2007), a série de Taylor é dada por

$$f(x) = f(a) + \sum_{n=1}^{\infty} \frac{(x-a)^n}{n!} f^{(n)}(a),$$

com $f : [a, b] \rightarrow \mathbb{R}$, n vezes derivável no intervalo (a, b) e todas as $(n-1)$ -ésimas derivadas contínuas em $[a, b]$, com $x \in [a, b]$.

Nestas condições, existe um valor $c \in (a, b)$ tal que, fazendo uma expansão em série de Taylor de grau $n = 1$, tem-se

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2 f''(c)}{2!},$$

²⁵ O método é atribuído a Sir. Isaac Newton (1643-1727) e Joseph Raphson (1648-1715).

em que $a < c < x$.

Substituindo-se x por x_ξ e, como $f(x_\xi) = 0$ tem-se

$$0 = f(a) + (x_\xi - a)f'(a) + \frac{(x_\xi - a)^2 f''(c)}{2!}.$$

Se a aproxima-se de x_ξ , o último termo à direita da expressão anterior será pequeno em comparação à soma dos dois primeiros termos e, por isso, pode-se desprezá-lo e aproximar a expressão para

$$\begin{aligned} 0 \approx f(a) + (x_\xi - a)f'(a) &\Rightarrow x_\xi \cdot f'(a) \approx -f(a) + a \cdot f'(a) \\ &\Rightarrow x_\xi \approx a - \frac{f(a)}{f'(a)}. \end{aligned}$$

Essa expressão é usada para definir a aproximação da raiz da função f , chamando $a = x_n$ e $x_\xi = x_{n+1}$, dada por

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, \dots, \quad (4.2.2)$$

em que:

- a) n indica a n -ésima iteração do algoritmo;
- b) $f(x_n)$ é o valor da função f no ponto x_n ;
- c) $f'(x_n)$ é o valor da derivada da função f no ponto x_n ;
- d) x_{n+1} é uma aproximação de x_ξ .

Geometricamente, o método de Newton-Raphson é equivalente a substituir um pequeno arco da curva $f(x)$ por uma reta tangente traçada a partir de um ponto x_n deste arco, como mostra a Figura 8. O número x_{n+1} é a abscissa do ponto em que a tangente ao gráfico de f no ponto x_n encontra o eixo horizontal.

Segundo Lima (2007), o método de Newton-Raphson baseia-se no fato que, se a tangente aproxima a curva então sua interseção com o eixo das abscissas aproxima o ponto de interseção da curva com esse eixo, ou seja, o ponto x_ξ raiz de f .

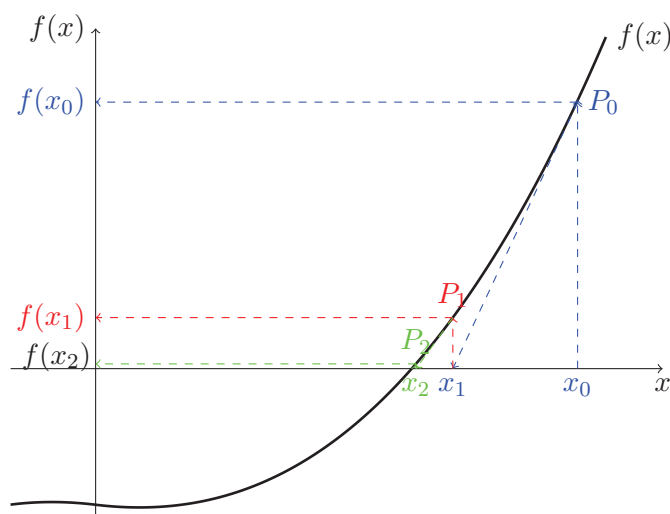


Figura 8 Representação geométrica do método de Newton-Raphson

Para obter a aproximação da raiz x_ξ desejada no método de Newton-Raphson, os seguintes passos devem ser executados de forma iterativa:

1. toma-se um valor para $x_0 \in [a, b]$ e seu respectivo valor $f(x_0)$;
2. calcula-se o valor da equação da tangente (derivada) da função nesse ponto, ou seja, $f'(x_0)$;

3. determina-se o intercepto da reta tangente nesse ponto ao eixo das abscissas, denominado de x_1 . Através da definição de derivada de uma função num ponto que é a inclinação da reta tangente naquele ponto, isto é

$$f'(x_0) = \frac{\Delta y}{\Delta x} = \frac{f(x_0) - 0}{x_0 - x_1} \Rightarrow x_1 = x_0 - \frac{f(x_0)}{f'(x_0)};$$

4. calcula-se o valor da função nesse novo ponto x_1 , ou seja $f(x_1)$;
5. repete-se o processo n vezes, até que o valor de x_n convirja para a raiz da função x_ξ , com a precisão $\epsilon = x_n - x_\xi$ desejada.

4.3 Método do jacobiano de transformação de variáveis

O método do jacobiano é um método utilizado em cálculo diferencial em diversas situações, como para a mudança de variáveis em integrais múltiplas. Em estatística, o método do jacobiano é utilizado para obter a distribuição de transformações de variáveis aleatórias contínuas.

Definição 1: Dada a função $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, definida por um vetor de m componentes, sendo cada componente uma função $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$. Se todas as funções f_i e suas derivadas parciais são funções contínuas do vetor aleatório $X = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, as derivadas parciais dessas funções podem ser organizadas numa matriz $m \times n$, que é denominada matriz jacobiana.²⁶

A matriz jacobiana é a matriz formada pelas derivadas parciais de primeira ordem de uma função vetorial. A matriz jacobiana tem em cada uma das suas m linhas as derivadas parciais de uma f_i , com $i = 1, \dots, m$, em relação aos n valores de x_j , com $j = 1, \dots, n$, ou seja,

²⁶ Adaptado de Khuri (2003) e Leithold (1994).

$$\mathbf{J}_f(\mathbf{X}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

A matriz jacobiana pode ser representada também por $\mathbf{J}_f(x_1, x_2, \dots, x_n)$, ou ainda por $\frac{\partial(f_1, f_2, \dots, f_m)}{\partial(x_1, x_2, \dots, x_n)}$.

Definição 2: Denomina-se de jacobiano o determinante da matriz jacobiana quando $m = n$ e este determinante é diferente de zero, e representa-se por $|\mathbf{J}_f(\mathbf{X})|$.

Dado o vetor aleatório de transformações

$$\begin{aligned} \mathbf{Y} &= [y_1, y_2, \dots, y_n] \\ &= [f_1(x_1, x_2, \dots, x_n), f_2(x_1, x_2, \dots, x_n), \dots, f_n(x_1, x_2, \dots, x_n)] \\ &= [f_1(\mathbf{X}), f_2(\mathbf{X}), \dots, f_n(\mathbf{X})], \\ &= \mathbf{f}(\mathbf{X}), \end{aligned}$$

com $|\mathbf{J}_f(\mathbf{X})| \neq 0$, se tais transformações são bijetoras, então

$$\begin{aligned} \mathbf{X} &= [x_1, x_2, \dots, x_n] \\ &= [h_1(y_1, y_2, \dots, y_n), h_2(y_1, y_2, \dots, y_n), \dots, h_n(y_1, y_2, \dots, y_n)] \\ &= [h_1(\mathbf{Y}), h_2(\mathbf{Y}), \dots, h_n(\mathbf{Y})] \\ &= \mathbf{h}(\mathbf{Y}), \end{aligned}$$

ou seja

$$\mathbf{Y}^{-1} = \mathbf{f}^{-1}(\mathbf{X}) \Rightarrow \mathbf{X} = \mathbf{h}(\mathbf{Y}).$$

Assim, a função $f_Y(\mathbf{y}) = f_X[\mathbf{h}(\mathbf{Y})] \cdot |\mathbf{J}_h(\mathbf{Y})|$, com $|\mathbf{J}_h(\mathbf{Y})| = |\mathbf{J}_f(\mathbf{X})|^{-1}$

$$\mathbf{J}_h(\mathbf{y}) = \begin{pmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} & \dots & \frac{\partial h_1}{\partial y_n} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} & \dots & \frac{\partial h_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_n}{\partial y_1} & \frac{\partial h_n}{\partial y_2} & \dots & \frac{\partial h_n}{\partial y_n} \end{pmatrix}.$$

Uma das condições para a aplicação do método do jacobiano é de que a função $\mathbf{f} : \mathbf{X} \rightarrow \mathbf{Y}$ seja bijetora. Se este não for o caso, mas se o domínio da função puder ser particionado em subconjuntos disjuntos, de forma que em cada subconjunto a função seja bijetora, pode-se aplicar o método do jacobiano em cada subconjunto e somar o resultado para obter a função desejada. Assim, sendo $i = 1, \dots, k$ os i -ésimos subconjuntos do domínio de $\mathbf{f}$ e $f^{(i)}$ a função relativa ao subconjunto P_i , com $h^{(i)}$ a sua respectiva função inversa, tem-se

$$f_Y(\mathbf{y}) = \sum_{i=1}^k f_X[h^{(i)}(\mathbf{Y})] \cdot |J_{h^{(i)}}(\mathbf{Y})|.$$

5 DISTRIBUIÇÕES DE PROBABILIDADE

Nesta seção é apresentada uma revisão das distribuições de probabilidades que são utilizadas para obter as estatísticas do teste de Dunnett unilateral e bilateral, com dados balanceados e não-balanceados.

A variável S é definida como a razão entre a variância residual amostral combinada e a variância populacional e sua distribuição está relacionada com a distribuição χ^2 . Ela é utilizada nos testes de Dunnett pois a distribuição da estatística para comparar as diversas médias com a média da testemunha considera os graus de liberdade associados à variância residual amostral.

A distribuição t de *Student* é necessária para melhor compreender a sua distribuição multivariada. Também é apresentada a distribuição t não-central univariada com sua função densidade e de distribuição e algumas de suas características. A obtenção desta distribuição tem a finalidade de completude, pois o teste de Dunnett para o caso não-central generaliza a situação univariada para uma multivariada, porém considerando as correlações entre as variáveis.

Como as estatísticas dos testes de Dunnett são aplicações das distribuições do máximo da t multivariada, para o teste unilateral, e do máximo do módulo da t multivariada, para o teste bilateral, fez-se uma apresentação detalhada da obtenção de suas funções de densidade e de distribuição. Para chegar a estas distribuições foi necessário obter a função de densidade e a de distribuição do $\max(\mathbf{Z}) = \max_i(Z_i)$ e do $\max(|\mathbf{Z}|) = \max_i(|Z_i|)$, o que também é mostrado com pormenores.

5.1 Distribuição da variável aleatória $S = \hat{\sigma}/\sigma$

Considerando uma amostra aleatória de tamanho n da distribuição normal $Y_1, Y_2, \dots, Y_n$, com média μ e variância σ^2 . O estimador não viesado da variância populacional σ^2 é dado por

$$\hat{\sigma}^2 = \frac{\sum_{k=1}^n (Y_k - \bar{Y})^2}{n-1}.$$

A média estimada $\bar{Y}$ é independente da variância amostral estimada $\hat{\sigma}^2$. A variável

$$U = \frac{\nu \hat{\sigma}^2}{\sigma^2} \quad (5.1.1)$$

possui uma distribuição qui-quadrado com $\nu = n - 1$ graus de liberdade, ou seja, $U \sim \chi_{(\nu)}^2$. A função densidade de probabilidade de U é dada por

$$f(u; \nu) = \frac{1}{2^{\nu/2} \Gamma(\frac{\nu}{2})} u^{\nu/2-1} e^{-u/2}, \quad (5.1.2)$$

²⁶ U pode ser definida também como

$$U = \frac{\sum_{k=1}^n (Y_k - \bar{Y})^2}{\hat{\sigma}^2} = \sum_{k=1}^n (Z_k - \bar{Z})^2,$$

sendo que $Z_k \sim N(0,1)$, para todo k e $\bar{Z} = \frac{\sum_{k=1}^n Z_k}{n}$.

e está representada na Figura 9, em que $u > 0$. A média e a variância de U^{26} são $\mu_U = \nu$ e $\sigma_U^2 = 2\nu$

A partir da distribuição de U pode-se obter a distribuição de $S = \hat{\sigma}/\sigma$ dada por

$$S = \sqrt{\frac{U}{\nu}} = \sqrt{\frac{\nu \hat{\sigma}^2}{\sigma^2}} = \frac{\hat{\sigma}}{\sigma}. \quad (5.1.3)$$

Teorema 5.1.1. (Distribuição de $S = \hat{\sigma}/\sigma$)

Definindo a variável aleatória S como em (5.1.3), $S \sim \sqrt{\chi_{(\nu)}^2/\nu}$, a sua função densidade de probabilidade é dada por

$$f(s; \nu) = \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2}, \quad s \geq 0. \quad (5.1.4)$$

para $U \sim \chi_{(\nu)}^2$ cuja função densidade é apresentada em (5.1.2).

Demonstração:

Sendo $S = \sqrt{U/\nu}$, a sua transformação inversa é $U = \nu S^2$. A função densidade de S deriva da função densidade da variável resultante da transformação com o uso do jacobiano da transformação, que é a derivada de primeira ordem de u com relação a s , ou seja,

$$J = \frac{du}{ds} = 2\nu s$$

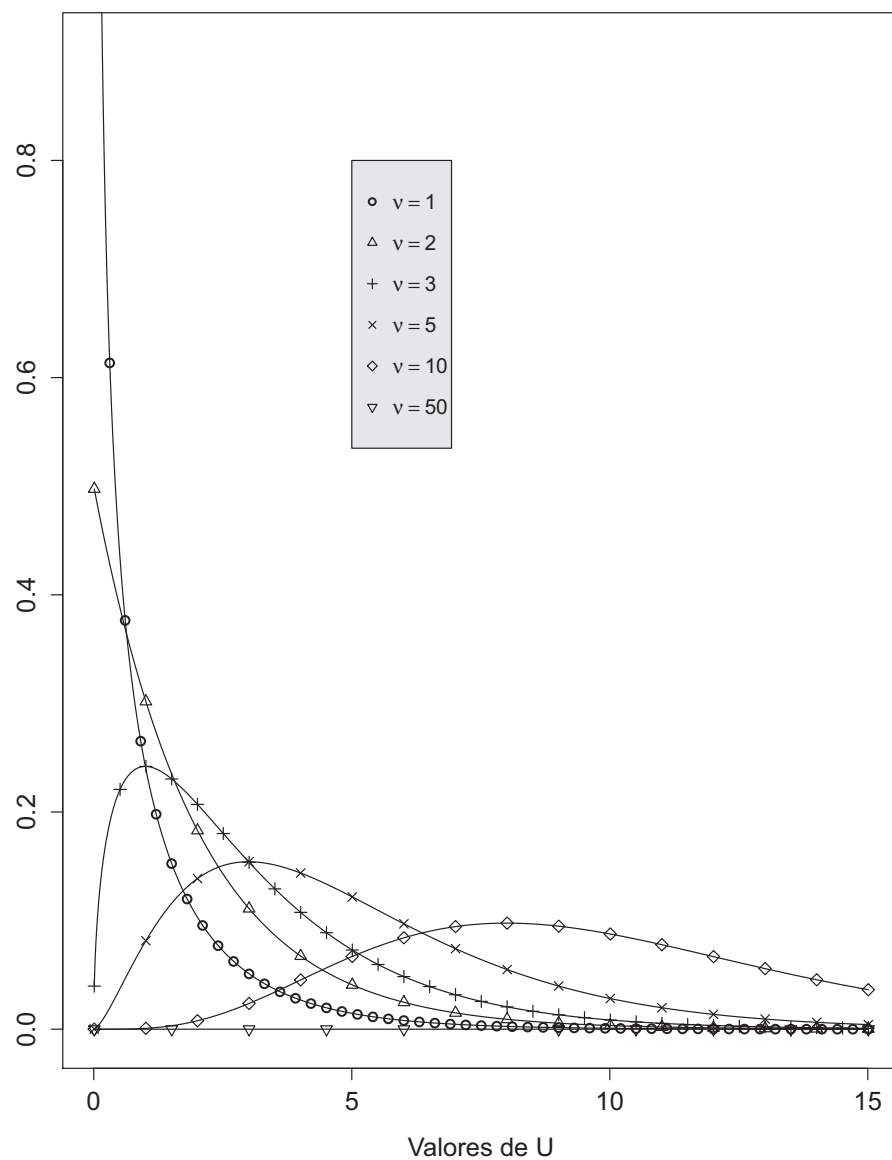


Figura 9 Representação gráfica da densidade da variável aleatória $U = \nu \hat{\sigma}^2 / \sigma^2$, dada pela equação (5.1.2), para diferentes valores de graus de liberdade ν

Assim, a função densidade de S , com $u = \nu s^2$ é obtida por

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

como se queria mostrar.

Na Figura 10 está a representação gráfica da função de densidade da variável aleatória S , com $s \geq 0$.

5.2 Distribuição t de *Student*

Segundo Ferreira (2009), a distribuição t de *Student* é definida pela razão entre duas variáveis aleatórias independentes, uma normal padronizada por outra originada da raiz quadrada de uma variável aleatória qui-quadrado dividida pelos seus graus de liberdade. Assim, considerando uma amostra aleatória de tamanho n da distribuição normal, $Y_1, Y_2, \dots, Y_n$, com média μ e variância σ^2 , então

$$Z = \frac{\bar{Y} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad (5.2.1)$$

possui distribuição normal padrão, $Z \sim \mathbb{N}(0,1)$, cuja função densidade de probabilidade é

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad (5.2.2)$$

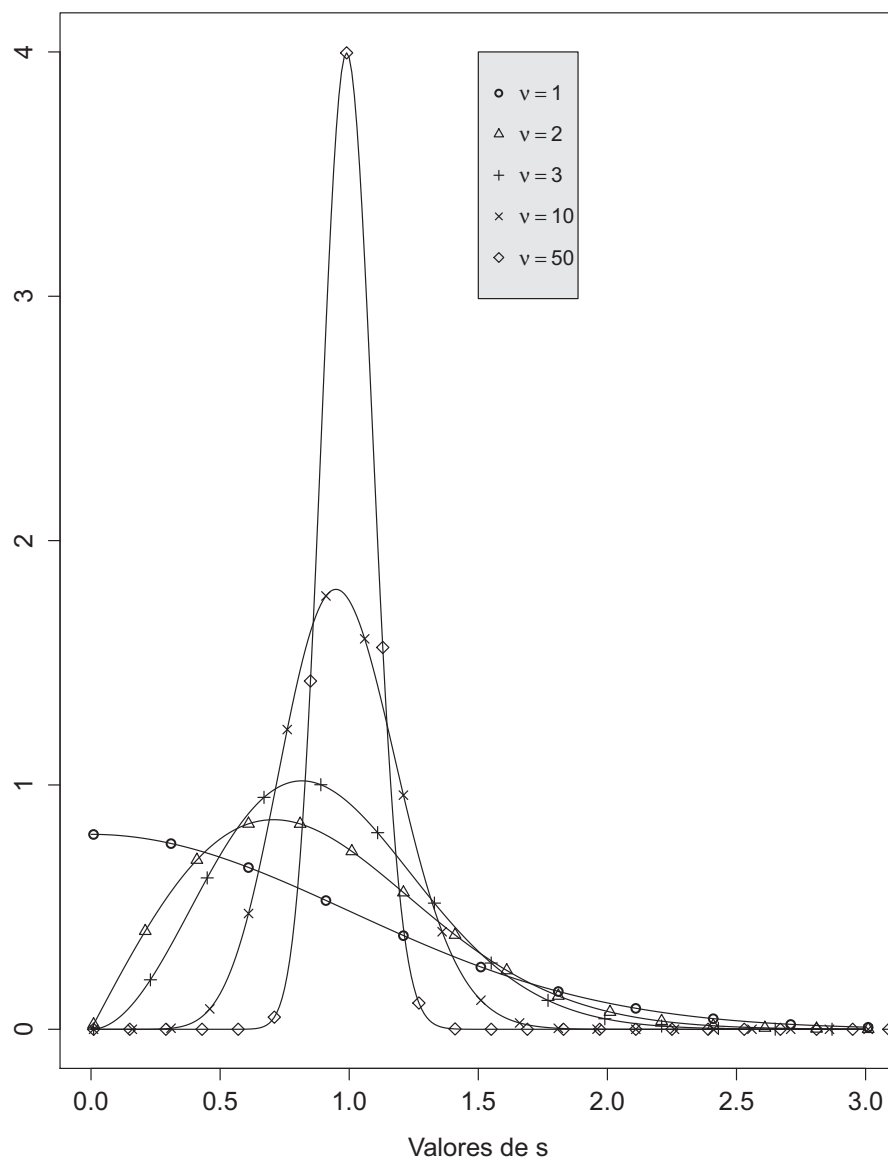


Figura 10 Representação gráfica da densidade da variável aleatória $S = \hat{\sigma}/\sigma = \sqrt{U/v}$, dada pela equação (5.1.4), para diferentes valores de graus de liberdade v

e a variável aleatória U dada por (5.1.1), com $U \sim \chi^2_{(\nu)}$. A partir de Z e U pode-se definir

$$X = \frac{Z}{\sqrt{\frac{U}{\nu}}} = \frac{\frac{\bar{Y} - \mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{\nu \hat{\sigma}^2}{\sigma^2}}}} = \frac{\bar{Y} - \mu}{\frac{\hat{\sigma}}{\sqrt{n}}}. \quad (5.2.3)$$

Teorema 5.2.1. (Função densidade da distribuição t de Student)

Definindo a variável aleatória X como em (5.2.3), X tem distribuição t de Student, $X \sim t(\nu)$, cuja função densidade de probabilidade é dada por

$$f_X(x; \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{\nu\pi}} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad (5.2.4)$$

com ν graus de liberdade o único parâmetro da distribuição e $S = \sqrt{U/\nu} \sim \sqrt{\chi^2_{(\nu)}/\nu}$ cuja função densidade é apresentada em (5.1.4).

Demonstração:

Como Z e U são independentes, tem-se que

$$\begin{aligned} f_{Z,S}(z,s;\nu) &= f_S(s;\nu) \cdot \phi(z) \\ &= \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}. \end{aligned}$$

Usando a transformação $Z = XS$ e $S = S$, a função densidade conjunta de X e S deriva da função densidade conjunta de Z e S com o uso do Jacobiano da transformação. O jacobiano da transformação é dado por

$$\mathbf{J} = \begin{vmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial s} \\ \frac{\partial s}{\partial x} & \frac{\partial s}{\partial s} \end{vmatrix} = \begin{vmatrix} s & x \\ 0 & 1 \end{vmatrix} = s.$$

Assim $f_{X,S}(x,s;\nu) = f_S(s;\nu)\phi(xs)\text{abs}(\mathbf{J})$, ou seja

$$\begin{aligned} f_{X,S}(x,s;\nu) &= \frac{\nu^{\nu/2}}{2^{(\nu/2)-1}\Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-(x^2 s^2)/2} s \\ &= \frac{\nu^{\nu/2} s^\nu e^{-\nu s^2/2} e^{-(x^2 s^2)/2}}{\sqrt{2\pi} 2^{(\nu/2)-1}\Gamma(\frac{\nu}{2})} \\ f_{X,S}(x,s;\nu) &= \frac{\nu^{\nu/2} s^\nu e^{-s^2(x^2+\nu)/2}}{\sqrt{2\pi} 2^{(\nu/2)-1}\Gamma(\frac{\nu}{2})}. \end{aligned}$$

A função densidade marginal da variável X é obtida integrando-se $f_{X,S}(x,s;\nu)$ em relação a s , de 0 a ∞ que é o domínio da variável S . Assim tem-se:

$$\begin{aligned} f_X(x;\nu) &= \int_0^\infty \frac{\nu^{\nu/2} s^\nu e^{-s^2(x^2+\nu)/2}}{\sqrt{2\pi} 2^{(\nu/2)-1}\Gamma(\frac{\nu}{2})} ds \\ &= \frac{\nu^{\nu/2}}{\sqrt{2\pi} 2^{(\nu/2)-1}\Gamma(\frac{\nu}{2})} \int_0^\infty s^\nu e^{-s^2(x^2+\nu)/2} ds. \end{aligned}$$

Chamando de $t = s^2(x^2 + \nu)/2$, então $dt = s(x^2 + \nu)ds$.

Também

$$t^{\frac{(\nu-1)}{2}} = \left(\frac{s^2(x^2 + \nu)}{2} \right)^{\frac{(\nu-1)}{2}} = \frac{s^{(\nu-1)}(x^2 + \nu)^{(\nu-1)/2}}{2^{(\nu-1)/2}}.$$

Logo, pode-se escrever

$$\begin{aligned} f_X(x;\nu) &= \frac{\nu^{\nu/2} s}{\sqrt{\pi} s(x^2 + \nu)(x^2 + \nu)^{(\nu-1)/2}\Gamma(\frac{\nu}{2})} \int_0^\infty \frac{s^{\nu-1}(x^2 + \nu)^{(\nu-1)/2} e^{-t} s(x^2 + \nu)}{2^{(\nu-1)/2}} ds \\ &= \frac{\nu^{\nu/2}}{\sqrt{\pi}(x^2 + \nu)^{(\nu+1)/2}\Gamma(\frac{\nu}{2})} \int_0^\infty t^{(\nu-1)/2} e^{-t} dt. \end{aligned}$$

Mas, pela definição de função gama tem-se que $\Gamma(\frac{\nu+1}{2}) = \int_0^\infty t^{(\nu-1)/2} e^{-t} dt$.

Assim, a função densidade de X é

$$\begin{aligned}
 f_X(x; \nu) &= \frac{\nu^{\nu/2} \Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{\nu}{2}\right) (x^2 + \nu)^{(\nu+1)/2}} \\
 &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{\nu}{2}\right)} \frac{\nu^{\nu/2}}{(x^2 + \nu)^{(\nu+1)/2}} \frac{\nu^{1/2}}{\nu^{1/2}} \\
 &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\nu^{1/2} \sqrt{\pi} \Gamma\left(\frac{\nu}{2}\right)} \left[\frac{\nu}{(x^2 + \nu)} \right]^{(\nu+1)/2} \\
 &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi} \Gamma\left(\frac{\nu}{2}\right)} \left[1 + \frac{x^2}{\nu} \right]^{-(\nu+1)/2},
 \end{aligned}$$

como se queria mostrar.

Na Figura 11 está a representação gráfica da função de densidade da variável aleatória X . Observa-se que com o aumento dos graus de liberdade ν a variável X se distribui como uma variável aleatória normal padrão.

A média e a variância da distribuição t de *Student* são dadas por

$$\mu = 0 \text{ para } \nu > 1 \text{ e } \sigma^2 = \frac{\nu}{\nu-2} \text{ para } \nu > 2.$$

5.3 Distribuição t não-central univariada

Johnson, Kotz e Balakrishnan (1995) definem a distribuição t não-central univariada com ν graus de liberdade e a constante $\delta \in \mathbb{R}$, denominada de parâmetro de não-centralidade, como a distribuição da variável aleatória X obtida pela razão de duas outras variáveis aleatórias independentes, uma $Z \sim \mathbb{N}(0,1)$ e a outra $U \sim \chi_\nu^2$ com $\nu = n - 1$ graus de liberdade, respectivamente, representada por

$$X = \frac{Z + \delta}{\sqrt{\frac{U}{\nu}}}. \quad (5.3.1)$$

Segundo estes autores, a não-centralidade da distribuição t está associada ao δ do numerador e geralmente a variável com esta distribuição é simbolizada por $t'_\nu(\delta)$.

As variáveis aleatórias $S = \sqrt{U/\nu}$ e Z envolvidas na definição (5.3.1) foram apresentadas nas expressões (5.1.3) e (5.2.1), respectivamente.

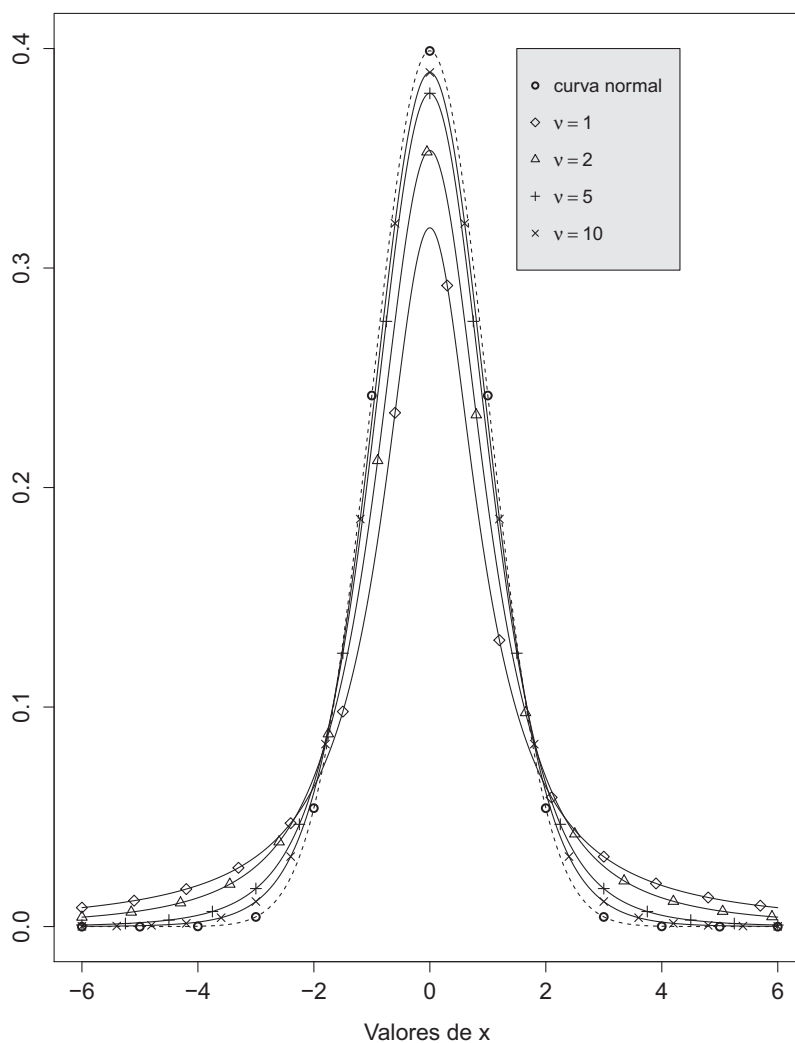


Figura 11 Representação gráfica da densidade da variável aleatória X com distribuição t de Student, dada pela equação (5.2.4), para diferentes valores de graus de liberdade ν

Teorema 5.3.1. (*Função densidade e função de distribuição da t não-central univariada*)

Definindo a variável aleatória X como em (5.3.1), a sua função de densidade é dada por

$$f_X(x; \nu, \delta) = \int_0^\infty s \phi(xs - \delta) f_S(s; \nu) ds \quad (5.3.2)$$

e a sua função de distribuição é dada por

$$F_X(x; \nu, \delta) = \int_0^\infty \Phi(xs - \delta) f_S(s; \nu) ds, \quad (5.3.3)$$

em que $\nu > 0$ são os graus de liberdade relacionados a variável U e $\delta \in \mathbb{R}$ é o parâmetro de não-centralidade para X .

Demonstração:

As funções de densidade da variável normal padrão Z e da variável $S = \sqrt{U/\nu}$ são dadas em (5.2.2) e (5.1.4), respectivamente. A partir da definição da variável aleatória X dada em (5.3.1), relativa à distribuição t não-central, pode-se escrever

$$Z = XS - \delta.$$

Como as variáveis aleatórias Z e U são independentes, as variáveis Z e S também o são e, portanto, a sua distribuição conjunta é dada por

$$f_{Z,S}(z, s; \nu, \delta) = \phi(z) f_S(s; \nu).$$

Usando as transformações de variáveis $Z = XS - \delta$ e $S = S$, cujo jacobiano da transformação é

$$J = \begin{vmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial s} \\ \frac{\partial s}{\partial x} & \frac{\partial s}{\partial s} \end{vmatrix} = \begin{vmatrix} s & x \\ 0 & 1 \end{vmatrix} = s,$$

tem-se que a distribuição conjunta das variáveis S e X é

$$\begin{aligned} f_{S,X}(s, x; \nu, \delta) &= f_{Z,S}(z, s; \nu, \delta) \text{abs}(J) \\ &= \phi(z) f_S(s; \nu) s \end{aligned}$$

Mas, como $z = xs - \delta$ tem-se que

$$f_{S,X}(s, x; \nu, \delta) = s \cdot \phi(xs - \delta) f_S(s; \nu)$$

A função densidade de X é obtida integrando-se esta expressão em relação a s , de 0 a ∞ que é o domínio desta variável, obtendo-se

$$\begin{aligned} f_X(x; \nu, \delta) &= \int_0^\infty f_{S,X}(s, x; \nu, \delta) ds \\ &= \int_0^\infty s \phi(xs - \delta) f_S(s; \nu) ds \end{aligned}$$

que é a função (5.3.2) que se queria mostrar.

Para obter a função de distribuição de X integra-se a função densidade obtida em relação a x no domínio de $-\infty$ a x , chamando $w = x$ e $dw = dx$ no integrando para poder resolver a integral, obtendo-se

$$F_X(x; \nu, \delta) = \int_{-\infty}^x \int_0^\infty \phi(ws - \delta) s ds f_S(s; \nu) dw.$$

Com uma simples mudança na ordem de integração, tem-se

$$F_X(x; \nu, \delta) = \int_0^\infty \int_{-\infty}^x \phi(ws - \delta) s dw f_S(s; \nu) ds.$$

Mas,

$$\int_{-\infty}^x \phi(ws - \delta) s dw = \Phi(xs - \delta).$$

Assim, chega-se a

$$F_X(x; \nu, \delta) = \int_0^\infty \Phi(xs - \delta) f_S(s; \nu) ds.$$

que é a função (5.3.3) que se queria mostrar.

Substituindo as funções dadas em (5.2.2) e (5.1.4) na função de distribuição (5.3.3) obtida, chega-se a mesma função de distribuição da t não-central apresentada por Johnson, Kotz e Balakrishnan (1995), que é

$$F_X(x; \nu, \delta) = \frac{\nu^{(\nu/2)-1}}{2^{(\nu/2)-1} \Gamma(\nu/2)} \int_0^\infty \left\{ s^{\nu-1} e^{-\nu s^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{xs-\delta} e^{-\frac{z^2}{2}} dz \right\} ds.$$

Outras formas de definição da estatística t não-central aparecem na literatura, mas geralmente em função de séries infinitas o que torna difícil o seu manuseio. Por exemplo, Krishnamoorthy (2006) define a distribuição t não-central como a distribuição da variável aleatória X obtida pela razão da variável aleatória $K \sim \mathbb{N}(\delta, 1)$ e por outra $\hat{\sigma}^2 \sim \chi_\nu^2$ com ν graus de liberdade, independentes, representada por $X = K/(\hat{\sigma}/\sqrt{\nu})$, e apresenta a respectiva função de densidade de probabilidade como :

$$f(x; \nu, \delta) = \frac{\nu^{\frac{\nu}{2}} \exp(-\frac{\delta^2}{2})}{\sqrt{\pi} \Gamma(\frac{\nu}{2}) \left(1 + \frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}} \sum_{i=0}^{\infty} \frac{\Gamma(\frac{\nu+1+i}{2})}{i!} \left(\frac{x \delta \sqrt{2}}{\sqrt{\nu + x^2}} \right)^i,$$

com $-\infty < x < \infty$, $-\infty < \delta < \infty$ e com $\left(\frac{x\delta\sqrt{2}}{\sqrt{\nu+x^2}}\right)^0 = 1$ para todos os valores de x e δ , incluindo o zero.

Em ambas as definições, de Johnson, Kotz e Balakrishnan (1995) e Krishnamoorthy (2006), a média e a variância da t não-central são dadas pelas expressões:

$$\mu_{t'_\nu(\delta)} = \frac{\Gamma(\frac{\nu-1}{2})\sqrt{\frac{\nu}{2}}}{\Gamma(\frac{\nu}{2})}\delta \quad \text{e} \quad \text{var}(t'_\nu(\delta)) = \frac{\nu}{\nu-2}(1+\delta^2) - \left[\frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})}\right]^2 \frac{\nu}{2}\delta^2.$$

Algumas propriedades da $t'_\nu(\delta)$ apresentadas por esses autores são:

1. A distribuição não-central $t'_\nu(\delta)$ especializa-se na distribuição t com ν graus de liberdade quando $\delta = 0$;
2. $P[t'_\nu(\delta) \leq 0] = P[Z \leq -\delta] = \Phi(-\delta)$, em que Z é uma variável aleatória normal padrão, e os eventos $[t'_\nu(\delta) \leq 0]$ e $[Z + \delta < 0]$ são idênticos;
3. $P[t'_\nu(\delta) \leq t] = P[t'_\nu(-\delta) \geq -t] = 1 - P[t'_\nu(-\delta) \leq -t]$;

5.3.1 Aplicações da distribuição t não-central

A estatística

$$t = \frac{\bar{Y} - \mu}{\frac{\hat{\sigma}}{\sqrt{n}}},$$

com $\bar{Y}$ e $\hat{\sigma}$ calculados a partir de uma amostra aleatória de tamanho n , tem distribuição t de *Student* com $n - 1$ graus de liberdade e é usada para testar a hipótese de que a média de uma população é μ .

Entretanto, se a média populacional da qual a amostra foi retirada não é μ e sim μ' , tem-se $\delta = \mu - \mu'$, e a estatística t é dada por

$$t'_{n-1} = \sqrt{n} \left[\frac{\delta}{\sigma} \right],$$

com σ o desvio padrão populacional. O poder do teste é calculado usando a função densidade desta distribuição t não-central univariada.

De modo semelhante, a estatística usada para testar a igualdade de médias de duas populações normais μ_A e μ_B , com variância comum σ^2 mas desconhecida, a partir de amostras de tamanho n_A e n_B é dada por

$$t_{(\bar{X}_A - \bar{X}_B)} = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{(n_A-1)S_A^2 + (n_B-1)S_B^2}{n_A + n_B - 2} \left[\frac{1}{n_A} + \frac{1}{n_B} \right]}}.$$

Se as duas média populacional são iguais, a estatística $t_{(\bar{X}_A - \bar{X}_B)}$ aproxima-se da distribuição t de *Student*. Mas se $\mu_A - \mu_B = \delta$, a estatística t é dada por

$$t'_{n_A+n_B-2} = \sqrt{\frac{1}{n_A} + \frac{1}{n_B}} \left[\frac{\delta}{\sigma} \right].$$

Neste caso também pode ser calculado o poder do teste pela distribuição t não-central apropriada.

5.4 Distribuições multivariadas

Denomina-se distribuição multivariada a distribuição de um vetor aleatório $\mathbf{X} = [X_1, X_2, \dots, X_r]^\top$, em que $r (\geq 2)$ representa o número de variáveis e cujos elementos X_i são variáveis aleatórias univariadas com função de distribuição $F_{X_i}(x_i)$. A função definida por:

$$F_{\mathbf{X}}(\mathbf{x}) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_r \leq x_r)$$

representa a função de distribuição conjunta de $X_1, X_2, \dots, X_r$.

Se $F_{\mathbf{X}}(\mathbf{x})$ é contínua em $-\infty < x_i < \infty$ para todo $i = 1, \dots, r$, então existe a função

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\partial^r F_{\mathbf{X}}(\mathbf{x})}{\partial x_1 \partial x_2 \dots \partial x_r}$$

denominada de função densidade de probabilidade do vetor aleatório contínuo $\mathbf{X}$.

Se os componentes $X_1, X_2, \dots, X_r$ possuem independência estatística, então

$$F_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^r F_{X_i}(x_i)$$

ou, equivalentemente

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^r f_{X_i}(x_i).$$

Tomando uma partição de $\mathbf{X}$ dada por

$$\begin{aligned} \mathbf{X} &= [X_1, X_2, \dots, X_q | X_{q+1}, X_{q+2}, \dots, X_r]^\top \\ &= [X_{(1)}^\top | X_{(2)}^\top]^\top, \end{aligned}$$

pode-se obter a função densidade marginal do subconjunto $\mathbf{X}_{(1)}$ integrando $f_{\mathbf{X}}(\mathbf{x})$ sobre o domínio das variáveis do subconjunto $\mathbf{X}_{(2)}$, para $q < r$ da seguinte forma

$$f_{\mathbf{X}_{(1)}}(\mathbf{x}_{(1)}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x}) dx_{q+1} \dots dx_r.$$

Pode-se também conhecer a probabilidade condicional de um evento quando se conhece a ocorrência de outro pela função densidade de probabilidade de $\mathbf{X}_{(1)}$ quando $\mathbf{X}_{(2)} = \mathbf{x}_{(2)}$ é conhecido, dada pela expressão

$$f_{\mathbf{X}_{(1)}}(\mathbf{x}_{(1)}|\mathbf{X}_{(2)} = \mathbf{x}_{(2)}) = \frac{f_{\mathbf{X}}(\mathbf{x})}{f_{\mathbf{X}_{(2)}}(\mathbf{x}_{(2)})},$$

considerando que $f_{\mathbf{X}_{(2)}}(\mathbf{x}_{(2)}) \neq 0$.

5.4.1 Distribuições elípticas simétricas

Segundo Ferreira (2008) as distribuições multivariadas pertencentes a família de distribuições elípticas simétricas possuem a característica de apresentarem função de densidade com contornos de formato elíptico.

Considerando um vetor aleatório $\mathbf{X} = (X_1, X_2, \dots, X_r)^\top \in \mathbb{R}^r$, com média $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_r]^\top \in \mathbb{R}^r$ e matriz de covariâncias que é função da matriz simétrica e positiva definida $\boldsymbol{\Sigma}$ ($r \times r$), pode-se definir as distribuições elípticas simétricas como funções da forma quadrática $(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})$ dadas por

$$f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = |\boldsymbol{\Sigma}|^{-\frac{1}{2}} g[(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})]$$

em que g é uma função com domínio em $[0, \infty)$ que para todo $\mathbf{z} \in \mathbb{R}^r$ satisfaz a condição $\int_{\mathbb{R}^r} g(\mathbf{z}^\top \mathbf{z}) d\mathbf{z} = 1$, com $\mathbf{z} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\mu})$.

Se for representado uma distribuição elíptica por $E_r(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$, em que a função g pode ou não ser omitida dependendo do contexto particular, a distribuição normal multivariada $\mathbb{N}_r(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ é dada por

$$f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-\frac{r}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right\}, \quad (5.4.1)$$

para $\mathbf{x} \in \mathbb{R}^r$ e $g(k) = (2\pi)^{-\frac{r}{2}} \exp \left\{ -\frac{k}{2} \right\}$, com $k = (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \geq 0$. Se $\boldsymbol{\mu} = \mathbf{0}$ e $\boldsymbol{\Sigma} = \mathbf{I}$, tem-se a distribuição normal multivariada esférica simétrica.

5.4.2 Distribuição t de Student multivariada

A distribuição t de *Student* multivariada com ν graus de liberdade pertence a família de distribuições elípticas. A função densidade de probabilidade de um vetor aleatório $\mathbf{X}$ com distribuição t multivariada é dada por

$$f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = \frac{\Gamma(\frac{\nu+r}{2})}{(\pi\nu)^{\frac{r}{2}} \Gamma(\frac{\nu}{2}) |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \left[1 + \frac{1}{\nu} (\mathbf{x} - \boldsymbol{\mu})^{\top} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right]^{-\frac{\nu+r}{2}}, \quad (5.4.2)$$

em que g é dada por

$$g(k; \nu) = \frac{\Gamma(\frac{\nu+r}{2})}{(\pi\nu)^{\frac{r}{2}} \Gamma(\frac{\nu}{2})} \left(1 + \frac{k}{\nu} \right)^{-\frac{\nu+r}{2}},$$

com $k = (\mathbf{x} - \boldsymbol{\mu})^{\top} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \geq 0$. A média da variável aleatória $\mathbf{X}$ é o vetor dado por $E(\mathbf{X}) = \boldsymbol{\mu}$ e a matriz de covariâncias é dada por $Cov(\mathbf{X}) = \nu \boldsymbol{\Sigma} / (\nu - 2)$, no caso de $\nu > 2$.

Partindo da expressão (5.4.2), se for efetuada uma transformação dada por $\mathbf{Z} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{X} - \boldsymbol{\mu})$ obtém-se a distribuição t de *Student* multivariada esférica simétrica²⁷, também denominada de distribuição t de *Student* padrão multivariada, cuja densidade é dada por

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{\Gamma(\frac{\nu+r}{2})}{(\pi\nu)^{\frac{r}{2}} \Gamma(\frac{\nu}{2})} \left(1 + \frac{1}{\nu} \mathbf{z}^{\top} \mathbf{z} \right)^{-\frac{\nu+r}{2}}.$$

A variável aleatória $\mathbf{Z}$ terá vetor de médias nulo, matriz de covariâncias $\nu \mathbf{I} / (\nu - 2)$ para $\nu > 2$, e a função densidade terá contornos de densidades constantes e esféricos.

²⁷ Esta distribuição também pode ser obtida a partir da expressão (5.4.2) usando a transformação $Z_i = X_i / (S / \sqrt{\nu})$, para cada variável X_i com $i = 1, \dots, r$ do vetor aleatório $\mathbf{X}$, sendo que $X \sim N_r(0, \mathbf{I})$, $S^2 \sim \chi_{\nu}^2$, e $\mathbf{X}$ e S são independentes (MARDIA; KENT; BIBBY, 2003).

Teorema 5.4.1. Dado o vetor aleatório $\mathbf{X} = (X_1, X_2, \dots, X_r)^\top$ com função densidade conjunta dada por (5.4.2), $\mathbf{X} \sim t_r(\nu, \boldsymbol{\mu}, \boldsymbol{\Sigma})$. Se for efetuada a transformação $\mathbf{Y} = \mathbf{V}^{-1/2}\mathbf{X}$, em que $\mathbf{V} = \text{diag}(\boldsymbol{\Sigma})$, a função densidade de probabilidade do vetor aleatório $\mathbf{Y} = (Y_1, Y_2, \dots, Y_r)^\top \in \mathbb{R}^r$ é dada por

$$f_{\mathbf{Y}}(\mathbf{y}; \boldsymbol{\mu}^*, \boldsymbol{\rho}, \nu) = \frac{\Gamma(\frac{\nu+r}{2})}{(\pi\nu)^{(\frac{r}{2})}\Gamma(\frac{\nu}{2})|\boldsymbol{\rho}|^{\frac{1}{2}}} \left[1 + \frac{1}{\nu}(\mathbf{y} - \boldsymbol{\mu}^*)^\top \boldsymbol{\rho}^{-1}(\mathbf{y} - \boldsymbol{\mu}^*) \right]^{-\frac{\nu+r}{2}}, \quad (5.4.3)$$

ou seja, $\mathbf{Y} \sim t_r(\nu, \boldsymbol{\mu}^*, \boldsymbol{\rho})$, com ν graus de liberdade, vetor de médias $\boldsymbol{\mu}^* = [\mu_1^*, \mu_2^*, \dots, \mu_r^*]^\top \in \mathbb{R}^r$ e matriz de correlação $\boldsymbol{\rho}$.

Demonstração:

Usando a transformação $\mathbf{Y} = \mathbf{V}^{-1/2}\mathbf{X}$, toma-se o vetor aleatório $\mathbf{Y}$ como a padronização do vetor aleatório $\mathbf{X}$, em que $\mathbf{V}$ é uma matriz diagonal definida como $\mathbf{V} = \text{diag}(\boldsymbol{\Sigma}) = \text{diag}(\sigma_{ii})$, ou seja, uma matriz cuja diagonal contém as variâncias da p -variáveis X_i . Tem-se que $\mathbf{V}^{-1/2} = \text{diag}(1/\sqrt{\sigma_{ii}})$. A esperança do vetor aleatório $\mathbf{Y}$ é o vetor coluna dado por

$$E(\mathbf{Y}) = E(\mathbf{V}^{-1/2}\mathbf{X}) = \mathbf{V}^{-1/2}E(\mathbf{X}) = \mathbf{V}^{-1/2}\boldsymbol{\mu} = \frac{\mu_i}{\sqrt{\sigma_{ii}}} = \boldsymbol{\mu}^*$$

e a matriz de covariâncias de $\mathbf{Y}$ é dada por

$$\begin{aligned} \text{Cov}(\mathbf{Y}) &= \text{Cov}(\mathbf{V}^{-1/2}\mathbf{X}) \\ &= \mathbf{V}^{-1/2}\text{Cov}(\mathbf{X})\mathbf{V}^{-1/2} \\ &= \mathbf{V}^{-1/2}\nu/(\nu-2)\boldsymbol{\Sigma}\mathbf{V}^{-1/2} \\ &= \nu/(\nu-2)\mathbf{V}^{-1/2}\boldsymbol{\Sigma}\mathbf{V}^{-1/2} \\ &= \frac{\nu\boldsymbol{\rho}}{(\nu-2)}, \end{aligned}$$

pois a matriz de correlações é definida por $\boldsymbol{\rho} = \mathbf{V}^{-1/2}\boldsymbol{\Sigma}\mathbf{V}^{-1/2}$.

Para determinar a função densidade de probabilidade conjunta do vetor aleatório $\mathbf{Y}$ a partir da função densidade de $\mathbf{X}$ dada por (5.4.2), usa-se o jacobiano da transformação que é

$$\mathbf{J} = \left| \frac{d\mathbf{x}}{d\mathbf{y}} \right| = |\mathbf{V}^{1/2}| = |\mathbf{V}|^{1/2} = \prod_{i=1}^p \sqrt{\sigma_{ii}}$$

e tem-se que $f_Y(\mathbf{y}) = f_X(\mathbf{x})\text{abs}(\mathbf{J}) = f_X(\mathbf{x})\text{abs}(|\mathbf{V}|)^{1/2}$, com $\mathbf{x} = \mathbf{V}^{1/2}\mathbf{y}$. Assim

$$\begin{aligned} f_Y(\mathbf{y}) &= \frac{|\mathbf{V}|^{\frac{1}{2}} \Gamma(\frac{\nu+r}{2})}{(\pi\nu)^{(\frac{r}{2})} \Gamma(\frac{\nu}{2}) |\Sigma|^{\frac{1}{2}}} \left[1 + \frac{1}{\nu} (\mathbf{V}^{1/2}\mathbf{y} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{V}^{1/2}\mathbf{y} - \boldsymbol{\mu}) \right]^{-\frac{\nu+r}{2}} \\ &= \frac{\Gamma(\frac{\nu+r}{2})}{(\pi\nu)^{(\frac{r}{2})} \Gamma(\frac{\nu}{2}) \frac{|\Sigma|^{\frac{1}{2}}}{|\mathbf{V}|^{\frac{1}{2}}}} \left[1 + \frac{1}{\nu} (\mathbf{y} - \mathbf{V}^{-1/2}\boldsymbol{\mu})^\top [\mathbf{V}^{\frac{1}{2}} \Sigma^{-1} \mathbf{V}^{\frac{1}{2}}] (\mathbf{y} - \mathbf{V}^{-1/2}\boldsymbol{\mu}) \right]^{-\frac{\nu+r}{2}} \\ &= \frac{\Gamma(\frac{\nu+r}{2})}{(\pi\nu)^{(\frac{r}{2})} \Gamma(\frac{\nu}{2}) \frac{|\Sigma|^{\frac{1}{2}}}{|\mathbf{V}|^{\frac{1}{2}}}} \left[1 + \frac{1}{\nu} (\mathbf{y} - \boldsymbol{\mu}^*)^\top \boldsymbol{\rho}^{-1} (\mathbf{y} - \boldsymbol{\mu}^*) \right]^{-\frac{\nu+r}{2}}. \end{aligned}$$

Como

$$\begin{aligned} |\Sigma| &= |\mathbf{V}^{1/2} \boldsymbol{\rho} \mathbf{V}^{1/2}| = |\mathbf{V}^{1/2}| |\boldsymbol{\rho}| |\mathbf{V}^{1/2}| = |\mathbf{V}|^{1/2} |\boldsymbol{\rho}| |\mathbf{V}|^{1/2} \\ &= |\mathbf{V}| |\boldsymbol{\rho}| \end{aligned}$$

tem-se que,

$$\begin{aligned} |\Sigma|^{1/2} &= |\mathbf{V}|^{1/2} |\boldsymbol{\rho}|^{1/2} \\ \Rightarrow \frac{|\Sigma|^{1/2}}{|\mathbf{V}|^{1/2}} &= |\boldsymbol{\rho}|^{1/2}, \end{aligned}$$

ou seja,

$$f_Y(\mathbf{y}; \boldsymbol{\mu}^*, \boldsymbol{\rho}, \nu) = \frac{\Gamma(\frac{\nu+r}{2})}{(\pi\nu)^{(\frac{r}{2})} \Gamma(\frac{\nu}{2}) |\boldsymbol{\rho}|^{\frac{1}{2}}} \left[1 + \frac{1}{\nu} (\mathbf{y} - \boldsymbol{\mu}^*)^\top \boldsymbol{\rho}^{-1} (\mathbf{y} - \boldsymbol{\mu}^*) \right]^{-\frac{\nu+r}{2}},$$

como se queria mostrar.

A definição (5.4.3) é dada por Kotz e Nadarajah (2004) para a distribuição t de *Student* multivariada.

Um caso particular da distribuição t de *Student* multivariada é quando $\nu = 1$, em que a distribuição é denominada de *Cauchy* multivariada (MARDIA; KENT; BIBBY, 2003). Segundo Kotz e Nadarajah (2004), o parâmetro ν (graus de liberdade) é denominado parâmetro de forma, porque o grau de achatamento

da curva (curtose) pode ser diminuído, conservado ou aumentado, variando v . A forma limite da equação (5.4.3) quando $v \rightarrow \infty$ é a função distribuição de probabilidade da normal multivariada com vetor de médias μ e matriz de covariância Σ , ou seja, a distribuição t multivariada para infinitos graus de liberdade tem distribuição normal multivariada como limite.

5.4.3 Distribuição do máximo e do máximo do módulo da normal multivariada

Para obter a distribuição das estatísticas do teste de Dunnett é necessário obter primeiro a distribuição $\max(\mathbf{Z}) = \max_i(Z_i)$, que é utilizada para os testes unilaterais e a distribuição do $\max(|\mathbf{Z}|) = \max_i(|Z_i|)$ que é utilizada para os testes bilaterais.

A função de distribuição acumulada de um vetor aleatório normal padrão multivariado $\mathbf{Z}$ é definida em termos da função de distribuição de probabilidade conjunta das r variáveis aleatórias dada por

$$\begin{aligned}\Phi_{Z_1, Z_2, \dots, Z_r}(z_1, z_2, \dots, z_r) &= P(\mathbf{Z} < \mathbf{z}) = P(Z_1 < z_1, Z_2 < z_2, \dots, Z_r < z_r) \\ &= \int_{-\infty}^{z_1} \dots \int_{-\infty}^{z_r} \phi_{Z_1, \dots, Z_r}(u_1, \dots, u_r) du_1 \dots du_r,\end{aligned}$$

em que $\phi_{\mathbf{Z}}(\mathbf{z}; \mathbf{0}, \boldsymbol{\rho}) = (2\pi)^{-\frac{r}{2}} |\boldsymbol{\rho}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{z}^\top \boldsymbol{\rho}^{-1} \mathbf{z}) \right\}$.

Do mesmo modo, a distribuição do $\max(\mathbf{Z})$ e, de forma similar a do $\max(|\mathbf{Z}|)$, recaem em integrais de r -dimensões que podem ser avaliadas numericamente por uma quadratura r -dimensional.

Essas quadraturas podem ser simplificadas se os Z_i 's puderem ser expressos como combinações lineares de variáveis aleatórias normais padrões independentes. A escolha dessas combinações lineares deve recair em um fator obtido de $\boldsymbol{\rho}_{r \times r}$. O teorema a seguir mostra uma combinação linear em que foi reduzida a avaliação numérica de r integrais necessárias para o cálculo da $\Phi(\mathbf{z}; \mathbf{0}, \boldsymbol{\rho})$ para uma integral univariada.

Teorema 5.4.2. (Função de distribuição do máximo da normal multivariada)

Dado o vetor aleatório $\mathbf{Z} = (Z_1, Z_2, \dots, Z_r)^\top$ com distribuição normal padrão multivariada cuja função densidade da distribuição conjunta das r variáveis é dada por $\phi(\mathbf{z}; \mathbf{0}, \boldsymbol{\rho}) = (2\pi)^{-\frac{r}{2}} |\boldsymbol{\rho}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{z}^\top \boldsymbol{\rho}^{-1} \mathbf{z}) \right\}$.

Se a matriz $\boldsymbol{\rho}$ satisfaz a condição de estrutura de correlação produto dada por $\rho_{ij} = \lambda_i \lambda_j$ para $1 \leq i \neq j \leq r$, em que λ_i e $\lambda_j \in (-1, 1)$, representada por

$$\boldsymbol{\rho} = \begin{pmatrix} 1 & \lambda_1 \lambda_2 & \lambda_1 \lambda_3 & \dots & \lambda_1 \lambda_r \\ \lambda_2 \lambda_1 & 1 & \lambda_2 \lambda_3 & \dots & \lambda_2 \lambda_r \\ \lambda_3 \lambda_1 & \lambda_3 \lambda_2 & 1 & \dots & \lambda_3 \lambda_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_r \lambda_1 & \lambda_r \lambda_2 & \lambda_r \lambda_3 & \dots & 1 \end{pmatrix}, \quad (5.4.4)$$

tomando-se outro vetor aleatório $\mathbf{Y} = (Y_1, Y_2, \dots, Y_r, Y_0)^\top$, com distribuição normal padrão multivariada, e a matriz de combinações lineares dada por

$$\mathbf{\Lambda} = \begin{pmatrix} \sqrt{1 - \lambda_1^2} & 0 & 0 & \dots & 0 & -\lambda_1 \\ 0 & \sqrt{1 - \lambda_2^2} & 0 & \dots & 0 & -\lambda_2 \\ 0 & 0 & \sqrt{1 - \lambda_3^2} & \dots & 0 & -\lambda_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \sqrt{1 - \lambda_p^2} & -\lambda_p \end{pmatrix}, \quad (5.4.5)$$

então pode-se escrever o vetor $\mathbf{Z} = \mathbf{\Lambda} \mathbf{Y}$, cujo máximo dado por $Z = \max_i Z_i$ tem a seguinte função de distribuição de probabilidade:

$$F_Z(z; \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \left[\prod_{i=1}^p \Phi \left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}} \right) \right] \phi(y) dy, \quad (5.4.6)$$

em que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_p)$.

Demonstração:

Sendo o vetor aleatório $\mathbf{Z} \sim \mathbb{N}_p(\mathbf{0}, \boldsymbol{\rho})$ com a matriz $\boldsymbol{\rho}$ satisfazendo a condição de estrutura de correlação produto dada por (5.4.4), toma-se o vetor aleatório $\mathbf{Y} = (Y_1, Y_2, \dots, Y_r, Y_0)^\top$, também normal padrão multivariado, e a matriz de combinações lineares dada por (5.4.5) e escreve-se o vetor $\mathbf{Z}$ como

$$\mathbf{Z} = \mathbf{\Lambda Y} \Rightarrow \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_r \end{pmatrix} = \begin{pmatrix} \sqrt{1 - \lambda_1^2} Y_1 - \lambda_1 Y_0 \\ \sqrt{1 - \lambda_2^2} Y_2 - \lambda_2 Y_0 \\ \vdots \\ \sqrt{1 - \lambda_r^2} Y_r - \lambda_r Y_0 \end{pmatrix}.$$

de modo que $E(\mathbf{Z}) = E(\mathbf{\Lambda Y}) = \mathbf{\Lambda E(Y)} = \mathbf{\Lambda 0} = \mathbf{0}$ e

$$\text{Cov}(\mathbf{Z}) = \text{Cov}(\mathbf{\Lambda Y}) = \mathbf{\Lambda Cov(Y) \Lambda^\top} = \mathbf{\Lambda \mathbf{\Lambda}^\top} = \mathbf{\Lambda \Lambda}^\top$$

ou seja,

$$\begin{aligned} \text{Cov}(\mathbf{Z}) &= \begin{pmatrix} [\sqrt{1 - \lambda_1^2}]^2 + \lambda_1^2 & \lambda_1 \lambda_2 & \dots & \lambda_1 \lambda_r \\ \lambda_2 \lambda_1 & [\sqrt{1 - \lambda_2^2}]^2 + \lambda_2^2 & \dots & \lambda_2 \lambda_r \\ \lambda_3 \lambda_1 & \lambda_3 \lambda_2 & \dots & \lambda_3 \lambda_r \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_r \lambda_1 & \lambda_r \lambda_2 & \dots & [\sqrt{1 - \lambda_r^2}]^2 + \lambda_r^2 \end{pmatrix} \\ \text{Cov}(\mathbf{Z}) &= \begin{pmatrix} 1 & \lambda_1 \lambda_2 & \lambda_1 \lambda_3 & \dots & \lambda_1 \lambda_r \\ \lambda_2 \lambda_1 & 1 & \lambda_2 \lambda_3 & \dots & \lambda_2 \lambda_r \\ \lambda_3 \lambda_1 & \lambda_3 \lambda_2 & 1 & \dots & \lambda_3 \lambda_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_r \lambda_1 & \lambda_r \lambda_2 & \lambda_r \lambda_3 & \dots & 1 \end{pmatrix} = \boldsymbol{\rho}. \end{aligned}$$

Assim, de forma geral tem-se:

$$Z_i = \sqrt{1 - \lambda_i^2} Y_i - \lambda_i Y_0 \quad 1 \leq i \leq r, \quad \text{com} \quad \lambda_i^2 = \rho_i.$$

Considerando a afirmação probabilística

$$\begin{aligned} \Phi_{Z_i}(z; \lambda_i) &= P(Z_i \leq z) = P\left(\sqrt{1 - \lambda_i^2} Y_i - \lambda_i Y_0 \leq z\right) \\ &= P\left(Y_i \leq \frac{\lambda_i Y_0 + z}{\sqrt{1 - \lambda_i^2}}\right), \end{aligned}$$

cada $Y_i \sim \mathbb{N}(0,1)$, mas a acumulada $\Phi_{Z_i}(z)$ depende de Y_0 , que é uma variável aleatória. Condiçãoando ao conhecimento do valor de Y_0 que será denotado por y tem-se,

$$\Phi_{Z_i}(z|Y_0 = y) = P\left(Y_i \leq \frac{\lambda_i Y_0 + z}{\sqrt{1 - \lambda_i^2}} \middle| Y_0 = y\right) = P\left(Y_i \leq \frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}}\right),$$

logo,

$$\begin{aligned}\Phi_{Z_i}(z; \lambda_i) &= \int_{-\infty}^{\infty} \Phi(z|Y_0 = y) \phi(y) dy \\ \Phi_{Z_i}(z; \lambda_i) &= \int_{-\infty}^{\infty} \Phi\left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}}\right) \phi(y) dy.\end{aligned}$$

Usando o resultado da distribuição do máximo de variáveis aleatórias independentes Y_i 's tem-se

$$P(Z_1 \leq z, Z_2 \leq z, \dots, Z_r \leq z) = P\left(\max_i Z_i \leq z\right) = \prod_{i=1}^r \Phi\left(\frac{\lambda_i Y_0 + z}{\sqrt{1 - \lambda_i^2}} \middle| Y_0 = y\right),$$

e com $Z = \max_i Z_i$ tem-se que função de distribuição do máximo da normal multivariada é dada por

$$F_Z(z; \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \left[\prod_{i=1}^r \Phi\left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}}\right) \right] \phi(y) dy,$$

como se queria mostrar.

A função de distribuição do máximo da normal multivariada é utilizada no teste de Dunnett unilateral quando os graus de liberdade crescem tendendo para o infinito.

Teorema 5.4.3. (Função densidade do máximo da normal multivariada)

Dadas as mesmas condições do teorema 5.3.2., o máximo do vetor $\mathbf{Z} = \boldsymbol{\Lambda}\mathbf{Y}$ dado por $Z = \max_i Z_i$ tem a seguinte função densidade de probabilidade:

$$f_Z(z; \lambda) = \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}} \right) \frac{1}{\sqrt{1 - \lambda_i^2}} \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k y + z}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) dy. \quad (5.4.7)$$

em que $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_r)$

Demonstração:

A função densidade de probabilidade de $Z = \max_i(Z_i)$ é obtida derivando-se a função de distribuição dada em (5.4.6) em relação a z , ou seja

$$f_Z(z; \lambda) = \frac{\partial}{\partial z} \left\{ \int_{-\infty}^{\infty} \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}} \right) \right] \phi(y) dy \right\}.$$

A derivada do produto de funções é dada por:

$$\begin{aligned} \frac{\partial [f_1(q) \overbrace{f_2(q) \dots f_r(q)}^{Q_{-1}}]}{\partial q} &= f_1'(q) \cdot Q_{-1}(q) + f_1(q) Q'_{-1} \\ &= f_1'(q) \cdot Q_{-1}(q) + f_1(q) [f_2'(q) \cdot Q_{-1,-2}(q) + f_2(q) Q'_{-1,-2}(q)], \end{aligned}$$

com $Q_{-1,-2}(q) = f_3(q) \dots f_r(q)$

$$\frac{\partial [f_1(q) f_2(q) \dots f_r(q)]}{\partial q} = f_1'(q) \cdot Q_{-1}(q) + f_2'(q) \cdot Q_{-2}(q) + f_1(q) f_2(q) Q'_{-1,-2}(q),$$

com $Q_{-2}(q) = f_1(q) f_3(q) \dots f_r(q)$

$$\begin{aligned} \frac{\partial [f_1(q) f_2(q) \dots f_r(q)]}{\partial q} &= f_1'(q) \cdot Q_{-1}(q) + f_2'(q) \cdot Q_{-2}(q) + f_1(q) f_2(q) [f_3'(q) Q_{-1,-2,-3}(q) \\ &+ f_3(q) Q'_{-1,-2,-3}(q)], \text{ com } Q_{-1,-2,-3}(q) = f_4(q) \dots f_r(q) \end{aligned}$$

$$\begin{aligned} \frac{\partial [f_1(q) f_2(q) \dots f_r(q)]}{\partial q} &= f_1'(q) \cdot Q_{-1}(q) + f_2'(q) \cdot Q_{-2}(q) + f_3'(q) \cdot Q_{-3}(q) + f_1(q) f_2(q) \times \\ &\times f_3(q) Q'_{-1,-2,-3}(q), \text{ com } Q_{-3}(q) = f_1(q), f_2(q), f_4(q), \dots, f_r(q) \end{aligned}$$

e, de forma generalizada

$$\frac{\partial[f_1(q)f_2(q)\dots f_r(q)]}{\partial q} = \sum_{i=1}^r f'_i(q) \prod_{\substack{k=1 \\ k \neq i}}^r f_k(q). \quad (5.4.8)$$

Assim, a função densidade de probabilidade de $Z = \max_i Z_i$ é

$$f_Z(z; \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \phi\left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}}\right) \frac{1}{\sqrt{1 - \lambda_i^2}} \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi\left(\frac{\lambda_k y + z}{\sqrt{1 - \lambda_k^2}}\right) \right] \right\} \phi(y) dy,$$

como se queria mostrar.

Teorema 5.4.4. *(Função de distribuição do máximo do módulo da normal multivariada)*

Dadas as mesmas condições do teorema 5.3.2., o máximo do módulo do vetor $\mathbf{Z} = \boldsymbol{\Lambda}\mathbf{Y}$ dado por $|Z| = \max_i |Z_i|$ tem a seguinte função de distribuição de probabilidade:

$$F_{|Z|}(z; \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \left\{ \prod_{i=1}^r \left[\Phi\left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}}\right) - \Phi\left(\frac{\lambda_i y - z}{\sqrt{1 - \lambda_i^2}}\right) \right] \right\} \phi(y) dy, \quad (5.4.9)$$

em que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_r)$.

Demonstração:

De modo similar ao demonstrado no teorema 5.3.2, porém usando o resultado da distribuição do máximo do módulo de uma variável aleatória em que

$$\begin{aligned} P(|Z_1| \leq z, |Z_2| \leq z, \dots, |Z_r| \leq z) &= P\left(\max_i |Z_i| \leq z\right) = P(-z \leq \max_i Z_i \leq z) \\ &= P\left(\frac{\lambda_i Y_0 - z}{\sqrt{1 - \lambda_i^2}} \leq Y_i \leq \frac{\lambda_i Y_0 + z}{\sqrt{1 - \lambda_i^2}}\right), \end{aligned}$$

tem-se

$$\begin{aligned} P\left(\max_i |Z_i| \leq z\right) &= \prod_{i=1}^r P\left(\frac{\lambda_i Y_0 - z}{\sqrt{1 - \lambda_i^2}} \leq Y_i \leq \frac{\lambda_i Y_0 + z}{\sqrt{1 - \lambda_i^2}}\right) \\ &= \prod_{i=1}^r \int_{-\infty}^{\infty} \left\{ \Phi\left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}}\right) - \Phi\left(\frac{\lambda_i y - z}{\sqrt{1 - \lambda_i^2}}\right) \right\} \phi(y) dy \end{aligned}$$

e com $|Z| = \max_i |Z_i|$ tem-se que a função de distribuição do máximo do módulo da normal multivariada é dada por

$$F_{|Z|}(z; \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \left\{ \prod_{i=1}^r \left[\Phi\left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}}\right) - \Phi\left(\frac{\lambda_i y - z}{\sqrt{1 - \lambda_i^2}}\right) \right] \right\} \phi(y) dy,$$

como se queria mostrar.

A função de distribuição do máximo do módulo da normal multivariada é utilizada no teste de Dunnett bilateral quando os graus de liberdade crescem tendendo para o infinito.

Teorema 5.4.5. (*Função densidade do máximo do módulo da normal multivariada*)

Dadas as mesmas condições do teorema 5.3.2., o máximo do vetor $\mathbf{Z} = \boldsymbol{\Lambda}\mathbf{Y}$ dado por $|Z| = \max_i (|Z_i|)$ tem a seguinte função densidade de probabilidade:

$$\begin{aligned} f_{|Z|}(z; \boldsymbol{\lambda}) &= \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi\left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}}\right) + \phi\left(\frac{\lambda_i y - z}{\sqrt{1 - \lambda_i^2}}\right) \right] \left(\frac{1}{\sqrt{1 - \lambda_i^2}} \right) \right) \right. \\ &\quad \times \left. \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi\left(\frac{\lambda_k y + z}{\sqrt{1 - \lambda_k^2}}\right) - \Phi\left(\frac{\lambda_k y - z}{\sqrt{1 - \lambda_k^2}}\right) \right] \right\} \phi(y) dy. \end{aligned} \quad (5.4.10)$$

em que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_r)$.

Demonstração:

A função densidade de probabilidade de $|Z| = \max_i (|Z_i|)$ é obtida derivando-se (5.4.9) em relação a z , conforme (5.4.8) obtendo-se

$$\begin{aligned}
f_{|Z|}(z; \boldsymbol{\lambda}) &= \frac{\partial}{\partial z} \left\{ \int_{-\infty}^{\infty} \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - z}{\sqrt{1 - \lambda_i^2}} \right) \right] \phi(y) dy \right\}, \\
&= \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}} \right) \frac{1}{\sqrt{1 - \lambda_i^2}} \right] - \left[\phi \left(\frac{\lambda_i y - z}{\sqrt{1 - \lambda_i^2}} \right) \frac{-1}{\sqrt{1 - \lambda_i^2}} \right] \right) \right. \\
&\quad \times \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + z}{\sqrt{1 - \lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - z}{\sqrt{1 - \lambda_k^2}} \right) \right] \left. \right\} \phi(y) dy, \\
&= \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}} \right) + \phi \left(\frac{\lambda_i y - z}{\sqrt{1 - \lambda_i^2}} \right) \right] \left(\frac{1}{\sqrt{1 - \lambda_i^2}} \right) \right) \right. \\
&\quad \times \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + z}{\sqrt{1 - \lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - z}{\sqrt{1 - \lambda_k^2}} \right) \right] \left. \right\} \phi(y) dy,
\end{aligned}$$

como se queria mostrar.

5.4.4 Distribuição do máximo e do máximo do módulo da t multivariada

Para se obter a distribuição do máximo da t multivariada, $\max(\mathbf{T}) = \max_i(T_i)$, sendo $\mathbf{T} = (T_1, T_2, \dots, T_r)^\top$ um vetor aleatório definido por

$$\mathbf{T} = \frac{\mathbf{Z}}{\sqrt{\frac{U}{\nu}}}, \quad (5.4.11)$$

em que $\mathbf{Z} \sim \mathbb{N}_p(\mathbf{0}, \boldsymbol{\rho})$ e $S = \sqrt{\frac{U}{\nu}}$ é uma variável aleatória com distribuição $\sqrt{\chi_\nu^2/\nu}$, cuja função densidade de probabilidade é dada pela equação (5.1.4), independente em relação a $\mathbf{Z}$.

Tomando-se

$$\max_i(T_i) = \frac{\max_i(Z_i)}{\sqrt{\frac{U}{\nu}}}, \quad (5.4.12)$$

e definindo $Z = \max_i(Z_i)$ e $D = \max_i(T_i)$, a variável aleatória de interesse é $D = Z/S$.

Teorema 5.4.6. (Função densidade do máximo da t multivariada)

Dado o vetor aleatório $\mathbf{T} = (T_1, T_2, \dots, T_r)^\top$ definido em (5.4.11), o máximo do vetor $\mathbf{T}$, dado por $D = \max_i(T_i)$ e definido em (5.4.12), tem a seguinte função densidade de probabilidade:

$$f_D(d; \boldsymbol{\lambda}, \nu) = \int_0^\infty \left[\int_{-\infty}^\infty \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i y + sd}{\sqrt{1 - \lambda_i^2}} \right) \frac{s}{\sqrt{1 - \lambda_i^2}} \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k y + sd}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \right] \times \phi(y) f(s; \nu) dy ds, \quad (5.4.13)$$

em que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_r)$.

Demonstração:

Para obter a função densidade de probabilidade de $D = \max_i(T_i)$, inicialmente deve-se obter a função densidade conjunta de Z e de S . Como Z e S são independentes, a função densidade conjunta de Z e S é obtida pelo produto de $f_Z(z; \boldsymbol{\lambda})$ e $f_S(s; \nu)$, dadas respectivamente pelas funções (5.4.7) e (5.1.4), ou seja

$$f_{Z,S}(z, s; \boldsymbol{\lambda}, \nu) = \int_{-\infty}^\infty \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}} \right) \frac{1}{\sqrt{1 - \lambda_i^2}} \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k y + z}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \times \phi(y) dy f(s; \nu).$$

Para obter a função densidade conjunta de D e S a partir da conjunta $f_{Z,S}(z, s; \boldsymbol{\lambda}, \nu)$, utilizam-se as transformações $Z = SD$ e $S = S$, cujo jacobiano é dado por

$$\mathbf{J} = \begin{vmatrix} \frac{\partial z}{\partial d} & \frac{\partial z}{\partial s} \\ \frac{\partial s}{\partial d} & \frac{\partial s}{\partial s} \end{vmatrix} = \begin{vmatrix} s & d \\ 0 & 1 \end{vmatrix} = s,$$

com $abs(s) = s$, resultando em

$$f_{D,S}(d, s; \boldsymbol{\lambda}, \nu) = f_{Z,S}(z, s; \boldsymbol{\lambda}, \nu) abs(J) \\ = \int_{-\infty}^\infty \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i y + sd}{\sqrt{1 - \lambda_i^2}} \right) \frac{1}{\sqrt{1 - \lambda_i^2}} \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k y + sd}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \times \phi(y) f(s; \nu) s dy.$$

A função densidade de $D = \max_i(T_i)$ é a marginal da densidade conjunta $f_{D,S}(d,s;\lambda,\nu)$ quando se integra em relação a variável s , que varia de $[0,\infty)$, chegando-se a

$$f_D(d;\lambda,\nu) = \int_0^\infty \left[\int_{-\infty}^\infty \sum_{i=1}^r \left\{ \phi\left(\frac{\lambda_i y + sd}{\sqrt{1-\lambda_i^2}}\right) \frac{s}{\sqrt{1-\lambda_i^2}} \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi\left(\frac{\lambda_k y + sd}{\sqrt{1-\lambda_k^2}}\right) \right] \right\} \right] \times \phi(y) f(s;\nu) dy ds,$$

como se queria mostrar.

Teorema 5.4.7. (Função de distribuição do máximo da t multivariada)

Dado o vetor aleatório $\mathbf{T} = (T_1, T_2, \dots, T_r)^\top$ definido em (5.4.11), o máximo do vetor $\mathbf{T}$, dado por $D = \max_i(T_i)$ e definido em (5.4.12), tem a seguinte função de distribuição de probabilidade:

$$F_D(d;\lambda,\nu) = \int_0^\infty \left[\int_{-\infty}^\infty \prod_{i=1}^r \Phi\left(\frac{\lambda_i y + sd}{\sqrt{1-\lambda_i^2}}\right) \phi(y) dy \right] f(s;\nu) ds, \quad (5.4.14)$$

em que $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_r)$.

Demonstração:

A função de distribuição de probabilidade da variável $D = \max_i(T_i)$ é obtida integrando-se $f_D(d;\lambda,\nu)$ dada em (5.4.13), em relação a d , da seguinte forma

$$F_D(d;\lambda,\nu) = \int_{-\infty}^d \left\{ \int_0^\infty \left[\int_{-\infty}^\infty \sum_{i=1}^r \left\{ \phi\left(\frac{\lambda_i y + sw}{\sqrt{1-\lambda_i^2}}\right) \frac{s}{\sqrt{1-\lambda_i^2}} \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi\left(\frac{\lambda_k y + sw}{\sqrt{1-\lambda_k^2}}\right) \right] \right\} \right] \right\} \phi(y) f(s;\nu) dy ds dw,$$

obtendo-se

$$F_D(d;\lambda,\nu) = \int_0^\infty \left\{ \int_{-\infty}^\infty \prod_{i=1}^r \left[\Phi\left(\frac{\lambda_i y + sw}{\sqrt{1-\lambda_i^2}}\right) \right]_{w=-\infty}^{w=d} \phi(y) dy \right\} f(s;\nu) ds,$$

pois

$$\begin{aligned} \frac{\partial[f_1(q)f_2(q)\dots f_3(q)]}{\partial q} &= \sum_{i=1}^r f'_i(q) \prod_{\substack{k=1 \\ k \neq i}}^r f_k(q) \\ \Rightarrow \int_q \frac{\partial[f_1(q)f_2(q)\dots f_3(q)]}{\partial q} &= \int_q \sum_{i=1}^r f'_i(q) \prod_{\substack{k=1 \\ k \neq i}}^r f_k(q) = \left[\prod_{i=1}^r f_i(q) \right]_{q=a}^{q=b}. \end{aligned}$$

Como

$$\lim_{w \rightarrow -\infty} \Phi \left(\frac{\lambda_i y + sw}{\sqrt{1 - \lambda_i^2}} \right) = \Phi \left(\frac{-\infty}{\sqrt{1 - \lambda_i^2}} \right) = \Phi(-\infty) = 0,$$

tem-se que

$$F_D(d; \boldsymbol{\lambda}, \nu) = \int_0^\infty \left[\int_{-\infty}^\infty \prod_{i=1}^r \Phi \left(\frac{\lambda_i y + ds}{\sqrt{1 - \lambda_i^2}} \right) \phi(y) dy \right] f(s; \nu) ds,$$

como se queria mostrar.

A função de distribuição do máximo da t multivariada, dada por (5.4.14), é utilizada no teste de Dunnett unilateral quando os graus de liberdade são finitos.

Teorema 5.4.8. (*Função densidade e função de distribuição do máximo do módulo da t multivariada*)

Tomando o vetor aleatório $\mathbf{T} = (T_1, T_2, \dots, T_r)^\top$ dado em (5.4.11), define-se o máximo do módulo do vetor $\mathbf{T}$ como

$$\max_i(|T_i|) = \frac{\max_i(|Z_i|)}{\sqrt{\frac{U}{\nu}}}.$$

Considerando $|Z| = \max_i(|Z_i|)$, $S = \sqrt{U/\nu}$ uma variável aleatória com distribuição $\sqrt{\chi_\nu^2/\nu}$, cuja função densidade de probabilidade é dada pela equação (5.1.4) e a variável aleatória de interesse $|D| = \max_i(|T_i|)$, a função densidade de probabilidade de $|D|$ é

$$\begin{aligned}
f_{|D|}(|d|; \boldsymbol{\lambda}, \nu) &= \int_0^\infty \int_{-\infty}^\infty \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + |d|s}{\sqrt{1 - \lambda_i^2}} \right) + \phi \left(\frac{\lambda_i y - |d|s}{\sqrt{1 - \lambda_i^2}} \right) \right] \left(\frac{s}{\sqrt{1 - \lambda_i^2}} \right) \right) \right. \\
&\quad \times \left. \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + |d|s}{\sqrt{1 - \lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - |d|s}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) f(s; \nu) dy ds.
\end{aligned}
\tag{5.4.15}$$

e a sua função de distribuição é dada por

$$\begin{aligned}
F_{|D|}(|d|; \boldsymbol{\lambda}, \nu) &= \int_0^\infty \left\{ \int_{-\infty}^\infty \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + |d|s}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - |d|s}{\sqrt{1 - \lambda_i^2}} \right) \right] \right. \\
&\quad \times \left. \phi(y) dy \right\} f(s; \nu) ds,
\end{aligned}
\tag{5.4.16}$$

em que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_r)$.

Demonstração:

Para se obter a função densidade de probabilidade de $|D|$, inicialmente obtém-se a função densidade conjunta de $|Z|$ e de S . Como $|Z|$ e S são independentes, a função densidade de probabilidade conjunta de $|Z|$ e S é obtida pelo produto de $f_{|Z|}(z; \boldsymbol{\lambda})$ e $f_S(s; \nu)$, dadas respectivamente pelas funções (5.4.10) e (5.1.4), ou seja

$$\begin{aligned}
f_{|Z|,S}(z, s; \boldsymbol{\lambda}, \nu) &= \int_{-\infty}^\infty \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + z}{\sqrt{1 - \lambda_i^2}} \right) + \phi \left(\frac{\lambda_i y - z}{\sqrt{1 - \lambda_i^2}} \right) \right] \left(\frac{1}{\sqrt{1 - \lambda_i^2}} \right) \right) \right. \\
&\quad \times \left. \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + z}{\sqrt{1 - \lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - z}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) f(s; \nu) dy.
\end{aligned}$$

A função densidade de probabilidade conjunta de $|D|$ e S é obtida através da transformação $|Z| = |D|S$ e $S = S$, cujo Jacobiano é s , resultando em

$$\begin{aligned}
f_{|D|,S}(|d|,s;\boldsymbol{\lambda},\nu) &= f_{|Z|,S}(z,s;\boldsymbol{\lambda},\nu)abs(J) \\
&= \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + |d|s}{\sqrt{1-\lambda_i^2}} \right) + \phi \left(\frac{\lambda_i y - |d|s}{\sqrt{1-\lambda_i^2}} \right) \right] \left(\frac{1}{\sqrt{1-\lambda_i^2}} \right) \right) \right. \\
&\quad \times \left. \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + |d|s}{\sqrt{1-\lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - |d|s}{\sqrt{1-\lambda_k^2}} \right) \right] \right\} \phi(y) f(s;\nu) s dy.
\end{aligned}$$

A função densidade de $|D|$ é a marginal da densidade conjunta $f_{|D|,S}(|d|,s;\boldsymbol{\lambda},\nu)$ quando se integra em relação a variável s , que varia de $[0,\infty)$, chegando-se a

$$\begin{aligned}
f_{|D|}(|d|;\boldsymbol{\lambda},\nu) &= \int_0^{\infty} \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + |d|s}{\sqrt{1-\lambda_i^2}} \right) + \phi \left(\frac{\lambda_i y - |d|s}{\sqrt{1-\lambda_i^2}} \right) \right] \left(\frac{s}{\sqrt{1-\lambda_i^2}} \right) \right) \right. \\
&\quad \times \left. \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + |d|s}{\sqrt{1-\lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - |d|s}{\sqrt{1-\lambda_k^2}} \right) \right] \right\} \phi(y) f(s;\nu) dy ds,
\end{aligned}$$

que é a (5.4.15), como se queria mostrar.

A função de distribuição de probabilidade da variável $|D|$ é obtida integrando-se a função densidade $f_{|D|}(|d|;\boldsymbol{\lambda},\nu)$ obtida em relação a $|d|$, da seguinte forma

$$\begin{aligned}
F_{|D|}(|d|;\boldsymbol{\lambda},\nu) &= \int_{-\infty}^{|d|} \left\{ \int_0^{\infty} \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + ws}{\sqrt{1-\lambda_i^2}} \right) + \phi \left(\frac{\lambda_i y - ws}{\sqrt{1-\lambda_i^2}} \right) \right] \left(\frac{s}{\sqrt{1-\lambda_i^2}} \right) \right) \right. \right. \\
&\quad \times \left. \left. \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + ws}{\sqrt{1-\lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - ws}{\sqrt{1-\lambda_k^2}} \right) \right] \right\} \phi(y) f(s;\nu) dy ds \right\} dw,
\end{aligned}$$

resultando em

$$F_{|D|}(|d|;\boldsymbol{\lambda},\nu) = \int_0^{\infty} \left\{ \int_{-\infty}^{\infty} \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + |d|s}{\sqrt{1-\lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - |d|s}{\sqrt{1-\lambda_i^2}} \right) \right] \phi(y) dy \right\} f(s;\nu) ds,$$

que é a (5.4.16), como se queria mostrar.

A função de distribuição do máximo do módulo da t multivariada, dada por (5.4.16) é utilizada no teste de Dunnett bilateral quando os graus de liberdade são finitos.

Em síntese, considerando uma matriz de correlação ρ que satisfaz a estrutura de correlação produto, denota-se por pontos críticos superiores α da distribuição do máximo da t multivariada $\max_i(T_i)$ ou do máximo do módulo da t multivariada $\max_i(|T_i|)$, os valores de d ou $|d|$ soluções das equações

$$\begin{aligned} F_D(d; \lambda, \nu) &= \int_0^\infty \left\{ \int_{-\infty}^\infty \prod_{i=1}^r \Phi \left[\frac{\lambda_i y + sd}{\sqrt{1 - \lambda_i^2}} \right] \phi(y) dy \right\} f(s; \nu) ds \\ &= 1 - \alpha \end{aligned}$$

e

$$\begin{aligned} F_{|D|}(|d|; \lambda, \nu) &= \int_0^\infty \left\{ \int_{-\infty}^\infty \prod_{i=1}^r \left\{ \Phi \left[\frac{\lambda_i y + |d|s}{\sqrt{1 - \lambda_i^2}} \right] - \Phi \left[\frac{\lambda_i y - |d|s}{\sqrt{1 - \lambda_i^2}} \right] \right\} \phi(y) dy \right\} f(s; \nu) ds \\ &= 1 - \alpha, \end{aligned}$$

respectivamente.

6 METODOLOGIA

Esta tese tem a finalidade de ampliar o algoritmo proposto por Dunlap, Marx e Agamy (1981), atendendo aos testes de Dunnett unilaterais e bilaterais, balanceados ou não e possibilitar a obtenção da probabilidade acumulada, dos quantis da distribuição, invertendo a função de distribuição, considerando a estatística do teste para a distribuição t multivariada não-central.

Portanto, nesta seção está desenvolvida a obtenção detalhada da distribuição das estatísticas do teste de Dunnett considerando um vetor de não-centralidade δ . Inicialmente obteve-se a distribuição do máximo e do máximo do módulo da normal multivariada não-central. A partir dessas distribuições chega-se a distribuição t multivariada não-central aplicável às estatísticas do teste de Dunnett.

Também são descritas a metodologia para a resolução numérica dessas integrais utilizando quadraturas gaussianas antes de construir o pseudocódigo. O ANEXO B apresenta todas as rotinas implementadas no *software R*.

6.1 Distribuição do máximo e do máximo do módulo da normal

multivariada não-central com parâmetro de não-centralidade $\delta = \mu$

Para obter as distribuições das estatísticas do teste de Dunnett considerando como parâmetro de não-centralidade $\delta = \mu$ utilizou-se as mesmas ideias do que é feito para as distribuições centrais. Nesta subseção desenvolve-se a obtenção das distribuições do máximo e do máximo do módulo da normal multivariada não-central.

Usando a matriz de combinações lineares dada por (5.4.5) e um vetor de constantes δ pode-se obter a transformação $\mathbf{Z} = \mathbf{A}\mathbf{Y} + \delta$, em que $\mathbf{Z} \sim \mathbb{N}_r(\delta, \rho)$ e a matriz ρ satisfaz a condição de estrutura de correlação produto dada por (5.4.4).

O máximo do vetor $\mathbf{Z}$, *dado por* $Z = \max_i Z_i$, tem a seguinte função de distribuição de probabilidade:

$$F_Z(z; \boldsymbol{\delta}, \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \phi(y) dy, \quad (6.1.1)$$

em que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_r)$ e $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_r)^\top$.

Demonstração:

Considerando as hipóteses, tem-se a transformação $\mathbf{Z} = \boldsymbol{\Lambda}\mathbf{Y} + \boldsymbol{\delta}$, ou seja

$$\begin{aligned} \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_r \end{pmatrix} &= \begin{pmatrix} \sqrt{1 - \lambda_1^2} & 0 & \dots & 0 & -\lambda_1 \\ 0 & \sqrt{1 - \lambda_2^2} & \dots & 0 & -\lambda_2 \\ 0 & 0 & \dots & 0 & -\lambda_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \sqrt{1 - \lambda_r^2} & -\lambda_r \end{pmatrix} \times \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_r \\ Y_0 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_r \end{pmatrix} \\ &\Rightarrow \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_r \end{pmatrix} = \begin{pmatrix} \sqrt{1 - \lambda_1^2} Y_1 - \lambda_1 Y_0 + \delta_1 \\ \sqrt{1 - \lambda_2^2} Y_2 - \lambda_2 Y_0 + \delta_2 \\ \vdots \\ \sqrt{1 - \lambda_r^2} Y_r - \lambda_r Y_0 + \delta_r \end{pmatrix}. \end{aligned}$$

Assim, $E(\mathbf{Z}) = E(\boldsymbol{\Lambda}\mathbf{Y} + \boldsymbol{\delta}) = E(\boldsymbol{\Lambda}\mathbf{Y}) + \boldsymbol{\delta} = \boldsymbol{\Lambda}E(\mathbf{Y}) + \boldsymbol{\delta} = \boldsymbol{\delta}$

$$\Rightarrow E(\mathbf{Z}) = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_r \end{pmatrix}$$

e $\text{Cov}(\mathbf{Z}) = \text{Cov}(\mathbf{\Lambda Y} + \boldsymbol{\delta}) = \mathbf{\Lambda} \text{Cov}(\mathbf{Y}) \mathbf{\Lambda}^\top = \mathbf{\Lambda} \mathbf{I} \mathbf{\Lambda}^\top = \mathbf{\Lambda} \mathbf{\Lambda}^\top$, ou seja

$$\text{Cov}(\mathbf{Z}) = \begin{pmatrix} \left[\sqrt{1 - \lambda_1^2} \right]^2 + \lambda_1^2 & \lambda_1 \lambda_2 & \dots & \lambda_1 \lambda_r \\ \lambda_2 \lambda_1 & \left[\sqrt{1 - \lambda_2^2} \right]^2 + \lambda_2^2 & \dots & \lambda_2 \lambda_r \\ \lambda_3 \lambda_1 & \lambda_3 \lambda_2 & \dots & \lambda_3 \lambda_r \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_r \lambda_1 & \lambda_r \lambda_2 & \dots & \left[\sqrt{1 - \lambda_r^2} \right]^2 + \lambda_r^2 \end{pmatrix}$$

$$\text{Cov}(\mathbf{Z}) = \begin{pmatrix} 1 & \lambda_1 \lambda_2 & \lambda_1 \lambda_3 & \dots & \lambda_1 \lambda_r \\ \lambda_2 \lambda_1 & 1 & \lambda_2 \lambda_3 & \dots & \lambda_2 \lambda_r \\ \lambda_3 \lambda_1 & \lambda_3 \lambda_2 & 1 & \dots & \lambda_3 \lambda_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_r \lambda_1 & \lambda_r \lambda_2 & \lambda_r \lambda_3 & \dots & 1 \end{pmatrix} = \boldsymbol{\rho}.$$

A matriz $\boldsymbol{\rho}$ satisfaz a condição de estrutura de correlação produto dada por (5.4.4). Assim, o vetor aleatório $\mathbf{Z} = (Z_1, Z_2, \dots, Z_r)^\top$, cuja forma geral do i -ésimo termo é

$$Z_i = \sqrt{1 - \lambda_i^2} Y_i - \lambda_i Y_0 + \delta_i, \quad \text{com } 1 \leq i \leq r,$$

tem distribuição normal multivariada $\mathbf{Z} \sim \mathbb{N}_r(\boldsymbol{\delta}, \boldsymbol{\rho})$ e a função de densidade de $\mathbf{Z}$ é dada por

$$f_{\mathbf{Z}}(\mathbf{z}; \boldsymbol{\delta}, \boldsymbol{\rho}) = (2\pi)^{-\frac{r}{2}} |\boldsymbol{\rho}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{z} - \boldsymbol{\delta})^\top \boldsymbol{\rho}^{-1} (\mathbf{z} - \boldsymbol{\delta}) \right\}.$$

Assim,

$$\begin{aligned}
 F_{Z_i}(z; \delta_i, \lambda_i) &= P(Z_i \leq z) \\
 &= P\left(\sqrt{1 - \lambda_i^2} Y_i - \lambda_i Y_0 + \delta_i \leq z\right) \\
 &= P\left(Y_i \leq \frac{\lambda_i Y_0 + z - \delta_i}{\sqrt{1 - \lambda_i^2}}\right).
 \end{aligned}$$

Cada $Y_i \sim \mathbb{N}(0,1)$, mas a acumulada $F_{Z_i}(z)$ depende de Y_0 , que é uma variável aleatória. Condiçãoando ao conhecimento do valor de Y_0 que será denotado por y tem-se

$$F_{Z_i}(z|Y_0 = y) = P\left(Y_i \leq \frac{\lambda_i Y_0 + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \middle| Y_0 = y\right) = P\left(Y_i \leq \frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}}\right),$$

Logo, usando o resultado de probabilidades condicionais, tem-se

$$F_{Z_i}(z; \delta_i, \lambda_i) = \int_{-\infty}^{\infty} \Phi\left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}}\right) \phi(y) dy.$$

Usando o resultado da distribuição do máximo de variáveis aleatórias, dado por

$$P(Z_1 \leq z, Z_2 \leq z, \dots, Z_r \leq z) = P\left(\max_i Z_i \leq z\right) = \prod_{i=1}^r \Phi\left(\frac{\lambda_i Y_0 + z - \delta_i}{\sqrt{1 - \lambda_i^2}}\right).$$

com $Z = \max_i Z_i$ tem-se que

$$F_Z(z; \boldsymbol{\delta}, \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \phi(y) dy,$$

como se queria mostrar.

Esta função da distribuição do máximo da normal multivariada não-central com parâmetro de não-centralidade $\boldsymbol{\delta} = \boldsymbol{\mu}$ é utilizada no teste de Dunnett unilaterial quando os graus de liberdade tendem ao infinito.

Teorema 6.1.2. (*Função densidade do máximo da normal multivariada não-central*)

Dadas as mesmas condições do teorema 6.2.1., o máximo do vetor $\mathbf{Z} = \boldsymbol{\Lambda}\mathbf{Y} + \boldsymbol{\delta}$ dado por $Z = \max_i Z_i$ tem a seguinte função densidade de probabilidade:

$$f_Z(z; \boldsymbol{\delta}, \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \frac{1}{\sqrt{1 - \lambda_i^2}} \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k y + z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) dy, \quad (6.1.2)$$

em que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_r)$ e $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_r)^\top$.

Demonstração:

A função densidade de probabilidade do máximo da normal multivariada não-central, $Z = \max_i(Z_i)$, é obtida derivando-se (6.1.1) em relação a z , conforme (5.4.8) obtendo-se

$$f_Z(z; \boldsymbol{\delta}, \boldsymbol{\lambda}) = \frac{\partial}{\partial z} \left\{ \int_{-\infty}^{\infty} \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \phi(y) dy \right\},$$

ou seja

$$f_Z(z; \boldsymbol{\delta}, \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \frac{1}{\sqrt{1 - \lambda_i^2}} \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k y + z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) dy,$$

como se queria mostrar.

Teorema 6.1.3. (*Função de distribuição do máximo do módulo da normal multivariada não-central*)

Dadas as mesmas condições do teorema 6.2.1., o máximo do módulo do vetor $\mathbf{Z} = \boldsymbol{\Lambda} \mathbf{Y} + \boldsymbol{\delta}$ dado por $|Z| = \max_i |Z_i|$ tem a seguinte função de distribuição de probabilidade:

$$F_{|Z|}(z; \boldsymbol{\delta}, \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \left\{ \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \right\} \phi(y) dy, \quad (6.1.3)$$

em que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_r)$ e $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_r)^\top$.

Demonstração:

De modo similar ao demonstrado no teorema 6.2.1, porém usando o resultado da distribuição do máximo do módulo de uma variável aleatória dado por

$$P(|Z_1| \leq z, |Z_2| \leq z, \dots, |Z_r| \leq z) = P\left(\max_i |Z_i| \leq z\right) = P(-z \leq \max_i Z_i \leq z,)$$

tem-se

$$P\left(\max_i |Z_i| \leq z\right) = \prod_{i=1}^r P\left(\frac{\lambda_i Y_0 - z - \delta_i}{\sqrt{1 - \lambda_i^2}} \leq Y_i \leq \frac{\lambda_i Y_0 + z - \delta_i}{\sqrt{1 - \lambda_i^2}}\right),$$

ou seja, com $|Z| = \max_i |Z_i|$

$$F_{|Z|}(z; \boldsymbol{\delta}, \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \left\{ \prod_{i=1}^r \left[\Phi\left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}}\right) - \Phi\left(\frac{\lambda_i y - z - \delta_i}{\sqrt{1 - \lambda_i^2}}\right) \right] \right\} \phi(y) dy,$$

como se queria mostrar.

A função da distribuição do máximo do módulo da normal multivariada não-central com parâmetro de não-centralidade $\boldsymbol{\delta} = \boldsymbol{\mu}$ é utilizada no teste de Dunnett bilateral quando os graus de liberdade tendem ao infinito.

Teorema 6.1.4. (*Função densidade do máximo do módulo da normal multivariada não-central*)

Dadas as mesmas condições do teorema 6.2.1., o máximo do vetor $\mathbf{Z} = \Lambda \mathbf{Y} + \boldsymbol{\delta}$ dado por $|Z| = \max_i(|Z_i|)$ tem a seguinte função densidade de probabilidade:

$$f_{|Z|}(z; \boldsymbol{\delta}, \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) + \phi \left(\frac{\lambda_i y - z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \left(\frac{1}{\sqrt{1 - \lambda_i^2}} \right) \right) \right. \\ \left. \times \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) dy. \quad (6.1.4)$$

em que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_r)$ e $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_r)^\top$.

Demonstração:

A função densidade de probabilidade de $|Z| = \max_i(|Z_i|)$ é obtida derivando-se (6.1.3) em relação a z , conforme (5.4.8) obtendo-se

$$f_{|Z|}(z; \boldsymbol{\delta}, \boldsymbol{\lambda}) = \frac{\partial}{\partial z} \left[\int_{-\infty}^{\infty} \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \phi(y) dy \right] \\ = \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \frac{1}{\sqrt{1 - \lambda_i^2}} \right] - \left[\phi \left(\frac{\lambda_i y - z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right. \right. \right. \\ \left. \left. \times \frac{-1}{\sqrt{1 - \lambda_i^2}} \right] \right) \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) dy.$$

Assim, a função de densidade do máximo da normal multivariada não-central é

$$f_{|Z|}(z; \boldsymbol{\delta}, \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) + \phi \left(\frac{\lambda_i y - z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \left(\frac{1}{\sqrt{1 - \lambda_i^2}} \right) \right) \right. \\ \left. \times \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) dy.$$

como se queria mostrar.

6.2 Distribuição do máximo e do máximo do módulo da t multivariada não-central com parâmetro de não-centralidade $\delta = \mu$

Para se obter a distribuição do máximo da t multivariada não-central com parâmetro de não-centralidade $\delta = \mu$, de forma semelhante às distribuições centrais toma-se $\mathbf{T} = [T_1, T_2, \dots, T_r]$ um vetor aleatório definido por

$$\mathbf{T} = \frac{\mathbf{Z}}{\sqrt{\frac{U}{\nu}}}, \quad (6.2.1)$$

em que $S = \sqrt{U/\nu} = \hat{\sigma}/\sigma$ é uma variável aleatória com distribuição de uma raiz de qui-quadrado dividido pelos graus de liberdade, cuja função densidade de probabilidade é dada pela equação (5.1.4), independente em relação a $\mathbf{Z}$.

Tomando-se

$$\max_i(T_i) = \frac{\max_i(Z_i)}{\sqrt{\frac{U}{\nu}}} \quad (6.2.2)$$

e chamando de $Z = \max_i(Z_i)$ e $D = \max_i(T_i)$, a variável aleatória de interesse é $D = Z/S$, em que cada $D_i = Q_i/S \sim t_\nu(\delta_i)$.

Teorema 6.2.1. (Função densidade do máximo da t multivariada não-central)

Dado o vetor aleatório $\mathbf{T} = (T_1, T_2, \dots, T_r)^\top$ definido em (6.2.1), o máximo do vetor $\mathbf{T}$, dado por $D = \max_i(T_i)$ e definido em (6.2.2), tem a seguinte função densidade de probabilidade:

$$f_D(d; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_0^\infty \left[\int_{-\infty}^\infty \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i y + sd - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \frac{s}{\sqrt{1 - \lambda_i^2}} \right. \right. \\ \left. \left. \times \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k y + sd - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \right] \phi(y) f(s; \nu) dy ds. \quad (6.2.3)$$

em que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_r)$ e $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_r)^\top$.

Demonstração:

Para obter a função densidade de probabilidade de $D = \max_i(T_i)$, inicialmente deve-se obter a função densidade conjunta de Z e de S . Como $Z = \max_i(Z_i)$ e S são independentes, a função densidade conjunta de Z e S é obtida pelo produto de $f_Z(z; \boldsymbol{\delta}, \boldsymbol{\lambda})$ e $f_S(s; \nu)$, dadas, respectivamente, pelas funções (6.1.2) e (5.1.4), ou seja

$$f_{Z,S}(z, s; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_{-\infty}^\infty \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \frac{1}{\sqrt{1 - \lambda_i^2}} \right. \\ \left. \times \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k y + z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) f(s; \nu) dy.$$

Utilizando as transformações $Z = DS$ e $S = S$, cujo Jacobiano é dado por

$$\mathbf{J} = \begin{vmatrix} \frac{\partial z}{\partial d} & \frac{\partial z}{\partial s} \\ \frac{\partial s}{\partial d} & \frac{\partial s}{\partial s} \end{vmatrix} = \begin{vmatrix} s & d \\ 0 & 1 \end{vmatrix} = s,$$

com $abs(s) = s$, obtém-se a função densidade conjunta de $D = \max_i(T_i)$ e S , dada por $f_{D,S}(d, s; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = f_{Z,S}(z, s; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) abs(\mathbf{J})$, ou seja,

$$f_{D,S}(d,s;\boldsymbol{\delta},\boldsymbol{\lambda},\nu) = \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \phi\left(\frac{\lambda_i y + sd - \delta_i}{\sqrt{1 - \lambda_i^2}}\right) \frac{1}{\sqrt{1 - \lambda_i^2}} \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi\left(\frac{\lambda_k y + sd - \delta_k}{\sqrt{1 - \lambda_k^2}}\right) \right] \right. \\ \left. \times \right\} \phi(y) f(s;\nu) s dy.$$

A função densidade de probabilidade de $D = \max_i(T_i)$ é a marginal da densidade conjunta $f_{D,S}(d,s;\boldsymbol{\delta},\boldsymbol{\lambda},\nu)$ quando se integra em relação a variável s , que varia de $[0,\infty)$, chegando-se a

$$f_D(d;\boldsymbol{\delta},\boldsymbol{\lambda},\nu) = \int_0^{\infty} \left[\int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \phi\left(\frac{\lambda_i y + sd - \delta_i}{\sqrt{1 - \lambda_i^2}}\right) \frac{s}{\sqrt{1 - \lambda_i^2}} \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi\left(\frac{\lambda_k y + sd - \delta_k}{\sqrt{1 - \lambda_k^2}}\right) \right] \right. \right. \\ \left. \left. \times \right\} \phi(y) f(s;\nu) dy ds, \right]$$

como se queria mostrar.

Teorema 6.2.2. (*Função de distribuição do máximo da t multivariada não-central*)

Dado o vetor aleatório $\mathbf{T} = (T_1, T_2, \dots, T_r)^\top$ definido em (6.2.1), o máximo do vetor $\mathbf{T}$, dado por $D = \max_i(T_i)$ e definido em (6.2.2), tem a seguinte função de distribuição de probabilidade:

$$F_D(d;\boldsymbol{\delta},\boldsymbol{\lambda},\nu) = \int_0^{\infty} \left[\int_{-\infty}^{\infty} \prod_{i=1}^r \Phi\left(\frac{\lambda_i y + sd - \delta_i}{\sqrt{1 - \lambda_i^2}}\right) \phi(y) dy \right] f(s;\nu) ds, \quad (6.2.4)$$

em que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_r)$ e $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_r)^\top$.

Demonstração:

A função de distribuição de probabilidade da variável $D = \max_i(T_i)$ é obtida integrando-se a função densidade $f_D(d; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu)$ em relação a d , da seguinte forma

$$F_D(d; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_{-\infty}^d \left\{ \int_0^{\infty} \left[\int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i y + sw - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \frac{s}{\sqrt{1 - \lambda_i^2}} \right. \right. \right. \\ \left. \left. \times \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k y + sw - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \right] dy \right\} f(s; \nu) ds dw,$$

obtendo-se

$$F_D(d; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_0^{\infty} \left[\int_{-\infty}^{\infty} \prod_{i=1}^r \Phi \left(\frac{\lambda_i y + sd - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \phi(y) dy \right] f(s; \nu) ds,$$

como se queria mostrar.

A função da distribuição do máximo da t multivariada não-central com parâmetro de não-centralidade $\boldsymbol{\delta} = \boldsymbol{\mu}$ é utilizada no teste de Dunnett unilateral quando os graus de liberdade são finitos.

Teorema 6.2.3. (Função densidade e função de distribuição do máximo do módulo da t multivariada)

Tomando o vetor aleatório $\mathbf{T} = (T_1, T_2, \dots, T_r)^\top$ dado em (6.2.1), define-se o máximo do módulo do vetor $\mathbf{T}$ como

$$\max_i(|T_i|) = \frac{\max_i(|Z_i|)}{\sqrt{\frac{U}{\nu}}}.$$

Considerando $|Z| = \max_i(|Z_i|)$, $S = \sqrt{U/\nu}$ uma variável aleatória com distribuição $\sqrt{\chi_\nu^2/\nu}$, cuja função densidade de probabilidade é dada pela equação (5.1.4) e a variável aleatória de interesse $|D| = \max_i(|T_i|)$, a função densidade de probabilidade de $|D|$ é

$$\begin{aligned} f_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) &= \int_0^\infty \int_{-\infty}^\infty \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) + \phi \left(\frac{\lambda_i y - |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \right. \right. \\ &\quad \times \left. \left(\frac{s}{\sqrt{1 - \lambda_i^2}} \right) \right) \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + |d|s - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - |d|s - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \\ &\quad \times \phi(y) f(s; \nu) dy ds. \end{aligned} \quad (6.2.5)$$

e a sua função de distribuição é dada por

$$\begin{aligned} F_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) &= \int_0^\infty \left\{ \int_{-\infty}^\infty \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \right. \\ &\quad \times \left. \phi(y) dy \right\} f(s; \nu) ds, \end{aligned} \quad (6.2.6)$$

em que $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_r)$ e $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_r)^\top$.

Demonstração:

Para obter a função densidade de probabilidade de $|D| = \max_i(|T_i|)$, inicialmente deve-se obter a função densidade conjunta de $|Z|$ e de S . Como $|Z|$ e S são independentes, a função densidade de probabilidade conjunta de $|Z|$ e S é dada por $f_{|Z|,S}(z, s; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = f_{|Z|}(z; \boldsymbol{\delta}, \boldsymbol{\lambda}) \cdot f_S(s; \nu)$, ou seja

$$\begin{aligned} f_{|Z|,S}(z, s; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) &= \int_{-\infty}^\infty \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) + \phi \left(\frac{\lambda_i y - z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \left(\frac{1}{\sqrt{1 - \lambda_i^2}} \right) \right) \right. \\ &\quad \times \left. \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) f(s; \nu) dy. \end{aligned}$$

A função densidade de probabilidade conjunta de $|D| = \max_i(|T_i|)$ e S é obtida através da transformação $|D| = |Z|S$ e $S = S$, cujo Jacobiano é s , resultando em $f_{|D|,S}(|d|,s;\boldsymbol{\delta},\boldsymbol{\lambda},\nu) = f_{|Z|,S}(z,s;\boldsymbol{\delta},\boldsymbol{\lambda},\nu)abs(\mathbf{J})$, ou seja

$$\begin{aligned} f_{|D|,S}(|d|,s;\boldsymbol{\delta},\boldsymbol{\lambda},\nu) &= \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi\left(\frac{\lambda_i y + |d|s - \delta_i}{\sqrt{1-\lambda_i^2}}\right) + \phi\left(\frac{\lambda_i y - |d|s - \delta_i}{\sqrt{1-\lambda_i^2}}\right) \right] \right. \right. \\ &\quad \times \left. \left(\frac{1}{\sqrt{1-\lambda_i^2}} \right) \right) \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi\left(\frac{\lambda_k y + |d|s - \delta_k}{\sqrt{1-\lambda_k^2}}\right) - \Phi\left(\frac{\lambda_k y - |d|s - \delta_k}{\sqrt{1-\lambda_k^2}}\right) \right] \\ &\quad \times \left. \right\} \phi(y) f(s;\nu) s dy. \end{aligned}$$

A função densidade de $|D|$ é a marginal da densidade conjunta $f_{|D|,S}(|d|,s;\boldsymbol{\delta},\boldsymbol{\lambda},\nu)$ quando se integra em relação a variável s , que varia de $[0,\infty)$, chegando-se a

$$\begin{aligned} f_{|D|}(|d|;\boldsymbol{\delta},\boldsymbol{\lambda},\nu) &= \int_0^{\infty} \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi\left(\frac{\lambda_i y + |d|s - \delta_i}{\sqrt{1-\lambda_i^2}}\right) + \phi\left(\frac{\lambda_i y - |d|s - \delta_i}{\sqrt{1-\lambda_i^2}}\right) \right] \right. \right. \\ &\quad \times \left. \left(\frac{s}{\sqrt{1-\lambda_i^2}} \right) \right) \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi\left(\frac{\lambda_k y + |d|s - \delta_k}{\sqrt{1-\lambda_k^2}}\right) - \Phi\left(\frac{\lambda_k y - |d|s - \delta_k}{\sqrt{1-\lambda_k^2}}\right) \right] \right\} \\ &\quad \times \phi(y) f(s;\nu) dy ds, \end{aligned}$$

que é a expressão (6.2.5), como se queria mostrar.

A função de distribuição de probabilidade da variável $|D| = \max_i(|T_i|)$ é obtida integrando-se $f_{|D|}(|d|;\boldsymbol{\delta},\boldsymbol{\lambda},\nu)$ em relação a d , da seguinte forma

$$\begin{aligned} F_{|D|}(|d|;\boldsymbol{\delta},\boldsymbol{\lambda},\nu) &= \int_{-\infty}^{|d|} \left\{ \int_0^{\infty} \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi\left(\frac{\lambda_i y + ws - \delta_i}{\sqrt{1-\lambda_i^2}}\right) + \phi\left(\frac{\lambda_i y - ws - \delta_i}{\sqrt{1-\lambda_i^2}}\right) \right] \right. \right. \right. \\ &\quad \times \left. \left. \frac{s}{\sqrt{1-\lambda_i^2}} \right) \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi\left(\frac{\lambda_k y + ws - \delta_k}{\sqrt{1-\lambda_k^2}}\right) - \Phi\left(\frac{\lambda_k y - ws - \delta_k}{\sqrt{1-\lambda_k^2}}\right) \right] \right\} \phi(y) f(s;\nu) dy ds \right\} dw \end{aligned}$$

resultando em

$$F_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_0^\infty \left\{ \int_{-\infty}^\infty \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \right. \\ \left. \times \phi(y) dy \right\} f(s; \nu) ds,$$

que é a expressão (6.2.6), como se queria mostrar.

A função da distribuição do máximo do módulo da t multivariada não-central com parâmetro de não-centralidade $\boldsymbol{\delta} = \boldsymbol{\mu}$ é utilizada no teste de Dunnett bilateral quando os graus de liberdade são finitos.

Em síntese, considerando uma matriz de correlação $\boldsymbol{\rho}$ que satisfaz a estrutura de correlação produto, denota-se por pontos críticos superiores α da distribuição do máximo da t multivariada não-central ou do máximo do módulo da t multivariada não-central, com parâmetro de não-centralidade $\boldsymbol{\delta} = \boldsymbol{\mu}$, os valores de d ou $|d|$, que sejam soluções das equações

$$F_D(d; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_0^\infty \left[\int_{-\infty}^\infty \prod_{i=1}^r \Phi \left(\frac{\lambda_i y + sd - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \phi(y) dy \right] f(s; \nu) ds = 1 - \alpha$$

e

$$F_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_0^\infty \left\{ \int_{-\infty}^\infty \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \right. \\ \left. \times \phi(y) dy \right\} f(s; \nu) ds = 1 - \alpha,$$

respectivamente.

6.3 Detalhes da implementação das funções

Abaixo estão descritos os parâmetros de entrada utilizados nas diversas funções desenvolvidas na biblioteca *nCDunnett*.

- a) d : a estatística calculada para o quantil da distribuição de Dunnett, dado por

$$d = \frac{\hat{\mu}_i - \hat{\mu}_{r+1}}{\hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}}, \text{ para cada } i = 1, \dots, r;$$

- b) λ : o vetor de correlações. A partir dos tamanhos amostrais, dados por $\mathbf{n} = (n_1, n_2, \dots, n_r, n_{r+1})$, em que as $i = 1, \dots, r$ primeiras amostras são dos tratamentos em teste e a $(r + 1)$ -ésima amostra é do tratamento controle, devem ser calculados os coeficientes de correlação entre cada contraste formando o vetor $\lambda = (\lambda_1^2, \lambda_2^2, \dots, \lambda_r^2)$, em que cada $\lambda_i^2 = \rho_i = n_i / (n_i + n_{r+1})$ para $i = 1, \dots, r$;
- c) δ : o vetor contendo os parâmetros de não-centralidade;
- d) n : número de pontos da quadratura;
- e) $1 - \alpha$: a probabilidade ou nível de confiança conjunta dos r contrastes simultaneamente testados;
- f) ν : os graus de liberdade da distribuição da variância amostral residual $\hat{\sigma}^2$.

As funções densidade e de distribuição de uma variável normal padrão utilizadas em todas as rotinas programadas são dadas de forma genérica por

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad \text{e} \quad \Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw.$$

No pacote proposto existem duas possibilidades para executar o teste de Dunnett descritas a seguir:

a) Alternativa 1: Para cada contraste entre o i -ésimo novo tratamento e o tratamento controle determinar a estimativa $\hat{d}_i$. Usando como parâmetros de entrada o valor $\hat{d}_i$; o vetor λ de correlações, o vetor de parâmetros de não-centralidade δ , se houver, os graus de liberdade ν da distribuição da variância amostral residual, a rotina fornece a probabilidade conjunta estimada, da qual se obtém o valor- p do teste. Compara-se o valor- p obtido no contraste com o nível nominal α estabelecido para o teste, e conclui-se sobre a significância da diferença de médias observadas. Repete-se este procedimento para cada um dos r contrastes. Desta forma determina-se quais dos r tratamentos em teste apresentam média significativamente diferente da média do tratamento controle, considerando um nível de significância nominal pré-especificado para as r comparações simultaneamente.

b) Alternativa 2: De outra forma, pode-se estabelecer uma probabilidade ou nível de significância α desejável para o teste simultâneo dos r contrastes. Usando como parâmetros de entrada o valor $1 - \alpha$ de probabilidade de confiança conjunta, o vetor λ de correlações, o vetor de parâmetros de não-centralidade δ , se houver, os graus de liberdade ν , a rotina fornece o quantil d , usando a inversa da função de distribuição da probabilidade $1 - \alpha$ fornecida. Compara-se o valor do quantil d fornecido pelo programa com as r estimativas dos contrastes, e conclui-se sobre a significância da diferença de médias de cada um dos r contrastes. Desta forma também determina-se quais dos tratamentos em teste apresentam média significativamente diferente da média do tratamento controle, considerando um nível de significância α pré-especificado para as r comparações simultaneamente.

O APÊNDICE A traz um exemplo que ilustra essas possibilidades.

6.3.1 Função para a distribuição do teste de Dunnett unilateral para dados não balanceados com infinitos graus de liberdade e parâmetro de não-centralidade δ , denominada de *pNDUD*

Nesta função *pNDUD*, os parâmetros de entrada são o quantil d , que será o mesmo do quantil z , o vetor λ de correlações, o vetor de não-centralidade δ e o número de pontos n a ser utilizado na quadratura, que por padrão é usado $n = 32$.

O valor fornecido pela função é o da probabilidade acumulada $1 - \alpha$. Quando os graus de liberdade tendem ao infinito, a distribuição t multivariada recai na distribuição normal multivariada, tanto no caso central como não-central. Assim, deseja-se resolver a função de distribuição do máximo da normal multivariada não-central com parâmetro de não-centralidade $\delta = \mu$ obtida em (6.1.1), que é dada por

$$F_Z(z; \boldsymbol{\delta}, \boldsymbol{\lambda}) = \int_{-\infty}^{\infty} \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \phi(y) dy,$$

com $Z = \max_i Z_i$.

Chamando o integrando dessa expressão de $g(y; z, \boldsymbol{\delta}, \boldsymbol{\lambda})$, tem-se

$$g(y; z, \boldsymbol{\delta}, \boldsymbol{\lambda}) = \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \phi(y)$$

que está representada na Figura 12 com os parâmetros $z = 1,35$, $\boldsymbol{\delta} = (0,2; 0,3; 0,5)$ e $\boldsymbol{\lambda} = (0,2; 0,5; 0,5)$. Pode-se observar que a função $g(y; z, \boldsymbol{\delta}, \boldsymbol{\lambda})$ tem a forma da função densidade da normal $\phi(y)$, reduzida pelo termo que contém o produtório, que é um número entre 0 e 1. A curva apresenta também um deslocamento da centralidade dada pelo vetor $\boldsymbol{\delta}$, com uma leve assimetria a direita.

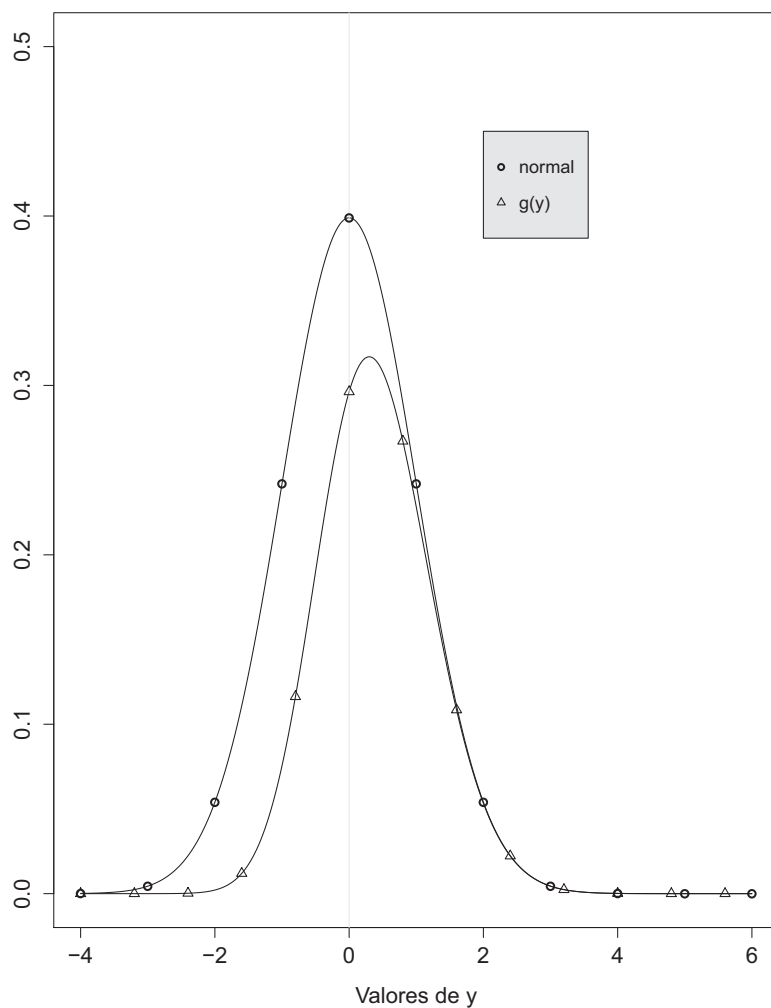


Figura 12 Representação gráfica da função $g(y; z, \delta, \lambda) = g(y; 1,35, (0,2; 0,3; 0,5), (0,2; 0,5; 0,5))$, integrando da expressão (6.1.1), em relação à função densidade da normal $\phi(y)$

Para resolver a integral utilizando quadraturas, optou-se por não utilizar a quadratura Gauss-Hermite e transformar a integral original numa integral com intervalo $[-1,1]$ e avaliá-la por uma quadratura de Gauss-Legendre. Esta decisão se deveu principalmente por dois fatores:

- a) na implementação das quadraturas Gauss-Hermite e Gauss-Laguerre surgem grandes valores (ou valores muito pequenos) dos nós das respectivas quadraturas. Nos computadores, a precisão é finita. Assim, determinadas funções apresentam imagens que computacionalmente dão resultados infinitos ou são arredondados erroneamente para zero, quando computadas para argumentos muito grandes ou muito pequenos. Isso decorre da precisão finita dos números reais representados por máquinas. Esses erros são denotados por *overflow* ou *underflow*. Determinados casos das distribuições da estatística do teste de Dunnett, quando submetidos a esses tipos de quadraturas, apresentaram problemas dessa natureza, o que conduziu a uma menor acurácia da integração, quando comparada com a quadratura Gauss-Legendre.
- b) as funções que serão integradas não são definidas para $x < 0$, o que pode acarretar dificuldade para avaliar estas integrais em intervalos infinitos, conforme descrito da seção (4.1.3). Uma opção para resolver estas integrais mantendo o intervalo infinito é o método de quadratura gaussiana proposto por Steel e Gelbard (1969).

Foi realizada uma mudança da variável y para a variável t , do tipo apresentada em (4.1.8), que é dada por

$$\int_{-\infty}^{\infty} g(y; z, \boldsymbol{\delta}, \boldsymbol{\lambda}) dy = \int_{-1}^1 g\left(\frac{t}{1-t^2}\right) \frac{1+t^2}{(1-t^2)^2} dt$$

em que $z = d$ é o valor da estatística calculada em cada contraste.

$$\begin{aligned} \int_{-\infty}^{\infty} g(y; d, \boldsymbol{\delta}, \boldsymbol{\lambda}) dy &\approx \sum_{j=1}^n w_j \left[\prod_{i=1}^r \Phi\left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2}\right) + d - \delta_i}{\sqrt{1 - \lambda_i^2}}\right) \right] \phi\left(\frac{t_j}{1-t_j^2}\right) \\ &\times \left[\frac{1+t_j^2}{(1-t_j^2)^2} \right], \end{aligned}$$

com t_j um nó da quadratura Gauss-Legendre e w_j o seu respectivo peso.

6.3.2 Função para a distribuição do teste de Dunnett bilateral para dados não balanceados com infinitos graus de liberdade e parâmetro de não-centralidade δ , denominada de *pNDBD*

Nesta função *pNDBD*, os parâmetros de entrada são o módulo do quantil $|d|$, que será o mesmo módulo do quantil $|z|$, o vetor λ de correlações, o vetor de não-centralidade δ e o número de pontos n a ser utilizado na quadratura, que por padrão é usado $n = 32$. O valor fornecido pela função é o da probabilidade acumulada $1 - \alpha$.

Neste caso, deseja-se resolver a função de distribuição do máximo do módulo da normal multivariada não-central com parâmetro de não-centralidade $\delta = \mu$ obtida em (6.1.3), que é dada por

$$F_{|Z|}(z; \delta, \lambda) = \int_{-\infty}^{\infty} \left\{ \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \right\} \phi(y) dy,$$

com $|Z| = \max_i |Z_i|$.

Chamando o integrando dessa expressão de $g(y; z, \delta, \lambda)$, tem-se

$$g(y; z, \delta, \lambda) = \left\{ \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \right\} \phi(y)$$

que está representada na Figura 13 com os parâmetros $\delta = (0,2; 0,3; 0,5)$, $z = 1,35$ e $\lambda = (0,2; 0,5; 0,5)$. Pode-se observar que a função é semelhante a da Figura 12, porém simétrica, com um deslocamento da centralidade dada pelo vetor δ .

Novamente a integral original será transformada numa integral com inter-valo $[-1,1]$ e avaliada por uma quadratura de Gauss-Legendre. Será realizada uma mudança da variável y para a variável t , do tipo obtida em (4.1.8), que resulta em

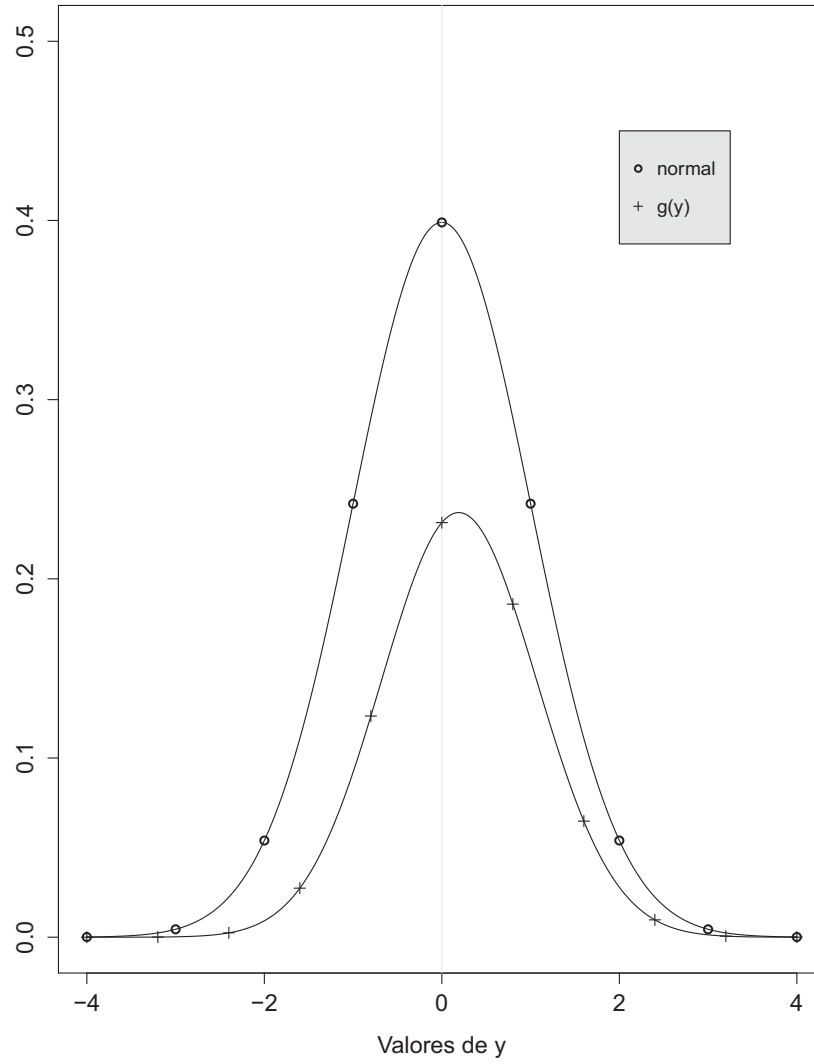


Figura 13 Representação gráfica da função $g(y; z, \delta, \lambda) = g(y; 1,35, (0,2; 0,3; 0,5), (0,2; 0,5; 0,5))$, integrando da expressão (6.1.3), comparando-a com a função densidade da normal $\phi(y)$

$$\int_{-\infty}^{\infty} g(y; z, \delta, \lambda) dy \approx \sum_{j=1}^n w_j \left\{ \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) + z - \delta_i}{\sqrt{1-\lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) - z - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \right\} \phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2},$$

em que cada t_j é um nó da quadratura Gauss-Legendre, w_j é o seu respectivo peso e sendo $z = |d|$ o valor do quantil calculado em cada contraste.

6.3.3 Função para a densidade do teste de Dunnett unilateral para dados não balanceados com infinitos graus de liberdade e parâmetro de não-centralidade δ , denominada *dNDUD*

Nesta função *dNDUD*, os parâmetros de entrada são o quantil d , que será o mesmo do quantil z , o vetor λ de correlações, o vetor de não-centralidade δ e o número de pontos n a ser utilizado na quadratura, que por padrão é usado $n = 32$.

O valor fornecido pela função é o da densidade de probabilidade no ponto z . Neste caso, deseja-se resolver a função densidade do máximo da normal multivariada não-central com parâmetro de não-centralidade $\delta = \mu$ obtida em (6.1.2), que é dada por

$$f_Z(z; \delta, \lambda) = \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \frac{1}{\sqrt{1 - \lambda_i^2}} \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k y + z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) dy,$$

com $Z = \max_i(Z_i)$.

Novamente a integral original será transformada numa integral com intervalo $[-1, 1]$ e avaliada por uma quadratura de Gauss-Legendre. Será realizada uma mudança da variável y para a variável t , do tipo obtida em (4.1.8), que resulta em

$$\begin{aligned} \int_{-\infty}^{\infty} g(y; z, \delta, \lambda) dy &\approx \sum_{j=1}^n w_j \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i \left(\frac{t_j}{1 - t_j^2} \right) + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \frac{1}{\sqrt{1 - \lambda_i^2}} \right. \\ &\quad \times \left. \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k \left(\frac{t_j}{1 - t_j^2} \right) + z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi \left(\frac{t_j}{1 - t_j^2} \right) \frac{1 + t_j^2}{(1 - t_j^2)^2}, \end{aligned}$$

em que cada t_j é um nó da quadratura Gauss-Legendre, w_j é o seu respectivo peso e sendo $z = d$ o valor do quantil calculado em cada contraste.

6.3.4 Função para a densidade do teste de Dunnett bilateral para dados não balanceados com infinitos graus de liberdade e parâmetro de não-centralidade δ , denominada *dNDBD*

Nesta função *dNDBD*, os parâmetros de entrada são o módulo do quantil $|d|$, que será o mesmo módulo do quantil $|z|$, o vetor λ de correlações, o vetor de não-centralidade δ e o número de pontos n a ser utilizado na quadratura, que por padrão é usado $n = 32$. O valor fornecido pela função é o da densidade de probabilidade no ponto z .

Neste caso, deseja-se resolver a função densidade do máximo do módulo da normal multivariada não-central com parâmetro de não-centralidade $\delta = \mu$ obtida em (6.1.4), que é dada por

$$f_{|Z|}(z; \delta, \lambda) = \int_{-\infty}^{\infty} \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i y + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) + \phi \left(\frac{\lambda_i y - z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \left(\frac{1}{\sqrt{1 - \lambda_i^2}} \right) \right) \right. \\ \left. \times \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) dy.$$

com $|Z| = \max_i |Z_i|$.

Novamente a integral original será transformada numa integral com intervalo $[-1, 1]$ e avaliada por uma quadratura de Gauss-Legendre. Será realizada uma mudança da variável y para a variável t , do tipo obtida em (4.1.8), que resulta em:

$$\int_{-\infty}^{\infty} g(y; z, \delta, \lambda) dy \approx \sum_{j=1}^n w_j \phi \left(\frac{t_j}{1 - t_j^2} \right) \frac{1 + t_j^2}{(1 - t_j^2)^2} \sum_{i=1}^r \left\{ \left(\left[\phi \left(\frac{\lambda_i \left(\frac{t_j}{1 - t_j^2} \right) + z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right. \right. \right. \\ \left. \left. + \phi \left(\frac{\lambda_i \left(\frac{t_j}{1 - t_j^2} \right) - z - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \frac{1}{\sqrt{1 - \lambda_i^2}} \prod_{\substack{k=1 \\ k \neq i}}^r \right. \\ \left. \left[\Phi \left(\frac{\lambda_k \left(\frac{t_j}{1 - t_j^2} \right) + z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k \left(\frac{t_j}{1 - t_j^2} \right) - z - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\}.$$

Cada t_j é um nó da quadratura Gauss-Legendre, w_j é o seu respectivo peso e sendo $z = |d|$ o valor do quantil calculado em cada contraste.

6.3.5 Função para a inversa da função de distribuição do teste de Dunnett unilateral para dados não balanceados com infinitos graus de liberdade e parâmetro de não-centralidade δ , denominada *qNDUD*

Nesta função *qNDUD*, os parâmetros de entrada são a probabilidade acumulada $1 - \alpha$, o vetor λ de correlações, o vetor de não-centralidade δ e o número de pontos n a ser utilizado na quadratura, que por padrão é usado $n = 32$. O valor fornecido pela função é o do quantil da distribuição.

Neste caso, deseja-se resolver a inversa da função de distribuição do máximo da normal multivariada não-central com parâmetro de não-centralidade $\delta = \mu$, ou seja, obter o quantil $d = z$ para o qual a expressão (6.1.1) se anula, que é dada por

$$F_Z(z; \delta, \lambda) - (1 - \alpha) = 0$$

$$\int_{-\infty}^{\infty} \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i y + d - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \phi(y) dy - (1 - \alpha) = 0,$$

com $Z = \max_i Z_i$.

Neste caso é necessário utilizar um método numérico para encontrar a raiz da equação. Como estão disponíveis a função distribuição e densidade do máximo da normal multivariada não-central, representadas por $F_Z(z; \delta, \lambda)$ dada em (6.1.1) e $f_Z(z; \delta, \lambda)$ dada em (6.1.2), respectivamente, pode-se utilizar o método de Newton-Raphson. A equação (4.2.2) é a forma geral para iterativamente obter a aproximação da raiz da equação.

Tomando como ponto inicial $x_0 = z_{1-(1-\alpha)^{1/r}}$, sendo $f(\cdot) = F_Z(z; \delta, \lambda) - (1 - \alpha)$ e $f'(\cdot) = f_Z(z; \delta, \lambda)$, tem-se

$$x_{n+1} = x_n - \frac{F_Z(z; \delta, \lambda) - (1 - \alpha)}{f_Z(z; \delta, \lambda)},$$

e fixando uma tolerância para os valores obtidos em iterações seguintes de 1×10^{-13} , com no máximo 5000 iterações, obtém-se o quantil desejado. As integrais são resolvidas pelas funções descritas nas subseções 6.3.1 e 6.3.3.

6.3.6 Função para a inversa da função de distribuição do teste de Dunnett bilateral para dados não balanceados com infinitos graus de liberdade e parâmetro de não-centralidade δ , denominada $qNDBD$

Nesta função $qNDBD$, os parâmetros de entrada são a probabilidade acumulada $1 - \alpha$, o vetor λ de correlações, o vetor de não-centralidade δ e o número de pontos n a ser utilizado na quadratura, que por padrão é usado $n = 32$. O valor fornecido pela função é o do quantil da distribuição.

Neste caso, deseja-se resolver a inversa da função de distribuição do máximo do módulo da normal multivariada não-central com parâmetro de não-centralidade $\delta = \mu$, ou seja, obter o quantil $|d| = z$ para o qual a expressão (6.1.3) se anula, que é dada por

$$F_{|Z|}(z; \delta, \lambda) - (1 - \alpha) = 0$$

$$\int_{-\infty}^{\infty} \left\{ \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + |d| - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - |d| - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \right\} \phi(y) dy - (1 - \alpha) = 0,$$

com $|Z| = \max_i |Z_i|$.

Novamente é necessário utilizar um método numérico para encontrar a raiz da equação. Como estão disponíveis a função distribuição e densidade do máximo da normal multivariada não-central, representadas por $F_{|Z|}(z; \delta, \lambda)$ dada em (6.1.3) e $f_{|Z|}(z; \delta, \lambda)$ dada em (6.1.4), respectivamente, pode-se utilizar o método de Newton-Raphson.

Tomando como ponto inicial $x_0 = z_{1-(1-\alpha)^{1/r}}$, sendo $f(\cdot) = F_{|Z|}(z; \delta, \lambda) - (1 - \alpha)$ e $f'(\cdot) = f_{|Z|}(z; \delta, \lambda)$, tem-se

$$x_{n+1} = x_n - \frac{F_{|Z|}(z; \delta, \lambda) - (1 - \alpha)}{f_{|Z|}(z; \delta, \lambda)},$$

e fixando uma tolerância para os valores obtidos em iterações seguintes de 1×10^{-13} , com no máximo 5000 iterações, obtém-se o quantil desejado. As integrais são resolvidas pelas funções descritas nas subseções 6.3.2 e 6.3.4.

6.3.7 Função para a distribuição do teste de Dunnett unilateral para dados não balanceados com finitos graus de liberdade e parâmetro de não-centralidade δ , denominada *pNDUDF*

Nesta função *pNDUDF*, os parâmetros de entrada são o quantil d , o vetor λ de correlações, os graus de liberdade ν da distribuição da variância amostral residual, o vetor de não-centralidade δ e o número de pontos n a ser utilizado na quadratura, que por padrão é usado $n = 32$. O valor fornecido pela função é o da probabilidade acumulada $1 - \alpha$.

Neste caso de graus de liberdade finitos deseja-se resolver a função de distribuição do máximo da t multivariada não-central com parâmetro de não-centralidade $\delta = \mu$ obtida em (6.2.4), que é dada por

$$F_D(d; \delta, \lambda, \nu) = \int_0^\infty \left[\int_{-\infty}^\infty \prod_{i=1}^r \Phi \left(\frac{\lambda_i y + sd - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \phi(y) dy \right] f(s; \nu) ds,$$

com $D = \max_i T_i$ e $f(s; \nu)$ dada pela equação (5.1.4), ou seja

$$f(s; \nu) = \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2}.$$

Chamando o integrando dessa expressão de $g(y; s, d, \delta, \lambda, \nu)$, tem-se

$$g(y; s, d, \delta, \lambda, \nu) = \prod_{i=1}^r \Phi \left(\frac{\lambda_i y + sd - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \phi(y) f(s; \nu)$$

que está representada na Figura 14 com os parâmetros $\delta = (0,2; 0,3; 0,5)$, $d = 1,35$, $\lambda = (0,2; 0,5; 0,5)$ e $\nu = 120$ para os valores de $s = (1; 1,05; 1,1)$.

A integral interna original será transformada numa integral com intervalo $[-1,1]$ e avaliada por uma quadratura de Gauss-Legendre. Será realizada uma mudança da variável y para a variável t , do tipo apresentada em (4.1.8), ou seja

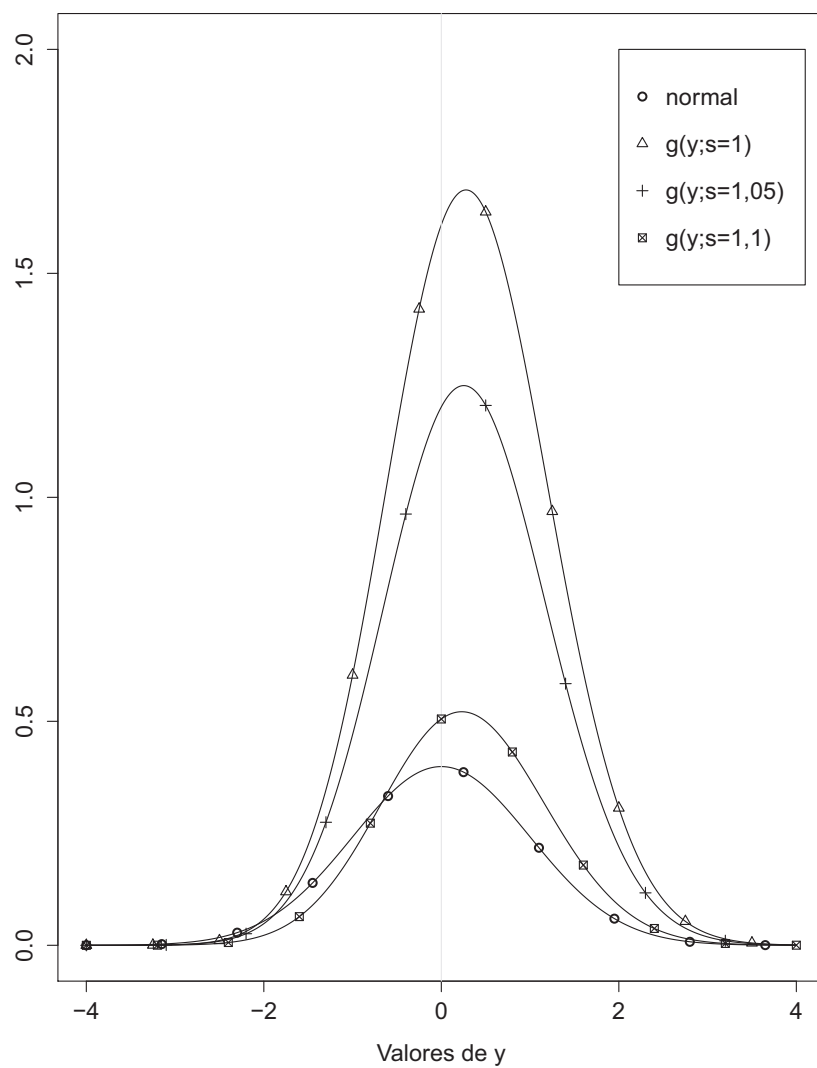


Figura 14 Representação gráfica da função $g(y;s,d,\delta,\lambda,v)$, integrando da expressão (6.2.4), com os parâmetros $d = 1,35$, $\delta = (0,2; 0,3; 0,5)$, $\lambda = (0,2; 0,5; 0,5)$, $v = 120$ para os valores de $s = (1; 1,05; 1,1)$, comparando-a com a função densidade da normal $\phi(y)$

$$\int_{-\infty}^{\infty} g(y; s, d, \delta, \lambda, \nu) dy = \int_{-1}^1 \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i \left(\frac{t}{1-t^2} \right) + sd - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \phi \left(\frac{t}{1-t^2} \right) \frac{1+t^2}{(1-t^2)^2} dt.$$

Assim,

$$\int_{-\infty}^{\infty} g(y; s, d, \delta, \lambda, \nu) dy \approx \sum_{j=1}^n w_j \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) + sd - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2},$$

em que cada t_j é um nó da quadratura Gauss-Legendre e w_j é o seu respectivo peso.

Para resolver a integral externa optou-se por não utilizar a quadratura Gauss-Laguerre mas dividir a integral original na soma de duas integrais, uma com intervalo $[0,1]$ e a outra com intervalo $[1,\infty]$. Ambas as integrais sofrerão mudanças de variáveis para ser resolvida no intervalo $[-1,1]$ por uma quadratura de Gauss-Legendre. Assim, inicialmente tem-se

$$F_D(d; \delta, \lambda, \nu) = \int_0^{\infty} \left[\int_{-\infty}^{\infty} g(y; s, d, \delta, \lambda, \nu) dy \right] \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2} ds,$$

em que a função a ser integrada é

$$h(s; d, \delta, \lambda, \nu) = \left[\int_{-\infty}^{\infty} g(y; s, d, \delta, \lambda, \nu) dy \right] \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2}.$$

Fazendo a integral externa como a soma de duas integrais tem-se

$$F_D(d; \delta, \lambda, \nu) = \int_0^1 h(s; d, \delta, \lambda, \nu) ds + \int_1^{\infty} h(s; d, \delta, \lambda, \nu) ds,$$

Para resolver a primeira integral será realizada uma mudança da variável s para a variável u , levando o intervalo finito $[0,1]$ para $[-1,1]$, do tipo apresentada em (4.1.5), ou seja

$$\int_0^1 h(s; d, \delta, \lambda, \nu) ds = \frac{1}{2} \int_{-1}^1 h \left(\frac{u+1}{2} \right) du.$$

Chamando

$$A(s; d, \delta, \lambda, \nu) = \int_{-\infty}^{\infty} g(y; s, d, \delta, \lambda, \nu) dy$$

$$\Rightarrow A(s; d, \delta, \lambda, \nu) \approx \sum_{j=1}^n w_j \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) + sd - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2},$$

tem-se que

$$\begin{aligned} \int_0^1 h(s) ds &\cong \int_0^1 A(s; d, \delta, \lambda, \nu) \frac{\nu^{\nu/2} s^{\nu-1} e^{-\nu s^2/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} ds \\ &\cong \frac{1}{2} \int_{-1}^1 A \left(\frac{u+1}{2} \right) \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} \left(\frac{u+1}{2} \right)^{\nu-1} e^{-\frac{\nu}{2} \left(\frac{u+1}{2} \right)^2} du \\ &\cong \frac{\nu^{\nu/2}}{2^{(\nu/2)} \Gamma(\frac{\nu}{2})} \int_{-1}^1 A \left(\frac{u+1}{2} \right) \left(\frac{u+1}{2} \right)^{\nu-1} e^{-\frac{\nu}{2} \left(\frac{u+1}{2} \right)^2} du \\ &\cong \frac{\nu^{\nu/2}}{2^{(\nu/2)} \Gamma(\frac{\nu}{2})} \left\{ \sum_{\ell=1}^n w_{\ell} \left(\frac{u_{\ell}+1}{2} \right)^{\nu-1} e^{-\frac{\nu}{2} \left(\frac{u_{\ell}+1}{2} \right)^2} A \left(\frac{u_{\ell}+1}{2} \right) \right\} \end{aligned}$$

$$\begin{aligned} \int_0^1 h(s) ds &\cong \frac{\nu^{\nu/2}}{2^{(\nu/2)} \Gamma(\frac{\nu}{2})} \left\{ \sum_{\ell=1}^n w_{\ell} \left(\frac{u_{\ell}+1}{2} \right)^{\nu-1} e^{-\frac{\nu}{2} \left(\frac{u_{\ell}+1}{2} \right)^2} \left\{ \sum_{j=1}^n w_j \right. \right. \\ &\quad \times \left. \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) + d \left(\frac{u_{\ell}+1}{2} \right) - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2} \right\} \right\} \end{aligned}$$

em que cada t_j e u_{ℓ} é um nó da quadratura Gauss-Legendre e w_j e w_{ℓ} os seus respectivos pesos.

Na segunda integral foi realizada duas mudanças de variáveis:

i) de s para a variável u , do tipo apresentada em (4.1.6), que é dada por

$$\int_1^{\infty} h(s; d, \delta, \lambda, \nu) ds = \int_0^{e^{-1}} \frac{h(-\ln(u))}{u} du,$$

ou seja,

$$\begin{aligned}
\int_1^\infty h(s)ds &= \frac{\nu^{\nu/2}}{2(\nu/2)-1\Gamma(\frac{\nu}{2})} \int_1^\infty A(s; d, \delta, \lambda, \nu) s^{\nu-1} e^{-\frac{\nu}{2}s^2} ds \\
&= \frac{\nu^{\nu/2}}{2(\nu/2)-1\Gamma(\frac{\nu}{2})} \int_0^{e^{-1}} \frac{1}{u} A(-\ln(u); d, \delta, \lambda, \nu) [-\ln(u)]^{\nu-1} e^{-\frac{\nu}{2}[\ln(u)]^2} du \\
&\cong \frac{\nu^{\nu/2}}{2(\nu/2)-1\Gamma(\frac{\nu}{2})} \int_0^{e^{-1}} \frac{1}{u} [-\ln(u)]^{\nu-1} e^{-\frac{\nu}{2}[\ln(u)]^2} \left\{ \sum_{j=1}^n w_j \right. \\
&\quad \times \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) - d(\ln(u)) - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2} \Big\} du;
\end{aligned}$$

ii) e depois de u para a variável m , com $f(u) = \frac{h(-\ln(u))}{u}$, do tipo apresentada em (4.1.7), que é dada por

$$\int_0^{e^{-1}} f(u)du = \frac{e^{-1}}{2} \int_{-1}^1 f \left(e^{-1} \frac{m+1}{2} \right) dm,$$

ou seja,

$$\begin{aligned}
\int_0^{e^{-1}} f(u)du &\cong \frac{e^{-1}\nu^{\nu/2}}{2(\nu/2)\Gamma(\frac{\nu}{2})} \int_{-1}^1 \frac{1}{(e^{-1}\frac{m+1}{2})} \left[-\ln \left(e^{-1} \frac{m+1}{2} \right) \right]^{\nu-1} \\
&\quad \times e^{-\frac{\nu}{2}[\ln(e^{-1}\frac{m+1}{2})]^2} \left\{ \sum_{j=1}^n w_j \left[\phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2} \right] \right. \\
&\quad \times \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) - d \left(\ln \left(e^{-1} \frac{m+1}{2} \right) \right) - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \Big\} dm.
\end{aligned}$$

Assim,

$$\begin{aligned} \int_1^\infty h(s)ds &\cong \frac{e^{-1}\nu^{\nu/2}}{2(\nu/2)\Gamma(\frac{\nu}{2})} \left\{ \sum_{\ell=1}^n w_\ell \left[\frac{e^{-\frac{\nu}{2} \left[\ln \left(e^{-1} \frac{m_\ell+1}{2} \right) \right]^2}}{\left(e^{-1} \frac{m_\ell+1}{2} \right)} \right] \right. \\ &\quad \times \left[-\ln \left(e^{-1} \frac{m_\ell+1}{2} \right) \right]^{\nu-1} \left\{ \sum_{j=1}^n w_j \left[\phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2} \right] \right. \\ &\quad \times \left. \left. \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) - d \left(\ln \left(e^{-1} \frac{m_\ell+1}{2} \right) \right) - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \right\} \right\} \end{aligned}$$

em que cada t_j e m_ℓ é um nó da quadratura Gauss-Legendre e w_j e w_ℓ os seus respectivos pesos.

Assim, todas as integrais envolvidas na função de distribuição são resolvidas por quadraturas de Gauss-Legendre.

6.3.8 Função para a distribuição do teste de Dunnett bilateral para dados não balanceados com finitos graus de liberdade e parâmetro de não-centralidade δ , denominada *pNDBDF*

Nesta função *pNDBDF*, os parâmetros de entrada são o quantil d , o vetor λ de correlações, os graus de liberdade ν da distribuição da variância amostral residual, o vetor de não-centralidade δ e o número de pontos n a ser utilizado na quadratura, que por padrão é usado $n = 32$. O valor fornecido pela função é o da probabilidade acumulada $1 - \alpha$.

Neste caso de graus de liberdade finitos deseja-se resolver a função de distribuição do máximo do módulo da t multivariada não-central com parâmetro de não-centralidade $\delta = \mu$ obtida em (6.2.6), que é dada por

$$\begin{aligned} F_{|D|}(|d|; \delta, \lambda, \nu) &= \int_0^\infty \left\{ \int_{-\infty}^\infty \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + |d|s - \delta_i}{\sqrt{1-\lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - |d|s - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \right. \\ &\quad \times \phi(y) dy \left. \right\} f(s; \nu) ds, \end{aligned}$$

com $|D| = \max_i |T_i|$ e $f(s; \nu)$ dada pela equação (5.1.4), ou seja

$$f(s; \nu) = \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2}.$$

Denominando o integrando dessa expressão de $g(y; s, d, \delta, \lambda, \nu)$, tem-se

$$g(y; s, d, \delta, \lambda, \nu) = \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \phi(y) f(s; \nu)$$

que está representada na Figura 15 com os parâmetros $d = 1,35$, $\delta = (0,2; 0,3; 0,5)$, $\lambda = (0,2; 0,5; 0,5)$, $\nu = 120$ para os valores de $s = (1; 1,05; 1,1)$.

A integral interna original será transformada numa integral com intervalo $[-1,1]$ e avaliada por uma quadratura de Gauss-Legendre. Será realizada uma mudança da variável y para a variável t , do tipo obtida em (4.1.8), ou seja

$$\begin{aligned} \int_{-\infty}^{\infty} g(y) dy &= \int_{-1}^1 \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i \left(\frac{t}{1-t^2} \right) + |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i \left(\frac{t}{1-t^2} \right) - |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \\ &\quad \times \phi \left(\frac{t}{1-t^2} \right) \frac{1+t^2}{(1-t^2)^2} dt, \end{aligned}$$

com $g(y) = g(y; s, d, \delta, \lambda, \nu)$.

Assim

$$\begin{aligned} \int_{-\infty}^{\infty} g(y) dy &\approx \sum_{j=1}^n w_j \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) + |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) - |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \\ &\quad \times \phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2}, \end{aligned}$$

em que cada t_j é um nó da quadratura Gauss-Legendre e w_j é o seu respectivo peso.

Do mesmo modo que na seção anterior, para resolver a integral externa optou-se por não utilizar a quadratura Gauss-Laguerre mas dividir a integral original na soma de duas integrais, uma com intervalo $[0,1]$ e a outra com intervalo $[1,\infty]$. Ambas as integrais sofrerão mudanças de variáveis para ser resolvida no intervalo $[-1,1]$ por uma quadratura de Gauss-Legendre. Assim, inicialmente tem-

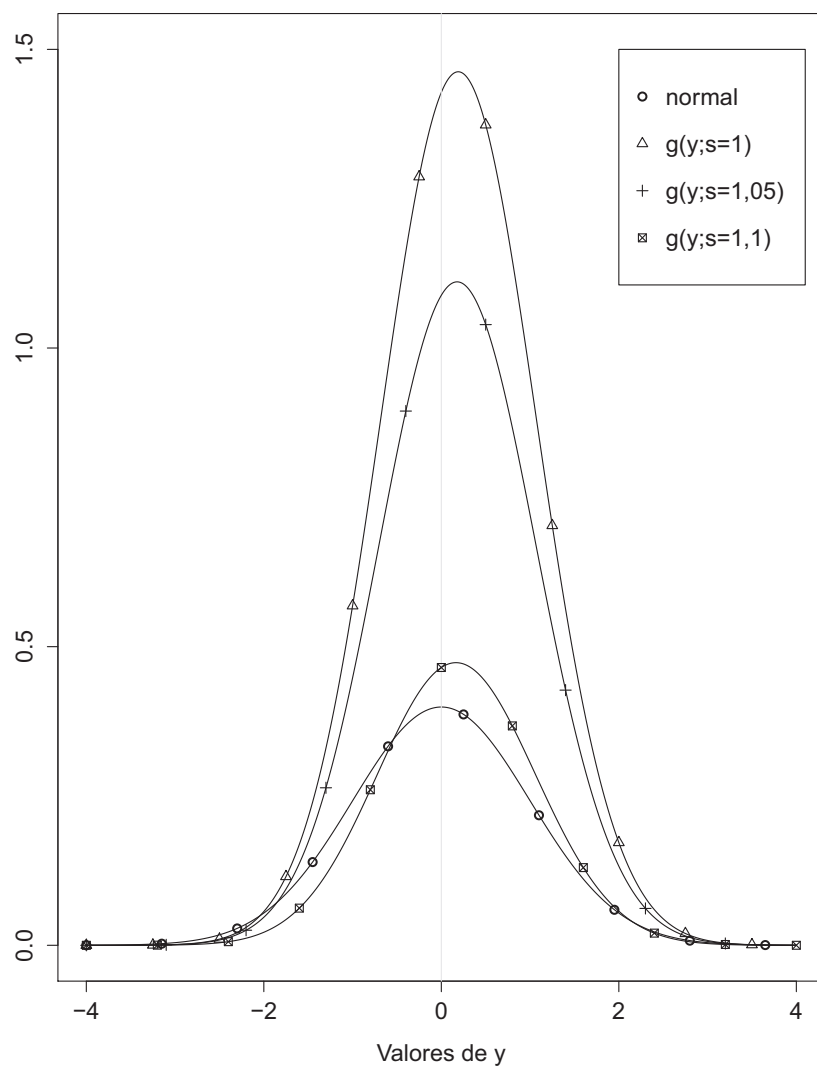


Figura 15 Representação gráfica da função $g(y; s, d, \delta, \lambda, \nu)$, integrando da expressão (6.2.6), com os parâmetros $d = 1,35$, $\delta = (0,2; 0,3; 0,5)$, $\lambda = (0,2; 0,5; 0,5)$, $\nu = 120$ para os valores de $s = (1; 1,05; 1,1)$, comparando-a com a função densidade da normal $\phi(y)$

se

$$F_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_0^\infty \left[\int_{-\infty}^\infty g(y; s, d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) dy \right] \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2} ds,$$

em que a função a ser integrada é

$$h(s; d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \left[\int_{-\infty}^\infty g(y; s, d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) dy \right] \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2}.$$

Fazendo a integral externa como a soma de duas integrais tem-se

$$F_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_0^1 h(s; d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) ds + \int_1^\infty h(s; d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) ds.$$

Na primeira integral foi realizada uma mudança da variável s para a variável u , do tipo obtida em (4.1.5), ou seja

$$\int_0^1 h(s; d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) ds = \frac{1}{2} \int_{-1}^1 h\left(\frac{u+1}{2}\right) du.$$

Sendo

$$A(y; s, d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_{-\infty}^\infty g(y; s, d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) dy$$

$$\begin{aligned} A(y; s, d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) \approx \sum_{j=1}^n w_j \phi\left(\frac{t_j}{1-t_j^2}\right) \frac{1+t_j^2}{(1-t_j^2)^2} \left[\prod_{i=1}^r \Phi\left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2}\right) + |d|.s - \delta_i}{\sqrt{1-\lambda_i^2}}\right) \right. \\ \left. - \Phi\left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2}\right) - |d|.s - \delta_i}{\sqrt{1-\lambda_i^2}}\right) \right], \end{aligned}$$

tem-se que

$$\begin{aligned} \int_0^1 h(s; d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) ds &= \int_0^1 A(y; s, d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) \frac{\nu^{\nu/2} s^{\nu-1} e^{-\nu s^2/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} ds \\ &\cong \frac{1}{2} \int_{-1}^1 A\left(y; \left(\frac{u+1}{2}\right), d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu\right) \frac{\nu^{\nu/2} \left(\frac{u+1}{2}\right)^{\nu-1} e^{-\frac{\nu}{2} \left(\frac{u+1}{2}\right)^2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} du, \end{aligned}$$

ou seja,

$$\begin{aligned} \int_0^1 h(s) ds &\cong \frac{\nu^{\nu/2}}{2^{(\nu/2)} \Gamma(\frac{\nu}{2})} \left\{ \sum_{\ell=1}^n w_{\ell} \left[\left(\frac{u_{\ell}+1}{2} \right)^{\nu-1} e^{-\frac{\nu}{2} \left(\frac{u_{\ell}+1}{2} \right)^2} \right] \right. \\ &\quad \times \left\{ \sum_{j=1}^n w_j \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) + |d| \left(\frac{u_{\ell}+1}{2} \right) - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right. \right. \\ &\quad \left. \left. - \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) - |d| \left(\frac{u_{\ell}+1}{2} \right) - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2} \right\} \right\}, \end{aligned}$$

em que cada t_j e u_{ℓ} é um nó da quadratura Gauss-Legendre e w_j e w_{ℓ} os seus respectivos pesos.

Na segunda integral foi realizada duas mudanças de variáveis:

i) de s para a variável u , do tipo obtida em (4.1.6), que é dada por

$$\int_1^{\infty} h(s; d, \delta, \lambda, \nu) ds = \int_0^{e^{-1}} \frac{h(-\ln(u))}{u} dt;$$

ii) e depois de u para a variável m , com $f(u) = \frac{h(-\ln(u))}{u}$, do tipo obtida em (4.1.7), que é dada por

$$\int_0^{e^{-1}} f(u) du = \frac{e^{-1}}{2} \int_{-1}^1 f \left(e^1 \frac{m+1}{2} \right) dm.$$

Assim, tem-se que

$$\begin{aligned} \int_1^\infty h(s)ds &\cong \frac{e^{-1}\nu^{\nu/2}}{2(\nu/2)\Gamma(\frac{\nu}{2})} \left\{ \sum_{\ell=1}^n w_\ell \left[\frac{e^{-\frac{\nu}{2} \left[\ln \left(e^{-1} \frac{m_\ell+1}{2} \right) \right]^2}}{\left(e^{-1} \frac{m_\ell+1}{2} \right)} \right] \right. \\ &\quad \times \left[-\ln \left(e^{-1} \frac{m_\ell+1}{2} \right) \right]^{\nu-1} \left\{ \sum_{j=1}^n w_j \left[\phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2} \right] \right. \\ &\quad \times \left[\prod_{i=1}^r \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) - |d| \left(\ln \left(e^{-1} \frac{m_\ell+1}{2} \right) \right) - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right. \\ &\quad \left. \left. \left. - \Phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) + |d| \left(\ln \left(e^{-1} \frac{m_\ell+1}{2} \right) \right) - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \right] \right\} \right\} \end{aligned}$$

em que cada t_j e m_ℓ é um nó da quadratura Gauss-Legendre e w_j e w_ℓ os seus respectivos pesos.

Assim, todas as integrais envolvidas na função de distribuição são resolvidas por quadraturas de Gauss-Legendre.

6.3.9 Função para a densidade do teste de Dunnett unilateral para dados não balanceados com finitos graus de liberdade e parâmetro de não-centralidade δ , denominada *dNDUDF*

Nesta função *dNDUDF*, os parâmetros de entrada são o quantil d , o vetor λ de correlações, os graus de liberdade ν da distribuição da variância amostral residual, o vetor de não-centralidade δ e o número de pontos n a ser utilizado na quadratura, que por padrão é usado $n = 32$. O valor fornecido pela função é o da densidade de probabilidade no ponto d . Neste caso, deseja-se resolver a função densidade do máximo da t multivariada não-central com parâmetro de não-centralidade $\delta = \mu$ obtida em (6.2.3), que é dada por

$$f_D(d; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_0^\infty \left[\int_{-\infty}^\infty \sum_{i=1}^p \left\{ \phi \left(\frac{\lambda_i y + sd - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \frac{s}{\sqrt{1 - \lambda_i^2}} \right. \right. \\ \left. \left. \times \left[\prod_{\substack{k=1 \\ k \neq i}}^p \Phi \left(\frac{\lambda_k y + sd - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi(y) dy \right] f(s; \nu) ds,$$

com $D = \max_i(T_i)$.

A integral interna será transformada numa integral com intervalo $[-1, 1]$ e avaliada por uma quadratura de Gauss-Legendre. Será realizada uma mudança da variável y para a variável t , do tipo obtida em (4.1.8), que resulta em

$$\int_{-\infty}^\infty g(y) dy \approx \sum_{j=1}^n w_j \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) + sd - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \frac{s}{\sqrt{1 - \lambda_i^2}} \right. \\ \left. \times \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k \left(\frac{t_j}{1-t_j^2} \right) + sd - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2},$$

com $g(y) = g(y; s, d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu)$ e em que cada t_j é um nó da quadratura Gauss-Legendre e w_j o seu respectivo peso.

Para resolver a integral externa optou-se por não utilizar a quadratura Gauss-Laguerre mas dividir a integral original na soma de duas integrais, uma com intervalo $[0, 1]$ e a outra com intervalo $[1, \infty]$. Ambas as integrais sofrerão mudanças de variáveis para ser resolvida no intervalo $[-1, 1]$ por uma quadratura de Gauss-Legendre. Assim, inicialmente tem-se

$$f_D(d; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_0^\infty \left[\int_{-\infty}^\infty g(y; s, d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) dy \right] \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2} ds,$$

em que a função a ser integrada é

$$h(s; d, \delta, \lambda, \nu) = \left[\int_{-\infty}^{\infty} g(y; s, d, \delta, \lambda, \nu) dy \right] \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2}.$$

Fazendo a integral externa como a soma de duas integrais tem-se

$$f_D(d; \delta, \lambda, \nu) = \int_0^1 h(s; d, \delta, \lambda, \nu) ds + \int_1^{\infty} h(s; d, \delta, \lambda, \nu) ds,$$

Na primeira integral foi realizada uma mudança da variável s para a variável u , do tipo obtida em (4.1.5), ou seja

$$\int_0^1 h(s; d, \delta, \lambda, \nu) ds = \frac{1}{2} \int_{-1}^1 h\left(\frac{u+1}{2}\right) du.$$

Sendo

$$\begin{aligned} A(y; s, d, \delta, \lambda, \nu) &= \int_{-\infty}^{\infty} g(y; s, d, \delta, \lambda, \nu) dy \\ \Rightarrow A(y; s, d, \delta, \lambda, \nu) &\approx \sum_{j=1}^n w_j \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) + sd - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \frac{s}{\sqrt{1-\lambda_i^2}} \right. \\ &\quad \times \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k \left(\frac{t_j}{1-t_j^2} \right) + sd - \delta_k}{\sqrt{1-\lambda_k^2}} \right) \right] \left. \right\} \phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2}, \end{aligned}$$

tem-se que

$$\begin{aligned} \int_0^1 h(s; d, \delta, \lambda, \nu) ds &= \int_0^1 A(y; s, d, \delta, \lambda, \nu) \frac{\nu^{\nu/2} s^{\nu-1} e^{-\nu s^2/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} ds \\ &\cong \frac{1}{2} \int_{-1}^1 A\left(y; \left(\frac{u+1}{2}\right), d, \delta, \lambda, \nu\right) \frac{\nu^{\nu/2} \left(\frac{u+1}{2}\right)^{\nu-1} e^{-\frac{\nu}{2} \left(\frac{u+1}{2}\right)^2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} du, \end{aligned}$$

ou seja,

$$\int_0^1 h(s) ds \cong \frac{\nu^{\nu/2}}{2^{(\nu/2)} \Gamma(\frac{\nu}{2})} \left\{ \sum_{\ell=1}^n w_{\ell} \left[\left(\frac{u_{\ell}+1}{2} \right)^{\nu-1} e^{-\frac{\nu}{2} \left(\frac{u_{\ell}+1}{2} \right)^2} \right] \right\}$$

$$\begin{aligned} & \times \left\{ \sum_{j=1}^n w_j \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) + d \left(\frac{u_\ell+1}{2} \right) - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \frac{\left(\frac{u_\ell+1}{2} \right)}{\sqrt{1-\lambda_i^2}} \right. \right. \\ & \times \left. \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k \left(\frac{t_j}{1-t_j^2} \right) + d \left(\frac{u_\ell+1}{2} \right) - \delta_k}{\sqrt{1-\lambda_k^2}} \right) \right] \right\} \phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2} \left. \right\} \Bigg\}, \end{aligned}$$

com $h(s) = h(s; d, \delta, \lambda, \nu)$ e em que cada t_j e u_ℓ é um nó da quadratura Gauss-Legendre e w_j e w_ℓ os seus respectivos pesos.

Na segunda integral foi realizada duas mudanças de variáveis:

i) de s para a variável u , do tipo obtida em (4.1.6), que é dada por

$$\int_1^\infty h(s; d, \delta, \lambda, \nu) ds = \int_0^{e^{-1}} \frac{h(-\ln(u))}{u} dt;$$

ii) e depois de u para a variável m , com $f(u) = \frac{h(-\ln(u))}{u}$, do tipo obtida em (4.1.7), que é dada por

$$\int_0^{e^{-1}} f(u) du = \frac{e^{-1}}{2} \int_{-1}^1 f \left(e^{-1} \frac{m+1}{2} \right) dm.$$

Assim, obteve-se

$$\begin{aligned} \int_1^\infty h(s) ds & \cong \frac{e^{-1} \nu^{\nu/2}}{2^{(\nu/2)} \Gamma(\frac{\nu}{2})} \left\{ \sum_{\ell=1}^n w_\ell \left[\frac{e^{-\frac{\nu}{2}} [\ln(e^{-1} \frac{m_\ell+1}{2})]^2}{(e^{-1} \frac{m_\ell+1}{2})} \left[-\ln \left(e^{-1} \frac{m_\ell+1}{2} \right) \right]^{\nu-1} \right] \right. \\ & \times \left\{ \sum_{j=1}^n w_j \sum_{i=1}^r \left\{ \phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) - d [\ln(e^{-1} \frac{m_\ell+1}{2})] - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \left[\frac{[-\ln(e^{-1} \frac{m_\ell+1}{2})]}{\sqrt{1-\lambda_i^2}} \right] \right. \right. \\ & \times \left. \left[\prod_{\substack{k=1 \\ k \neq i}}^r \Phi \left(\frac{\lambda_k \left(\frac{t_j}{1-t_j^2} \right) - d [\ln(e^{-1} \frac{m_\ell+1}{2})] - \delta_k}{\sqrt{1-\lambda_k^2}} \right) \right] \right\} \phi \left(\frac{t_j}{1-t_j^2} \right) \frac{1+t_j^2}{(1-t_j^2)^2} \left. \right\} \Bigg\}, \end{aligned}$$

com $h(s) = h(s; d, \delta, \lambda, \nu)$ e em que cada t_j e m_ℓ é um nó da quadratura Gauss-Legendre e w_j e w_ℓ os seus respectivos pesos.

Assim, todas as integrais envolvidas na função densidade são resolvidas por quadraturas de Gauss-Legendre.

6.3.10 Função para a densidade do teste de Dunnett bilateral para dados não balanceados com finitos graus de liberdade e parâmetro de não-centralidade δ , denominada *dNDBDF*

Nesta função *dNDBDF*, os parâmetros de entrada são o quantil d , o vetor λ de correlações, os graus de liberdade ν da distribuição da variância amostral residual, o vetor de não-centralidade δ e o número de pontos n a ser utilizado na quadratura, que por padrão é usado $n = 32$. O valor fornecido pela função é o da densidade de probabilidade no ponto d .

Neste caso, deseja-se resolver a função densidade do máximo do módulo da t multivariada não-central com parâmetro de não-centralidade $\delta = \mu$ obtida em (6.2.5), que é dada por

$$f_{|D|}(|d|; \delta, \lambda, \nu) = \int_0^\infty \left\{ \int_{-\infty}^\infty \sum_{i=1}^r \left\{ \left(\phi \left(\frac{\lambda_i y + |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) + \phi \left(\frac{\lambda_i y - |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right) \right. \right. \\ \times \left(\frac{s}{\sqrt{1 - \lambda_i^2}} \right) \left. \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k y + |d|s - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k y - |d|s - \delta_k}{\sqrt{1 - \lambda_k^2}} \right) \right] \right\} \\ \times \phi(y) dy \Big\} f(s; \nu) ds,$$

com $|D| = \max_i |T_i|$.

A integral interna será transformada numa integral com intervalo $[-1, 1]$ e avaliada por uma quadratura de Gauss-Legendre. Será realizada uma mudança da variável y para a variável t , do tipo obtida em (4.1.8), que resulta em

$$\int_{-\infty}^\infty g(y) dy \approx \sum_{j=1}^n w_j \left[\phi \left(\frac{t_j}{1 - t_j^2} \right) \frac{1 + t_j^2}{(1 - t_j^2)^2} \right] \sum_{i=1}^r \left(\frac{s}{\sqrt{1 - \lambda_i^2}} \right)$$

$$\times \left\{ \left[\phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) + |d|s - \delta_i}{\sqrt{1-\lambda_i^2}} \right) + \phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) - |d|s - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \right. \\ \left. \times \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k \left(\frac{t_j}{1-t_j^2} \right) + |d|s - \delta_k}{\sqrt{1-\lambda_k^2}} \right) - \Phi \left(\frac{\lambda_k \left(\frac{t_j}{1-t_j^2} \right) - |d|s - \delta_k}{\sqrt{1-\lambda_k^2}} \right) \right] \right\}.$$

Cada t_j é um nó da quadratura Gauss-Legendre e w_j o seu respectivo peso e $g(y) = g(y; s, d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu)$.

Para resolver a integral externa optou-se por não utilizar a quadratura Gauss-Laguerre mas dividir a integral original na soma de duas integrais, uma com intervalo $[0,1]$ e a outra com intervalo $[1,\infty]$. Ambas as integrais sofrerão mudanças de variáveis para ser resolvida no intervalo $[-1,1]$ por uma quadratura de Gauss-Legendre. Assim, inicialmente tem-se

$$f_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_0^\infty \left[\int_{-\infty}^\infty g(y) dy \right] \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2} ds,$$

em que a função a ser integrada é

$$h(s; d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \left[\int_{-\infty}^\infty g(y; s, d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) dy \right] \frac{\nu^{\nu/2}}{2^{(\nu/2)-1} \Gamma(\frac{\nu}{2})} s^{\nu-1} e^{-\nu s^2/2}.$$

Fazendo a integral externa como a soma de duas integrais tem-se

$$f_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) = \int_0^1 h(s; d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) ds + \int_1^\infty h(s; d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) ds.$$

Na primeira integral foi realizada uma mudança da variável s para a variável u , do tipo obtida em (4.1.5), ou seja

$$\int_0^1 h(s; d, \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) ds = \frac{1}{2} \int_{-1}^1 h\left(\frac{u+1}{2}\right) du.$$

$$\int_1^\infty h(s)ds \cong \frac{e^{-1}\nu^{\nu/2}}{2^{(\nu/2)}\Gamma(\frac{\nu}{2})}\left\{\sum_{\ell=1}^nw_\ell\frac{e^{-\frac{\nu}{2}}\left[\ln\left(e^{-1}\frac{m_\ell+1}{2}\right)\right]^2}{\left(e^{-1}\frac{m_\ell+1}{2}\right)}\left[-\ln\left(e^{-1}\frac{m_\ell+1}{2}\right)\right]^{\nu-1}\right. \\ \left.\times\left\{\sum_{j=1}^nw_j\left[\phi\left(\frac{t_j}{1-t_j^2}\right)\frac{1+t_j^2}{(1-t_j^2)^2}\right]\sum_{i=1}^r\left\{\left(\frac{[-\ln\left(e^{-1}\frac{m_\ell+1}{2}\right)]}{\sqrt{1-\lambda_i^2}}\right)\right.\right.\right.$$

$$\begin{aligned}
& \times \left[\phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) - |d| \left[\ln \left(e^{-1} \frac{m_\ell+1}{2} \right) \right] - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right. \\
& \left. + \phi \left(\frac{\lambda_i \left(\frac{t_j}{1-t_j^2} \right) + |d| \left[\ln \left(e^{-1} \frac{m_\ell+1}{2} \right) \right] - \delta_i}{\sqrt{1-\lambda_i^2}} \right) \right] \\
& \times \prod_{\substack{k=1 \\ k \neq i}}^r \left[\Phi \left(\frac{\lambda_k \left(\frac{t_j}{1-t_j^2} \right) - |d| \left[\ln \left(e^{-1} \frac{m_\ell+1}{2} \right) \right] - \delta_k}{\sqrt{1-\lambda_k^2}} \right) \right. \\
& \left. - \Phi \left(\frac{\lambda_k \left(\frac{t_j}{1-t_j^2} \right) + |d| \left[\ln \left(e^{-1} \frac{m_\ell+1}{2} \right) \right] - \delta_k}{\sqrt{1-\lambda_k^2}} \right) \right] \Bigg\} \Bigg\} \Bigg\},
\end{aligned}$$

com $h(s) = h(s; d, \delta, \lambda, \nu)$ e em que cada t_j e m_ℓ é um nó da quadratura Gauss-Legendre e w_j e w_ℓ os seus respectivos pesos.

Assim, todas as integrais envolvidas na função densidade são resolvidas por quadraturas de Gauss-Legendre.

6.3.11 Função para a inversa da função de distribuição do teste de Dunnett unilateral para dados não balanceados com finitos graus de liberdade e parâmetro de não-centralidade δ , denominada *qNDUDF*

Nesta função *qNDUDF*, os parâmetros de entrada são a probabilidade acumulada $1 - \alpha$, o vetor λ de correlações, os graus de liberdade ν da distribuição da variância amostral residual, o vetor de não-centralidade δ e o número de pontos n a ser utilizado na quadratura, que por padrão é usado $n = 32$. O valor fornecido pela função é o do quantil da distribuição.

Neste caso, deseja-se resolver a inversa da função de distribuição do máximo da t multivariada não-central com parâmetro de não-centralidade $\delta = \mu$, ou seja, obter o quantil d para o qual a expressão (6.2.4) se anula, que é dada por

$$F_D(d; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) - (1 - \alpha) = 0$$

$$\int_0^\infty \left[\int_{-\infty}^\infty \prod_{i=1}^r \Phi \left(\frac{\lambda_i y + sd - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \phi(y) dy \right] f(s; \nu) ds - (1 - \alpha) = 0,$$

com $D = \max_i T_i$.

Neste caso é necessário utilizar um método numérico para encontrar a raiz da equação. Como estão disponíveis a função distribuição e densidade do máximo da t multivariada não-central, representadas por $F_D(d; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu)$ dada em (6.2.4) e $f_D(d; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu)$ dada em (6.2.3), respectivamente, pode-se utilizar o método de Newton-Raphson.

Tomando como ponto inicial $x_0 = t_{[(1-(1-\alpha)^{1/r}), \nu]}$ tem-se

$$x_{n+1} = x_n - \frac{F_D(d; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) - (1 - \alpha)}{f_D(d; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu)},$$

e fixando uma tolerância para os valores obtidos em iterações seguintes de 1×10^{-13} , com no máximo 5000 iterações, obtém-se o quantil desejado. As integrais são resolvidas pelas funções descritas nas subseções 6.4.7 e 6.4.9.

6.3.12 Função para a inversa da função de distribuição do teste de Dunnett bilateral para dados não balanceados com finitos graus de liberdade e parâmetro de não-centralidade δ , denominada $qNDBDF$

Nesta função $qNDBDF$, os parâmetros de entrada são a probabilidade acumulada $1 - \alpha$, o vetor $\boldsymbol{\lambda}$ de correlações, os graus de liberdade ν da distribuição da variância amostral residual, o vetor de não-centralidade $\boldsymbol{\delta}$ e o número de pontos n a ser utilizado na quadratura, que por padrão é usado $n = 32$. O valor fornecido pela função é o do quantil da distribuição.

Neste caso, deseja-se resolver a inversa da função de distribuição do máximo do módulo da t multivariada não-central com parâmetro de não-centralidade $\delta = \mu$, ou seja, obter o quantil $|d|$ para o qual a expressão (6.2.6) se anula, que é dada por $F_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) - (1 - \alpha) = 0$, ou seja

$$\int_0^\infty \left\{ \int_{-\infty}^\infty \prod_{i=1}^r \left[\Phi \left(\frac{\lambda_i y + |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) - \Phi \left(\frac{\lambda_i y - |d|s - \delta_i}{\sqrt{1 - \lambda_i^2}} \right) \right] \phi(y) dy \right\} \\ \times f(s; \nu) ds - (1 - \alpha) = 0,$$

com $|D| = \max_i |T_i|$.

Novamente é necessário utilizar um método numérico para encontrar a raiz da equação. Como estão disponíveis a função distribuição e densidade do máximo da normal multivariada não-central, representadas por $F_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu)$ dada em (6.2.6) e $f_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu)$ dada em (6.2.5), respectivamente, pode-se utilizar o método de Newton-Raphson.

Tomando como ponto inicial $x_0 = t_{[(1-(1-\alpha/2)^{1/r}, \nu]}$ tem-se

$$x_{n+1} = x_n - \frac{F_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu) - (1 - \alpha)}{f_{|D|}(|d|; \boldsymbol{\delta}, \boldsymbol{\lambda}, \nu)},$$

e fixando uma tolerância para os valores obtidos em iterações seguintes de 1×10^{-13} , com no máximo 5000 iterações, obtém-se o quantil desejado. As integrais são resolvidas pelas funções descritas nas subseções 6.4.8 e 6.4.10.

7 RESULTADOS

Nessa seção buscou-se avaliar o desempenho das funções implementadas no *software R* e comparar as suas acurácias com valores publicados na literatura ou oriundos de pacotes implementados no *sistema R*. Diferentes configurações dos argumentos dessas funções foram utilizadas. Portanto, a influência do número de tratamentos em teste r , das correlações λ_i 's, dos graus de liberdade ν e de parâmetro de não-centralidade δ foi avaliada. Também procurou-se avaliar a influência do número de pontos das quadraturas na precisão dos resultados das probabilidades dos casos unilaterais e bilaterais.

Inicialmente, uma questão que merece ser discutida refere-se à convergência da distribuição t multivariada para a normal na medida em que os graus de liberdade aumentam. Para mostrar isso, na Figura 16 as probabilidades da t para $q = 1,35$, $\lambda = (0.2, 0.5, 0.5)$ e $\delta = 0$, em função dos graus de liberdade ν , foram computadas para a distribuição da estatística do teste de Dunnett unilateral não balanceada, com $n = 200$ pontos na quadratura, por meio da função `pNDUDF`. A linha horizontal tracejada ($p = 0,791$) refere-se a probabilidade acumulada da normal trivariada, considerando o quantil $d = 1,35$ e demais parâmetros anteriormente citados. Ela foi calculada pela função `pNDUD`.

Nesse exemplo, verifica-se que com o aumento dos graus de liberdade a probabilidade da t trivariada se aproxima daquela obtida pela normal trivariada.

Com $\nu = 100$ os valores são praticamente idênticos. Por isso o teste de Dunnett é calculado pela normal multivariada quando os graus de liberdade aumentam para o infinito. Entretanto, para $\nu = 1$, os valores da probabilidade da t trivariada e da normal trivariada divergem bastante. Se for feito um paralelo da convergência da t à normal no caso univariado para esse caso trivariado, constata-se que a aproximação normal da t é bem mais lenta a medida que a dimensão aumenta.

Na Tabela 5 estão apresentados os valores das probabilidades das estatísticas do teste de Dunnett variando-se o número de pontos da quadratura n para infinitos graus de liberdade, a partir das funções $pNDUD$ no teste unilateral e $pNDBD$ no teste bilateral, que utilizam a distribuição normal multivariada. Fixou-se $\lambda = (0,2; 0,5; 0,5)$ e $d = 1,35$, considerando-se o caso central $\delta = (0; 0; 0)$ e não-central com $\delta = (0,5; 0,3; 0,2)$.

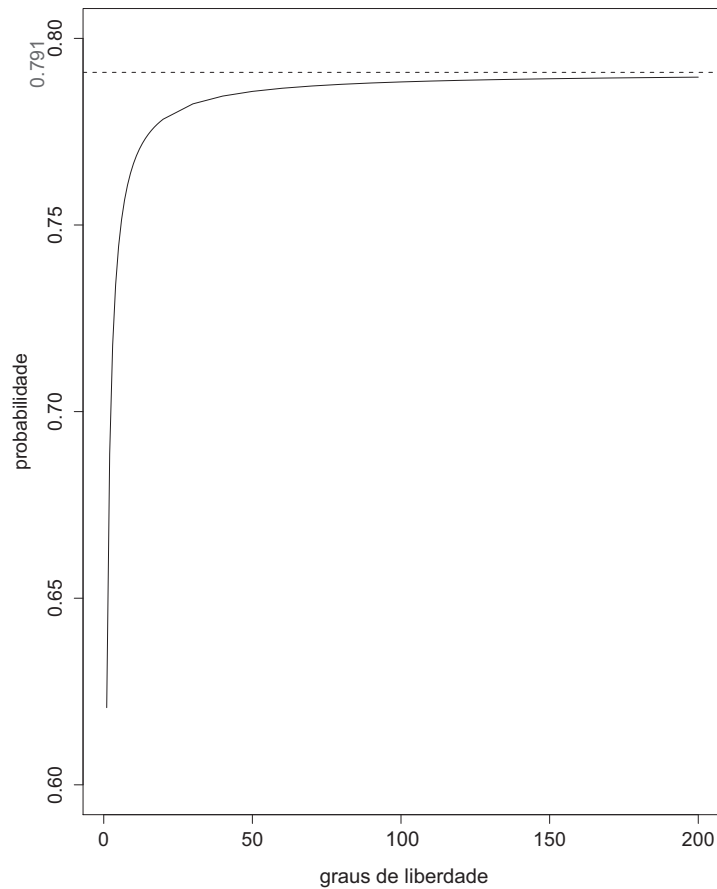


Figura 16 Representação da probabilidade t multivariada, função $pNDUDF$ ($d = 1,35$, $\lambda = (0,2; 0,5; 0,5)$, v , $\delta = 0$, $n = 200$), em função dos graus de liberdade v , em que a linha horizontal tracejada refere-se a probabilidade da normal multivariada nessas mesmas condições

Na Tabela 6 estão apresentados os resultados das probabilidades das estatísticas do teste de Dunnett, variando-se os graus de liberdade ν e o número de pontos da quadratura n , a partir das funções $pNDUDF$ e $pNDBDF$ para o teste unilateral e bilateral, respectivamente, utilizando a distribuição t multivariada. Em todos os casos fixou-se $\lambda = (0,2; 0,5; 0,5)$ e $d = 1,35$, considerando-se um caso central $\delta = (0; 0; 0)$ e os outros não-centrais com $\delta = (0,5; 0,3; 0,2)$.

Comparando-se os valores das probabilidades das Tabelas 5 e 6, em relação a distribuição central ($\delta = (0; 0; 0)$) e não central ($\delta = (0,5; 0,3; 0,2)$), observa-se maior valor das probabilidade para o caso central, tanto para teste unilateral como para bilateral, independentemente dos graus de liberdade (infinitos graus de liberdade na Tabela 5 e $\nu = 5$ na Tabela 6). Isto aconteceu porque neste caso particular o vetor de não centralidade $\delta = (0,5; 0,3; 0,2)$ da distribuição é composto somente de valores positivos, indicando um deslocamento da média no sentido positivo em relação à origem nas três dimensões. Isto acarreta uma diminuição do volume da região que indica a probabilidade acumulada.

Tabela 5 Probabilidade do teste de Dunnett em função do número de pontos da quadratura n , para infinitos graus de liberdade, $d = 1,35$ e $\lambda = (0,2; 0,5; 0,5)$, obtidas pelas funções $pNDUD(d, \lambda, \delta, n)$ para o teste unilateral e $pNDBD(d, \lambda, \delta, n)$ para o teste bilateral

n	$\delta = (0; 0; 0)$	
	$pNDUD$	$pNDBD$
30	0,7909168476601640	0,5898754466961390
31	0,7909078341197382	0,5898768511928386
32	0,7908952033283740	0,5898700158598975
33	0,7908899615613073	0,5898648842399930
34	0,7908927478023080	0,5898651532076940
35	0,7908979619159170	0,5898680509762940

“Tabela 5 , conclusão”

$\delta = (0,5; 0,3; 0,2)$		
n	$pNDUD$	$pNDBD$
30	0,6558814060528640	0,5452551385810563
31	0,6558700726113210	0,5452650491173250
32	0,6558596247713100	0,5452645943605655
33	0,6558570037562660	0,5452589126912320
34	0,6558603117450020	0,5452552142912210
35	0,6558644621804115	0,5452557141044370
50	0,6558638804428980	0,5452582689375716
100	0,6558638458713020	0,5452582316484290
200	0,6558638458713220	0,5452582316484950
500	0,6558638458713220	0,5452582316484945

Quando se comparam os valores das probabilidades da Tabela 6 em relação aos graus de liberdade, para o caso não-central, observa-se que as probabilidades aumentam com o aumento dos graus de liberdade, tanto para o teste unilateral como para o bilateral. Os resultados dessas probabilidades tendem para os valores obtidos quando os graus de liberdade são infinitos.

De modo geral, pelas Tabelas 5 e 6 verifica-se que a precisão dos valores obtidos é mantida no mínimo até a quarta casa decimal a partir de $n = 32$ pontos na quadratura. O erro de precisão nestes casos é no máximo de 1×10^{-5} . Assim, adotou-se na maioria das funções R implementadas $n = 32$ pontos na quadratura como padrão. Tomou-se como exemplo a distribuição não-central, com $\nu \rightarrow \infty$ na Tabela 5 e $\nu = 5$ na Tabela 6, para mostrar o aumento da precisão do valor obtido para a probabilidade com o aumento dos pontos na quadratura n . Observa-se que com $n = 50$ pontos da quadratura a precisão do valor da probabilidade obtida é da ordem da sétima casa decimal, nos dois casos, tanto para teste unilateral como bilateral. Com $n = 100$ pontos na quadratura a precisão do valor é da ordem da décima terceira casa decimal. A diferença de precisão obtida com $n = 200$ e $n = 500$ pontos para a quadratura é de 1×10^{-16} .

Portanto, uma grande vantagem do uso das quadraturas gaussianas na resolução das integrais é que a precisão pode ser aumentada aumentando-se o número de pontos nas quadraturas.

Na Tabela 7 estão valores de probabilidades do teste de Dunnett bilateral balanceado, para o caso central, obtidas pela função *pNDBD* e *pmvt* do pacote “*mvtnorm*” do *sistema R* e em algoritmos da literatura obtidos no artigo de Wang e Kennedy (1992). São fixados $\delta = 0$, $d = 2$, infinitos graus e liberdade e $n = 100$ pontos na quadratura e tomam-se diferente número de comparações e diferentes valores para a correlação: $\lambda = (0,7; 0,7; 0,7; 0,7)$ com $r = 4$ e $\lambda = (0,95; 0,95; 0,95)$ com $r = 3$. Observa-se que os valores das probabilidades obtidos pela função *pNDBD* tem precisão comparável com os resultados obtidos na literatura, enquanto que o valor obtido pela função *pmvt* tem a menor precisão dentre os quatro algoritmos comparados. Os erros apresentados na referida tabela foram obtidos das seguintes formas: para a função *pNDBD* foi calculado pelo módulo da diferença entre o valor obtido com $n = 99$ e $n = 100$ pontos na quadratura; para a função *pmvt* foi fornecido pelo programa; para os algoritmos de Schervish e de Wang & Kennedy foram obtidos na literatura.

Tabela 6 Probabilidade do teste de Dunnett em função do número de pontos da quadratura n e dos graus de liberdade v , para $\lambda = (0,2; 0,5; 0,5)$, $d = 1,35$, obtidas pelas funções *pNDUDF*(d, λ, v, δ, n) para o teste unilateral e *pNDBDF*(d, λ, v, δ, n) para o teste bilateral

$\delta = (0; 0; 0)$			
n	v	<i>pNDUDF</i>	<i>pNDBDF</i>
30	5	0,7443243514593456	0,5219718080549842
31		0,7443176099866940	0,5219777466641453
32		0,7443091458398384	0,5219793047649580
33		0,7443054727012124	0,5219773649612096
34		0,7443070288944140	0,5219751828497650

“Tabela 6, conclusão”

$\delta = (0,5; 0,3; 0,2)$			
n	ν	$pNDUDF$	$pNDBDF$
30	5	0,6124415105454004	0,4861078361033020
31		0,6124327652042770	0,4861142370493203
32		0,6124231636015610	0,4861176011546550
33		0,6124197163894910	0,4861167065346608
34		0,6124220180623233	0,4861143089720665
35		0,6124259283889940	0,4861130107715058
50		0,6124259491098984	0,4861142692127336
100		0,6124259087971170	0,4861142665012473
200		0,6124259087971380	0,4861142665012367
500		0,6124259087971380	0,4861142665012370
30	40	0,6499173751914193	0,5364410995665176
31		0,6499065090081490	0,5364506820083180
32		0,6498959612791815	0,5364512613093234
33		0,6498930220199706	0,5364465215357960
34		0,6498961792997022	0,5364428891854060
35		0,6499004050255507	0,5364428796555265
500		0,6498998915410090	0,5364452679365250
30	120	0,6538738889036770	0,5422694273521110
31		0,6538627063594710	0,5422694273521134
32		0,6538522069398794	0,5422693499503548
33		0,6538494711055880	0,5422640018325940
34		0,6538527337927306	0,5422603159842188
35		0,6538569175623976	0,5422606277944660
500		0,6538563331101050	0,5422631070586976
30	480	0,6553776578000187	0,5445006878399210
31		0,6553663571032960	0,5445105719226925
32		0,6553559027809526	0,5445102200659050
33		0,6553532543431540	0,5445046274985954
34		0,6553565502854934	0,5445009299884280
35		0,6553607093469049	0,5445013804128660
500		0,6553601009457210	0,5445038891977925

Tabela 7 Probabilidades do teste de Dunnett bilateral obtidas pela função $pNDBD(d, \lambda, \delta, n)$ e $pmvt$ do pacote “*mvtnorm*” do sistema R e em algoritmos da literatura, para $v \rightarrow \infty$, $\delta = 0$, $d = 2$ e $n = 100$ pontos na quadratura

programa	probabilidade	erro absoluto
$\lambda = (0,7; 0,7; 0,7; 0,7)$		
pNDBD	0,880221827686591	8.0984e-12
pmvt	0,880015360026146	0,000699
Schervish	0,880221922100000	0,000001
Wang & Kennedy	0,880221803700000	0,000001
$\lambda = (0,95; 0,95; 0,95)$		
pNDBD	0,9328458417978010	2.3419e-07
pmvt	0,9328138039268582	0,000562
Schervish	0,9328451942000000	0,000001
Wang & Kennedy	0,9328452295000000	0,000001

Na Tabela 8 estão apresentados valores das probabilidade do teste de Dunnett unilateral e bilateral, pelas funções propostas com $n = 32$ pontos na quadratura e pela função $pmvt$ do pacote “*mvtnorm*”, para o caso em que são comparados $r = 18$ tratamentos com o controle e o quantil $d = 1,35$. Pode-se citar como exemplo de uma situação experimental com graus de liberdade do resíduo $v = 5$ um experimento em blocos aumentados, em que somente a testemunha é repetida, por exemplo com 6 blocos e 2 testemunhas. Nesta tabela, comparando-se os valores das probabilidades para o caso balanceado com $\rho = 0,5$ e não-balanceado, observa-se que a probabilidade é maior para a situação balanceada, para o mesmo valor de graus de liberdade, tanto na distribuição central como na não-central.

Pela Tabela 8 também pode-se comparar os valores das probabilidades obtidas pelas funções propostas em relação às obtidas pela função $pmvt$ do pacote “*mvtnorm*”. Para a função $pmvt$ a rotina fornece o erro, que em geral é da ordem da quarta casa decimal. Observa-se que os valores das probabilidades

obtidas pelas funções propostas coincidem com os valores fornecidos pela função *pmvt* até a quarta decimal. No entanto, as funções propostas fornecem probabilidades de acurácia igual ou maior do que as fornecidas pela função *pmvt* do pacote “*mvtnorm*”, já que o seu limite de precisão pode aumentar conforme o aumento do número de pontos n tomados para a quadratura. Nessa tabela todos os cálculos foram realizados com $n = 32$.

Tabela 8 Probabilidade do teste de Dunnett obtidas pelas funções $pNDUDF$ para o teste unilateral e p bilateral e pela função $pmvt$ do pacote “*mvtnorm*”, para o caso central e não-central, com $d = 1$, 32 pontos na quadratura, variando-se os valores dos graus de liberdade ν

[illegible]

“Tabela 8, conclusão”

v	$pNDUDF$	$pmvt \text{ (unilateral)} (*)$	$pNDBDF$	$pmvt \text{ (bilateral)}$
		$\lambda = (0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5;$ $\delta = (0,5; 0,3; 0,2; 0,5; 0,3; 0,2; 0,5; 0,3; 0,2; 0,5; 0,3; 0,2; 0,5; 0,3; 0,2; 0,5; 0,3; 0,$		
5	0,3214773613365876	0,3215110810237615	0,14908253104951950	0,14878775
40	0,3312513297409567	0,3312907743690111	0,11534890954472570	0,11528691
120	0,3320351771417700	0,3319978730771600	0,10955782691901300	0,10954402
480	0,3323186824629374	0,3322842744330308	0,10716528325443420	0,10728022

(*) erro mínimo da ordem da quarta casa decimal.

Na Tabela 9 estão apresentadas probabilidades do teste bilateral de Dunnett balanceado em função do valor do quantil $|d|$, do vetor de correlações λ , dos graus de liberdade v e do número de comparações r , para o caso central ($\delta = 0$).

São comparados os valores obtidos pela função *pNDBDF*, pela função *mvstud* do pacote “*mvtnormpcs*” e pela função *pmvt* de o pacote “*mvtnorm*”. Apesar do pacote “*mvtnormpcs*” não estar mais disponível no sistema *R*, constata-se que a sua função *mvstud* fornecia valores precisos para as probabilidades de até $r = 50$ comparações.

Tabela 9 Probabilidades do teste bilateral de Dunnett balanceado obtidas pela função *pNDBDF* com $n = 100$ pontos na quadratura, pela função *mvstud* do pacote “*mvtnormpcs*” e pela função *pmvt* do pacote “*mvtnorm*”, para o quantil $|d| = 2,0$ em função da correlação ρ , dos graus de liberdade v e do número de comparações r , para o caso central

r	ρ	v	pNDBDF	mvstud	pmvt
3	0,2	10	0,8113716489082575	0,8113578557968140	0,8113830683823173
3	0,2	900	0,8714772347765425	0,8714449405670166	0,8714886645505286
3	0,5	10	0,8300169202508818	0,8300129771232605	0,8298486360228929
3	0,5	900	0,8844497368153013	0,8844429254531860	0,8844758882347792
3	0,8	10	0,8655867204390135	0,8655934333801270	0,8654305867908594
3	0,8	900	0,9094684275419850	0,9094656109809875	0,9094323161190694
20	0,2	10	0,4092176821076725	0,4092107117176056	0,4092211769103733
20	0,2	900	0,4440980770756253	0,4440965652465820	0,4441524178669163
20	0,5	10	0,5351623845394875	0,5351692438125610	0,5351985234395773
20	0,5	900	0,6072332519936443	0,6072408556938171	0,6069517337392991
20	0,8	10	0,7245095169119858	0,7245147228240967	0,7244203134028777
20	0,8	900	0,7911656902718861	0,7911650538444519	0,7913516849382036
30	0,2	10	0,3184388235151981	0,3184342980384827	0,3183995140879301
30	0,2	900	0,3121924490656889	0,3121851384639740	0,3124898885483320
30	0,5	10	0,4637106784100888	0,4637171626091003	0,4636319886765521
30	0,5	900	0,5231312309062763	0,5231428146362305	0,5229091885731137
30	0,8	10	0,6909920061801590	0,6909971833229065	0,6910525346027789
30	0,8	900	0,7606143133273131	0,7606118321418762	0,7604906853236557

“Tabela 9, conclusão”

r	ρ	v	pNDBDF	mvstud	pmvt
50	0,2	10	0,2198516722704251	0,2198455482721329	0,2196491370887752
50	0,2	900	0,1620519559370634	0,1620369553565979	0,1621115930233352
50	0,5	10	0,3774180200083504	0,3774199485778810	0,3772613353763954
50	0,5	900	0,4103645998243669	0,4103795289993286	0,4111604372575535
50	0,8	10	0,6479981227113798	0,6480036973953250	0,6471338724118832
50	0,8	900	0,7201440957850329	0,7201449871063230	0,7203688530739962
80	0,2	10	0,1486425026914517	-	0,1485718513912539
80	0,2	900	0,0650955630545228	-	0,0654777335959302
80	0,5	10	0,3048339530466407	-	0,3040131138961192
80	0,5	900	0,3051325364140102	-	0,3048833804252217
80	0,8	10	0,6080812262989408	-	0,6072048670608854
80	0,8	900	0,6812443358286571	-	0,6807791969856853

Pela Tabela 9 pode-se observar também que para o mesmo número de comparações e de graus de liberdade, o aumento da correlação aumenta o valor das probabilidades. No entanto, ao se comparar o valor das probabilidades mantendo-se fixo o número de comparações e a correlação, observa-se que o aumento do número de graus de liberdade influencia de forma diferente no valor da probabilidade. Quando o número de comparações é pequeno, com o aumento dos graus de liberdade ocorre um aumento na probabilidade, para qualquer valor da correlação. Conforme aumenta o número de comparações, observa-se que para correlações baixas, esta relação é inversa, ou seja, com o aumento dos graus de liberdade ocorre uma queda na probabilidade. Quando o número de comparações é maior isto ocorre com correlações maiores, inclusive com correlações maiores do que 0,5. Por exemplo, com um quantil $d = 2$ e $r = 180$ comparações, num experimento com correlação constante entre as comparações igual a 0,6, a probabilidade para o teste bilateral central é de pN DBDF = 0,28598 para $v = 10$ e pN DBDF = 0,28265 para $v = 900$. Em estudos posteriores pode-se melhor compreender como se dá o comportamento das distribuições de

probabilidade do teste em toda a amplitude de correlações possíveis, em função do número de comparações e dos graus de liberdade.

Nas Tabelas 10 e 11 estão apresentados valores dos quantis do teste bilateral de Dunnett balanceado, no caso central, com $p = 1 - \alpha = 0,95$ e $p = 1 - \alpha = 0,99$, respectivamente, e $\rho = 0,5$ em função do número de comparações r e dos graus de liberdade v . Os valores apresentados foram obtidos pelas funções $qNDBDF$ e $qNDBD$ com $n = 32$ pontos na quadratura, pela função $qmvt$ do pacote “mvtnorm” e pelas tabelas do Hsu (1999) nos casos de número de comparações abrangidas pelas referidas tabelas. Observa-se que, conforme o número de comparações aumenta, também aumenta o valor do quantil das distribuições, para o mesmo número de graus de liberdade. Considerando o aumento dos graus de liberdade para o mesmo número de comparações, observa-se que o valor do quantil diminui. Essas relações são observadas da mesma forma para as duas probabilidades $p = 0,95$ e $p = 0,99$. Essa informação de que o valor do quantil diminui com o aumento dos graus de liberdade é importante na fase de planejamento experimental, pois os graus de liberdade estão diretamente relacionados ao tamanho amostral dos tratamentos. Nesse caso, quando é viável amostras com maior número de repetições, mais facilmente pode ser comprovada a existência de diferença significativa no contraste.

Tabela 10 Quantis do teste bilateral de Dunnett balanceado, no caso central, com $\rho = 0,5$, $p = 1 - \alpha = 0,95$ e $n = 32$ pontos na quadratura, em função do número de comparações r e dos graus de liberdade v

v	HSU	qNDBDF	qmvt
r = 1			
9	2,262	2,262407653761983	2,262157162798204
40	2,021	2,021277350389747	2,021075390306273
120	1,980	1,980118281037748	1,979930405082422
$\infty(*)$	1,960	1,960143983738570	1,959983753662981

“Tabela 10, conclusão”

ν	<i>HSU</i>	<i>qNDBDF</i>	<i>qmvmt</i>
r = 15			
9	3,535	3,534850561320265	3,536222003512497
40	2,970	2,969505180782855	2,969919769425486
120	2,875	2,873942241770314	2,874945107636507
$\infty(*)$	2,829	2,827422503199287	2,829092111721982
r = 50			
9	-	4,029872787453915	4,012415500426735
40	-	3,311703047250819	3,311266500817176
120	-	3,191717551210948	3,192921466771687
$\infty(*)$	-	3,132538897268113	3,136810342535750
r = 100			
9	-	4,267895375298354	4,268178097413750
40	-	3,492743780917273	3,493234918265517
120	-	3,359215345054900	3,359772591030518
$\infty(*)$	-	3,292994346200360	3,296588767286848
r = 400			
10	-	4,611732928841250	4,609471540679550
40	-	3,826884086773704	3,827683293708740
120	-	3,667451229473050	3,667049399661720
$\infty(*)$	-	3,588477849203994	3,587600180153800

(*) obtido pela função $qNDBD(p, r, \text{delta}, n)$.

Nas Tabelas 12 e 13 estão os valores dos quantis do teste bilateral de Dunnett balanceado, no caso central, com $p = 1 - \alpha = 0,95$ e $p = 1 - \alpha = 0,99$, respectivamente, em função do número de comparações r , dos graus de liberdade ν , variando-se o valor da correlação $\rho = 0,2; 0,4; 0,5; 0,60; 0,75$. Os valores apresentados foram obtidos pelas funções *qNDBDF* com $n = 32$ pontos na quadratura, e pela função *qmvmt* do pacote “*mvtnorm*”. Inicialmente, observa-se que no caso de uma comparação, $r = 1$, as funções fornecem o mesmo valor do quantil, independente dos valores da correlação. Isso ocorre porque no caso

de uma só comparação não existe correlação entre comparações, pois a distribuição se simplifica na t univariada.

Tabela 11 Quantis do teste bilateral de Dunnett balanceado, no caso central com $\rho = 0,5$, $\delta = 0$, $p = 1 - \alpha = 0,99$ e $n = 32$ pontos na quadratura, em função do número de comparações r e dos graus de liberdade ν

ν	<i>HSU</i>	<i>qNDBDF</i>	<i>qmv</i>
r = 1			
9	3,251	3,250821980635707	3,249835541592125
40	2,704	2,705495559268186	2,704459267433162
120	2,617	2,618499382946313	2,617421145106865
$\infty(*)$	2,576	2,576932946472953	2,575870275392421
r = 15			
9	4,683	4,682435765923735	4,672939153121355
40	3,598	3,599831686973376	3,600577274727130
120	3,430	3,433240519342290	3,429224424492776
$\infty(*)$	3,350	3,355187660159780	3,352424734255180
r = 50			
9	-	5,248157348609074	5,262269182817700
40	-	3,939271566862792	3,942192993882864
120	-	3,738036461369261	3,734891989839134
$\infty(*)$	-	3,645595917567103	3,635953823187740
r = 100			
9	-	5,552994772842930	5,539795955304107
40	-	4,121477063796293	4,117640118730612
120	-	3,900761203837250	3,886612113055520
$\infty(*)$	-	3,800835422237547	3,787742252237294
r = 400			
9	-	6,122011738145607	6,103170758314130
40	-	4,461583938935660	4,465354038232636
120	-	4,202901225165413	4,200595262918237
$\infty(*)$	-	4,089253020427740	4,085193035751085

(*) obtido pela função $qNDBD(p, r, \delta, n)$.

Pelas Tabelas 12 e 13 comparando-se o valor do quantil em relação aos graus de liberdade, com a mesma correlação aumentando-se os graus de liberdade ocorre uma diminuição no valor do quantil. Por outro lado, com o aumento do número de comparações, para o mesmo número de graus de liberdade e mesma correlação, o valor do quantil aumenta. Essas relações são observadas da mesma forma para as duas probabilidades $p = 0,95$ e $p = 0,99$.

A partir das Tabelas 12 e 13 também se observa que para o mesmo número de comparações e graus de liberdade, conforme o valor da correlação aumenta o valor do quantil diminui, mantendo-se fixa a probabilidade. Ou seja, quanto menor a correlação, maior o valor do quantil. Dunnett(1964) considerou esta questão ao incluir, na publicação das suas novas tabelas para testes do tipo MCC, um método de correção para o quantil equicorrelacionado que a tabela fornecia nos casos com correlações menores do que 0,5.

Tabela 12 Quantis do teste bilateral de Dunnett balanceado, no caso central, com $p = 1 - \alpha = 0,95$ e $n = 32$ pontos na quadratura, em função do número de comparações r , dos graus de liberdade ν e do valor da correlação ρ

r	ν	$qNDBDF$	$qmv\tau$
$\rho = 0,2$			
1	9	2,262293382223230	2,262157162798205
15	120	2,968328616993119	2,968272676775902
50	300	3,292237404297974	3,292052272380782
50	1200	3,270753738730534	3,270027823728713
100	300	3,482959514776127	3,482217683874110
100	1200	3,458230533259006	3,457619666734760
400	300	3,838635648055526	3,836099781046326
400	1200	3,807249232208513	3,805305211939601

“Tabela 12, continuação”

r	ν	$qNDBDF$	$qmv\tau$
$\rho = 0,4$			
1	9	2,262374250190340	2,262157162798205
15	120	2,917051780006589	2,916930908448603
50	300	3,218104339955390	3,217260048846396
50	1200	3,199192505150986	3,200412840189197
100	300	3,392939923522793	3,392011015744620
100	1200	3,371454841014215	3,374212418685474
400	300	3,715704219342110	3,718523755916460
400	1200	3,689032667530900	3,692967432464550
$\rho = 0,5$			
1	9	2,262407653761983	2,262157162798205
15	120	2,873942241770314	2,874945107636507
50	300	3,156265401790887	3,157642778133778
50	1200	3,138491973846286	3,144480614516396
100	300	3,319627035312050	3,320904333481244
100	1200	3,299694863860150	3,305554257303160
400	300	3,620187649859925	3,623120216963587
400	1200	3,596451426118756	3,596748163042480
$\rho = 0,6$			
1	9	2,262405603825037	2,262157162798205
15	120	2,815698129045571	2,817618705960482
50	300	3,075230699407817	3,078503963598772
50	1200	3,059288607017559	3,060218651138311
100	300	3,225142451713763	3,226451919224349
100	1200	3,207915396552372	3,210243481494084
400	300	3,499759613609014	3,496365374506372
400	1200	3,480748948116408	3,475867659236415

“Tabela 12, conclusão”

<i>r</i>	<i>v</i>	<i>qNDBDF</i>	<i>qmv</i>
$\rho = 0,75$			
1	9	2,262220345157308	2,262157162798205
15	120	2,690989218625786	2,691812474543030
50	300	2,908825330684217	2,905566390342563
50	1200	2,898786685340200	2,889843058482880
100	300	3,032877085419235	3,025789452375880
100	1200	3,023392248844917	3,010407180214314
400	300	3,257609765343446	3,244975174277843
400	1200	3,229266813067587	3,228998539110050

Tabela 13 Quantis do teste bilateral de Dunnett balanceado, no caso central, com $p = 1 - \alpha = 0,99$ e $n = 32$ pontos na quadratura, em função do número de comparações r , dos graus de liberdade v e do valor da correlação ρ

<i>r</i>	<i>v</i>	<i>qNDBDF</i>	<i>qmv</i>
$\rho = 0,2$			
1	9	3,250368243925903	3,249835541592125
15	120	3,485538734597327	3,485042255779190
50	300	3,755615190692590	3,754693984610962
50	1200	3,722234215844463	3,719957183462010
100	300	3,931519611710035	3,931165430612506
100	1200	3,893823220047815	3,893534040416453
400	300	4,264296266342476	4,262115340659284
400	1200	4,217700562180558	4,215702814184188
$\rho = 0,4$			
1	9	3,250612826607794	3,249835541592125
15	120	3,459028000784110	3,457508424947267
50	300	3,718398361414369	3,712772792442421
50	1200	3,712772792442421	3,684739794013008
100	300	3,884870735426310	3,885139611022879
100	1200	3,851956329127058	3,846767368267685
400	300	4,196015375222595	4,190848678713910
400	1200	4,157702444143808	4,148649853911074

“Tabela 13, conclusão”

<i>r</i>	<i>v</i>	<i>qNDBDF</i>	<i>qmvt</i>
$\rho = 0,5$			
1	9	3,250821980635707	3,249835541592125
15	120	3,433240519342290	3,429224424492776
50	300	3,681632946641043	3,676725276579973
50	1200	3,654469068943323	3,645421556576080
100	300	3,839365019637973	3,830422350465660
100	1200	3,810228736794454	3,802681225747631
400	300	4,131651167024943	4,127187694172008
400	1200	4,099187942303606	4,077307846648280
$\rho = 0,6$			
1	9	3,251119934231895	3,249835541592125
15	120	3,394642024672320	3,396989591520984
50	300	3,626514426199454	3,614349998554530
50	1200	3,602975743391702	3,586641591984543
100	300	3,771975298909729	3,763932950536114
100	1200	3,747475801221997	3,725952249989691
400	300	4,038960035366154	4,031587455183168
400	1200	4,012804504327774	3,993872906243045
$\rho = 0,75$			
1	9	3,251786124399999	3,249835541592125
15	120	3,297070742721380	3,290895020129418
50	300	3,486905937941090	3,473594937139957
50	1200	3,468583414796159	3,448675503233097
100	300	3,604584560405272	3,583476212604860
100	1200	3,584992367664654	3,569381064480714
400	300	3,817511852761212	3,799212820944980
400	1200	3,792086158541832	3,788167087213330

Com isso, torna-se claro o erro que se comete ao utilizar indiscriminadamente tabelas com valor crítico para as estatísticas do teste de Dunnett equicorrelacionadas ($\rho = 0,5$) sem levar em consideração se, ao final do experimento, as amostras de todos os tratamentos, ou seja, do tratamento

controle e de todos os tratamentos em teste, mantém o mesmo número de repetições. Por exemplo, supondo uma análise em que foi obtido um quantil tabelado dado por $d_{(\rho=0,5; v=120; r=15)} = 2,874$ para $p = 0,95$. Para um contraste com valor da estatística $d = 2,90$ concluir-se-ia que o efeito é significativo. Porém se a correlação não fosse $\rho = 0,5$ mas fosse $\rho = 0,4$, o quantil tabelado correto é $d_{(\rho=0,4; v=120; r=15)} = 2,917$ para $p = 0,95$. Neste caso, para o contraste com valor da estatística $d = 2,90$ concluir-se-ia que o efeito não é significativo. Observa-se que as decisões são contrárias. No entanto, esta observação inicial se contrapõe a outra mais minuciosa, também relatada por Dunnett (1964). Quando se considera a correta correlação entre os contrastes, no caso de correlações menores do que 0,5 existe um ganho relativo na estatística de teste calculada. Devido ao maior número de observações do tratamento controle em relação aos tratamentos em teste, o resultado é uma diminuição do erro padrão entre a diferença dos tratamentos, que aparece no denominador da estatística calculada

$\hat{d}_i = (\hat{\mu}_i - \hat{\mu}_{r+1}) / (\hat{\sigma} \sqrt{n_i + 1/n_{r+1}})$. Este fato faz com que, apesar do aumento do valor crítico tabelado pela diminuição da correlação entre contrastes, o aumento da estatística calculada devido a diminuição do erro padrão leva a uma estatística que mais facilmente rejeite a hipótese nula.

No ANEXO A está apresentado um exemplo fictício de como o tamanho amostral maior do tratamento controle pode afetar o resultado do teste.

Neste trabalho estas observações são primárias, sugerindo que estudos futuros esclareçam se isto ocorre em qualquer situação de número de comparações, graus de liberdade, valores de correlações entre contrastes e de não-centralidade também. Estudos de poder do teste em função da relação entre os tamanhos amostrais do tratamento controle e dos tratamentos em teste nestas diversas situações também são interessantes.

8 CONCLUSÕES

As funções densidade e de distribuição das distribuições t multivariadas não-centrais que envolvem o teste de Dunnett foram matematicamente desenvolvidas e apresentadas de forma clara e compreensível.

Os algoritmos para a obtenção da função de distribuição, da função densidade e da inversa da função de distribuição, para obtenção de quantis das distribuições do máximo e do máximo do módulo da normal e da t multivariadas utilizando quadraturas para a resolução das integrais múltiplas foram propostos e implementados no *software R* com sucesso.

Os valores das probabilidades e dos quantis obtidos pelas rotinas propostas têm, no mínimo, a mesma precisão de valores obtidos para comparação em referências bibliográficas ou outras rotinas computacionais. Além disso, a precisão dos valores fornecidos pelas rotinas propostas pode aumentar conforme a necessidade do usuário, apenas aumentando o número de pontos para a quadratura.

A biblioteca *nCDunnett* foi implementada e submetida ao CRAN-R e encontra-se disponibilizada para instalação em todas as versões R iguais ou superiores a 2.15.0, com livre acesso por todos os usuários do programa no mundo. Esta biblioteca possibilita que qualquer pesquisador possa aplicar corretamente o teste de Dunnett, independente do valor da correlação entre comparações, sem dependência de tabelas limitadas disponíveis na literatura.

No aspecto teórico, foi sintetizado, redigido e disponibilizado um referencial didático minucioso sobre o teste de Dunnett em procedimentos unilaterais e bilaterais, para dados balanceados ou não. Buscou-se esclarecer suas pressuposições e mostrar todas as passagens das demonstrações dos teoremas a eles relacionadas. Espera-se que com isso haja uma contribuição para

o entendimento do tema, ampliando a bibliografia nacional que ainda é escassa nesta área.

REFERÊNCIAS

- ANDRADE, D. F.; OGLIARI, P. J. **Estatística para as ciências agrárias e biológicas**: com noções de experimentação. Florianópolis: UFSC, 2007. 438 p.
- BARROSO, L. C. et al. **Cálculo numérico com aplicações**. 2. ed. São Paulo: Harbra, 1987. 367 p.
- BENJAMINI, Y.; HOCHBERG, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of the Royal Statistical Society, Serie B**, London, n. 57, p. 289-300, 1995.
- BOHRER, R.; SCHERVISH, M. J. An error-bounded algorithm for normal probabilities of rectangular regions. **Journal of Applied Statistics**, Abingdon, v. 23, n. 3, p. 297-300, 1981.
- BOHRER, R.; SCHERVISH, M. J.; SHEFT, J. Algorithm as 184: non-central studentized maximum and related multiple t-probabilities. **Journal of Applied Statistics**, Abingdon, v. 31, n. 3, p. 309-317, 1982.
- BORGES, C. L.; FERREIRA, D. F. Poder e taxas de erro tipo I dos testes Scott-Knott, Tukey e Student-Newman-Keuls sob distribuição normal e não normais dos resíduos. **Revista de Matemática e Estatística**, Marília, v. 21, n. 1, p. 67-83, 2003.
- BRETZ, F.; HOTHORN, T.; WESTFALL, P. **Multiple comparisons using R**. New York: Chapman & Hall/CRC, 2011. 182 p.
- BURDEN, R. L.; FAIRES, J. D. **Análise numérica**. São Paulo: Thomson, 2003. 763 p.
- CHAPRA, S. C.; CANALE, R. P. **Numerical methods for engineers**. 6th ed. Boston: McGraw-Hill Higher Education, 2010. 968 p.
- DEAN, A. M.; VOSS, D. T. **Design and analysis of experiments**. New York: Springer-Verlag, 1999. 740 p.

DMITRIENKO, A.; TAMHANE, A. C.; BRETZ, F. **Multiple testing problems in pharmaceutical statistics**. New York: Chapman & Hall/CRC, 2010. 289 p.

DREZNER, Z. Computation of the multivariate normal integral. **ACM Transactions on Mathematical Software**, Davis, v. 18, n. 4, p. 470-480, 1992.

DUNLAP, W. P.; MARX, M. S.; AGAMY, G. J. FORTRAN IV functions for calculating probabilities associated with Dunnett's test. **Behavior Research Methods & Instrumentation, Psychonomic Society**, Austin, v. 13, n. 3, p. 363-366, 1981.

DUNNETT, C. W. Algorithm AS 251: multivariate normal probabilities integrals with product correlation structure. **Journal of Applied Statistics**, Abingdon, v. 38, p. 564-579, 1989.

_____. Multiple comparison procedure for comparing several treatments with a control. **Journal of the American Statistical Association**, New York, v. 50, n. 272, p. 1096-1121, 1955.

_____. New tables for multiple comparisons with a control. **Biometrika**, London, v. 20, n. 3, p. 482-491, 1964.

FERREIRA, D. F. **Estatística básica**. 2. ed. Lavras: UFLA, 2009. 664 p.

_____. **Estatística multivariada**. Lavras: UFLA, 2008. 661 p.

FRANCO, N. B. **Cálculo numérico**. São Paulo: Pearson Prentice Hall, 2007. 505 p.

GENZ, A.; BRETZ, F. Methods for the computation of multivariate t-probabilities. **Journal of Computational and Graphical Statistics**, Alexandria, v. 11, n. 2, p. 950-971, 2002.

_____. Numerical computation of the multivariate t-probabilities with applications to power calculation of multiple contrasts. **Journal of Statistical Computation and Simulation**, New York, v. 63, n. 4, p. 361-378, 1999.

GIL, A.; SEGURA, J.; TEMME, N. M. **Numerical methods for special functions**. Philadelphia: Siam, 2007. 417 p.

GIRARDI, L. H.; CARGNELUTTI FILHO, A.; STORCK, L. Erro tipo I e poder de cinco testes de comparações múltiplas de médias. **Revista Brasileira de Biometria**, São Paulo, v. 27, n. 1, p. 23-36, 2009.

HINKELMANN, K.; KEMPTHORNE, O. **Design and analysis of experiments**. 2nd ed. New York: J. Wiley, 2007. 631 p.

HOCHBERG, Y.; TAMHANE, A. C. **Multiple comparisons procedures**. New York: J. Wiley, 1987. 450 p.

HSU, J. C. **Multiple comparisons: theory and methods**. New York: Chapman & Hall/CRC, 1999. 277 p.

IZBICKI, R. **Classe de testes de hipóteses**. 2010. 48 p. Dissertação (Mestrado em Ciências) - Universidade de São Paulo, São Paulo, 2010.

JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous univariate distributions**. 2nd ed. New York: J. Wiley, 1995. 717 p.

KHURI, A. I. **Advanced calculus with applications in statistics**. 2nd ed. New York: J. Wiley, 2003. 673 p.

KOLMOGOROV, A. **Foundations of the theory of probability**. 2nd ed. New York: Chelsea, 1956. 84 p. Disponível em:
<<http://www.matematik.com/Kolmogorov/index.html>>. Acesso em: 10 dez. 2012.

KOTZ, S.; NADARAJAH, S. **Multivariate t-distributions and their applications**. Cambridge: Cambridge University, 2004. 272 p.

KRISHNAMOORTHY, K. **Handbook of statistical distributions with Applications**. Chicago: Taylor & Francis, 2006. 344 p.

LEITHOLD, L. **O Cálculo com geometria analítica**. 3. ed. São Paulo: Harbra, 1994. 546 p.

LIMA, E. L. **Análise real: função de uma variável**. 9. ed. Rio de Janeiro: IMPA, 2007. 198 p.

MAGALHÃES, M. N. **Probabilidade e variáveis aleatórias**. 2. ed. São Paulo: EDUSP, 2006. 411 p.

MAGALHÃES, M. N.; LIMA, A. C. P. de. **Noções de probabilidade e estatística**. 6. ed. São Paulo: EDUSP, 2005. 392 p.

MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. **Multivariate analysis**. 2nd ed. New York: Academic, 2003. 521 p.

MEYER, P. L. **Probabilidade: aplicações à estatística**. 2. ed. Rio de Janeiro: Livros Técnicos e Científicos, 1984. 426 p.

MONTGOMERY, D. C. **Design and analysis of experiments**. 5th ed. New York: J. Wiley, 2001. 684 p.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics**. 3rd ed. Singapore: McGraw-Hill, 1974. 564 p.

R CORE TEAM DEVELOPMENT. **R: a language and environment for statistical computing**. Vienna, 2012. Disponível em: <<http://www.Rproject.org/>>. Acesso em: 12 dez. 2012.

RALSTON, A.; RABINOWITZ, P. **A first course in numerical analysis**. Tokyo: McGraw-Hill, 1965. 578 p.

RAMALHO, M. A. P.; FERREIRA, D. F.; OLIVEIRA, A. **Experimentação em genética e melhoramento de plantas**. 2. ed. Lavras: UFLA, 2005. 322 p.

ROSS, S. M. **Introduction to probability models**. 9th ed. New York: Academic, 2007. 782 p.

SCHERVISH, M. J. Algorithm AS 195: multivariate normal probabilities with error bound. **Journal of the Royal Statistical Society. Serie C, Applied Statistics**, London, v. 33, n. 1, p. 81-87, 1984.

_____. Correction to algorithm AS 195: multivariate normal probabilities with error bound. **Journal of the Royal Statistical Society. Serie C, Applied Statistics**, London, v. 34, p. 103-104, 1985.

SPERANDIO, D.; MENDES, J. T.; SILVA, L. H. M. e. **Cálculo numérico: características matemáticas e computacionais dos métodos numéricos**. São Paulo: Pearson Prentice Hall, 2003. 354 p.

STEEN, M.; GELBARD, E. Gaussian quadratures for the integrals $\int_0^\infty e^{-x^2} f(x) dx$ and $\int_0^b e^{-x^2} f(x) dx$. **Mathematics of Computation**, Providence, v. 23, p. 661-671, 1969.

WANG, M.; KENNEDY, W. J. A numerical method for accurately approximating multivariate normal probabilities. **Computational Statistics and Data Analysis**, New York, v. 13, n. 2, p. 197-210, Mar. 1992.

APÊNDICES

APÊNDICE A – Aplicação do teste de Dunnett

Considerando os dados experimentais fictícios da Tabela 1, em que está sendo estudado o efeito de três catalisadores sobre o rendimento de uma reação.

Nos dois experimentos, um quarto tratamento, sem catalisador, é usado como controle. Testar a hipótese de diferença entre efeitos do rendimento da reação, comparando cada catalisador com o controle, com um nível de significância conjuntade 5% em cada um dos experimentos.

Tabela 1 Rendimento de uma reação em dois experimentos

Experimento I				Experimento II			
T_{11}	T_{12}	T_{13}	T_{14}	T_{21}	T_{22}	T_{23}	T_{24}
54,1	52,7	51,2	50,7	54,1	52,7	51,2	50,7
53,8	53,9	50,8	51,5	53,8	53,9	50,8	51,5
53,1	57,0	49,7	49,2	53,1	57,0	49,7	49,2
52,5	54,1	48,0	53,1	52,5	54,1	48,0	53,1
54,0	52,5	47,2	52,7	54,0	52,5	47,2	52,7
-	-	-	-	-	-	-	51,4
-	-	-	-	-	-	-	50,4
$\hat{\mu}_{11}$	$\hat{\mu}_{12}$	$\hat{\mu}_{13}$	$\hat{\mu}_{14}$	$\hat{\mu}_{21}$	$\hat{\mu}_{22}$	$\hat{\mu}_{23}$	$\hat{\mu}_{24}$
53,50	54,04	49,38	51,44	53,50	54,04	49,38	51,29

Ambos os experimentos tem $i = 1, 2, 3$ tratamentos em teste e o T_4 é o tratamento controle. O primeiro experimento tem todos os tratamentos com mesmo tamanho amostral, $n_i = n = 5$ para todo i , tratando-se de um experimento equicorrelacionado, com correlação constante igual a

$$\rho_1 = \frac{n_i}{n_i + n_4} = \frac{5}{5 + 5} = 0,5.$$

O segundo experimento tem 5 (cinco) repetições nos tratamentos em teste e o tratamento controle tem 7 repetições, tratando-se de um experimento com correlação constante entre as diferenças de médias igual a

$$\rho_2 = \frac{n_i}{n_i + n_4} = \frac{5}{5 + 7} = 0,4167.$$

Calculando o quadrado médio do erro ou variância residual dos experimentos tem-se que:

$$\hat{\sigma}_1^2 = QME_1 = \sum_{j=1}^4 \sum_{a=1}^5 \frac{(Y_{ja} - \bar{Y}_j)^2}{4(5-1)} = 2,300750$$

ou seja, $\hat{\sigma}_1 = 1,516822$, e

$$\hat{\sigma}_2^2 = QME_2 = \sum_{j=1}^4 \sum_{a=1}^5 \frac{(Y_{ja} - \bar{Y}_j)^2}{3 \cdot (5-1) + 6} = 2,096032$$

ou seja, $\hat{\sigma}_2 = 1,447768$.

Para cada contraste entre o i -ésimo novo tratamento e o tratamento controle determina-se a estatística $\hat{d}_C = \frac{\hat{\mu}_i - \hat{\mu}_{r+1}}{\hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_{r+1}}}}$.

Assim, no primeiro experimento se obtém:

$$\hat{d}_{C11} = \frac{53,50 - 51,44}{1,516822336 \sqrt{\frac{1}{5} + \frac{1}{5}}} = \frac{2,06}{0,95932268} = 2,1473484$$

$$\hat{d}_{C12} = \frac{54,04 - 51,44}{1,516822336 \sqrt{\frac{1}{5} + \frac{1}{5}}} = \frac{2,6}{0,95932268} = 2,71024552$$

$$\hat{d}_{C13} = \frac{49,38 - 51,44}{1,516822336 \sqrt{\frac{1}{5} + \frac{1}{5}}} = \frac{-2,06}{0,95932268} = -2,1473484$$

No segundo experimento se obtém:

$$\hat{d}_{C21} = \frac{53,50 - 51,29}{1,447768\sqrt{\frac{1}{5} + \frac{1}{7}}} = 2,612030$$

$$\hat{d}_{C22} = \frac{54,04 - 51,29}{1,447768\sqrt{\frac{1}{5} + \frac{1}{7}}} = 3,249028$$

$$\hat{d}_{C23} = \frac{49,38 - 51,29}{1,447768\sqrt{\frac{1}{5} + \frac{1}{7}}} = -2,248031$$

Com o nível de significância $\alpha = 0,05$ desejável para o teste simultâneo dos 3 contrastes e usando como parâmetros de entrada:

- cada valor $\hat{d}_C$ estimado;
- o vetor de correlações $\lambda_1 = (0,5; 0,5; 0,5)$ no experimento I e $\lambda_2 = (0,4167; 0,4167; 0,4167)$ no experimento II;
- o vetor de parâmetros de não-centralidade δ , que nestes casos é 0;
- os graus de liberdade $\nu = 16$ no experimento I e $\nu = 18$ no experimento II para a distribuição da variância amostral residual em cada caso.

O *scrip* do R para obter o quantil das estatísticas do teste de Dunnett bilateral para estes dois experimentos é dado a seguir, primeiro efetuando o procedimento pela alternativa 1 e depois pela alternativa 2, descritas na seção 6.3.

```
#####
# Script para obter o quantil do teste de Dunnett bilateral #
# Alternativa 1 de resolução #
#####
# Experimento I
> library(nCDunnett)
> q<-c(2.1473,2.7102,2.1473)
> rho<-c(0.5,0.5,0.5)
> nu<-16
> delta <- c(0,0,0)
> n<-32
> pNCDun(q, nu, rho, delta, n, TRUE)
[1] 0.8845471 0.9602866 0.8845471
>
# Experimento II
> q<-c(2.6120,3.2490,2.2480)
> rho<-c(0.4167,0.4167,0.4167)
> nu<-18
> delta <- c(0,0,0)
> n<-32
> pNCDun(q, nu, rho, delta, n, TRUE)
[1] 0.9529890 0.9876631 0.9037866
```

- $|\hat{d}_{C11}| = 2,1473484$ a probabilidade conjunta 0,8845, ou seja, valor- $p = 0,1155$. Comparando-se com $\alpha = 0,05$ conclui-se que o efeito deste catalisador não é significativo.
- $|\hat{d}_{C12}| = 2,71024552$ a probabilidade conjunta 0,9603, ou seja, valor- $p = 0,0397$. Comparando-se com $\alpha = 0,05$ conclui-se que o efeito deste catalisador é significativo.
- $|\hat{d}_{C13}| = 2,1473484$ a probabilidade conjunta 0,8845, ou seja, valor- $p = 0,1155$. Comparando-se com $\alpha = 0,05$ conclui-se que o efeito deste catalisador não é significativo.

No experimento II obteve-se, para:

- $|\hat{d}_{C21}| = 2,6120$ a probabilidade conjunta 0,9530, ou seja, valor- $p = 0,047$. Comparando-se com $\alpha = 0,05$ conclui-se que o efeito deste catalisador é significativo.
- $|\hat{d}_{C22}| = 3,2490$ a probabilidade conjunta 0,9877, ou seja, valor- $p = 0,0123$. Comparando-se com $\alpha = 0,05$ conclui-se que o efeito deste catalisador é significativo.

- $|\hat{d}_{C23}| = 2,2480$ a probabilidade conjunta 0,9038, ou seja, valor- $p = 0,0962$. Comparando-se com $\alpha = 0,05$ conclui-se que o efeito deste catalisador não é significativo.

```
#####
# Alternativa 2 de resolução #
#####
# Experimento I
> library(nCDunnett)
> rho<-c(0.5,0.5,0.5)
> nu<-16
> delta <- c(0,0,0)
> n<-32
> p<-0.95
> qNCDun(p, nu, rho, delta, n, TRUE)
[1] 2.592548
```

```
# Experimento II
> rho<-c(0.4167,0.4167,0.4167)
> nu<-18
> delta <- c(0,0,0)
> n<-32
> p<-0.95
> qNCDun(p, nu, rho, delta, n, TRUE)
[1] 2.581522
```

Neste caso, a função fornece o quantil $d_1 = 2,5925$ e $d_2 = 2,5815$, usando a inversa da função de distribuição da probabilidade 0,95 fornecida. Compara-se o valor do quantil fornecido pelo programa com os valores das estatísticas em cada contraste. Para o experimento I tem-se:

- $|\hat{d}_{11}| = 2,1473484 < d = 2,5925$ conclui-se que o efeito deste catalisador não é significativo.
- $|\hat{d}_{12}| = 2,71024552 > d = 2,5925$ conclui-se que o efeito deste catalisador é significativo.
- $|\hat{d}_{13}| = 2,1473484 < d = 2,5925$ conclui-se que o efeito deste catalisador não é significativo.

Para o experimento II tem-se:

- $|\hat{d}_{21}| = 2,612030 > d = 2,5815$ conclui-se que o efeito deste catalisador é significativo.
- $|\hat{d}_{22}| = 3,249028 > d = 2,5815$ conclui-se que o efeito deste catalisador é significativo.
- $|\hat{d}_{23}| = 2,248031 < d = 2,5815$ conclui-se que o efeito deste catalisador não é significativo.

Portanto, no experimento I apenas o rendimento médio da reação do catalisadore 3 é significativamente diferente do rendimento médio da reação utilizando o controle enquanto que no experimento II o rendimento médio da reação dos catalisadores 2 e 3 são significativamente diferentes do rendimento médio da reação utilizando o controle com tamanho amostral maior do que os tratamentos em teste, em um nível de significância de 5%.

APÊNDICE B - Rotinas R

```
#####
# Algoritmo implementado em R das funções para obter probabili-#
# dades, inversas e densidades da distribuição da estatística #
# do teste de Dunnett unilateral e bilateral, não-balanceado e #
# não-central, dados nu > 0 graus de liberdade reais, delta #
# (vetor em R^k) parâmetro de não-centralidade e lambda (vetor #
# R_+^k) das correlações, sendo k médias regulares. #
#####

library(statmod)
```

```
#####
# Algoritmo implementado em R para obter a função de distribui-#
# ção para o caso unilateral não-balanceado e não-central, #
# com nu = infity #
#####

pNDUD <- function(q, r, delta, n = 32)
{
  k <- length(r)
  x <- gauss.quad(n, kind="legendre")
  # transformando de -1 a 1 para -infity a +infity, t = x/(1-x^2),
  # x em -1 a 1
  y <- matrix(x$nodes/(1-(x$nodes)^2), n, 1)
  pArgYi <- function(y, r, delta)
  {
    ti <- pnorm((r^0.5 * y + q - delta)/(1-r)^0.5, log.p=TRUE)
    ti <- exp(sum(ti)+dnorm(y, log=TRUE))
    return(ti)
  }
  I <- apply(y, 1, pArgYi, r, delta)
  I <- I * ((1+x$nodes^2)/(1-(x$nodes)^2)^2)
  I <- sum(I * x$weights)
  return(I)
}
```

```
#####
# Algoritmo implementado em R para obter a função de distribui- #
# ção para o caso bilateral não-balanceado e não-central, com #
# nu = infty #
#####

pNDBD <- function(q, r, delta, n = 32)
{
  k <- length(r)
  x <- gauss.quad(n, kind="legendre")
  # transformando de -1 a 1 para -infty a +infty, t = x/(1-x^2),
  # x em -1 a 1
  y <- matrix(x$nodes/(1-(x$nodes)^2), n, 1)
  pArgYi <- function(y, r, delta)
  {
    ti <- pnorm((r^0.5 * y + q - delta)/(1-r)^0.5, log.p=FALSE)
    ti <- ti - pnorm((r^0.5 * y - q - delta)/(1-r)^0.5,
    log.p=FALSE)
    if (any(ti <= 0))
    {
      ti[ti <= 0] <- -743.7469
      ti[ti > 0] <- log(ti[ti > 0])
    } else ti <- log(ti)
    ti <- exp(sum(ti)+dnorm(y, log=TRUE))
    return(ti)
  }
  I <- apply(y, 1, pArgYi, r, delta)
  I <- I * ((1+x$nodes^2)/(1-x$nodes^2)^2)
  I <- sum(I * x$weights)
  return(I)
}
```

```
#####
# Algoritmo implementado em R para obter a função densidade para#
# o caso unilateral não-balanceado e não-central, com nu = infty#
#####

dNDUD <- function(q, r, delta, n = 32)
{
  k <- length(r)
  x <- gauss.quad(n, kind="legendre")
  # transformando de -1 a 1 para -infty a +infty, t = x/(1-x^2),
  # x em -1 a 1
  y <- matrix(x$nodes/(1-(x$nodes)^2), n, 1)
  rgdgi <- cbind(r, delta, (1:k))
  loopg <- function(rd, ym, q)
  {
    tg <- pnorm((rd[1]^0.5*ym+q-rd[2])/(1-rd[1])^0.5, log.p=TRUE)
    return(tg)
  }
  loopi <- function(rd, ym, q)
  {
    ti <- dnorm((rd[1]^0.5*ym+q-rd[2])/(1-rd[1])^0.5, log=TRUE)-
      0.5*log(1-rd[1])
    return(ti)
  }
  # loop dos y's
  loopy <- function(ym, q, rd)
  {
    Phy <- apply(rd, 1, loopg, ym, q) # loopg
    phy <- apply(rd, 1, loopi, ym, q) # loopi
    sPhy <- sum(Phy) # soma das diferenças dos Phi's
    sTi <- sum(exp(phy + sPhy - Phy)) * dnorm(ym)
    return(sTi)
  }
  I <- apply(y, 1, loopy, q, rgdgi)
  I <- I * ((1+x$nodes^2)/(1-x$nodes^2)^2)
  I <- sum(I * x$weights)
  return(I)
}
```

```
#####
# Algoritmo implementado em R para obter a função densidade para#
# o caso bilateral não-balanceado e não-central, com nu = infty #
#####

dNDBD <- function(q, r, delta, n = 32)
{
  k <- length(r);  x <- gauss.quad(n, kind="legendre")
  # transformando de -1 a 1 para -infty a +infty, t = x/(1-x^2),
  # x em -1 a 1
  y <- matrix(x$nodes/(1-(x$nodes)^2), n, 1)
  rgdgi <- cbind(r, delta, (1:k))
  loopg <- function(rd, ym, q)
  {
    tg <- pnorm((rd[1]^0.5 * ym + q - rd[2])/(1 - rd[1])^0.5) -
          pnorm((rd[1]^0.5 * ym - q - rd[2]) / (1 - rd[1])^0.5)
    if (tg <= 0) tg <- -743.7469 else tg <- log(tg)
    return(tg)
  }
  loopi <- function(rd, ym, q)
  {
    ti <- dnorm((rd[1]^0.5 * ym + q - rd[2])/(1 - rd[1])^0.5) +
          dnorm((rd[1]^0.5 * ym - q - rd[2]) / (1 - rd[1])^0.5)
    if (ti <= 0) ti <- -743.7469 else ti <- log(ti)
    aux <- ti - 0.5 * log(1 - rd[1])
    return(aux)
  } # loop dos y's
  loopy <- function(ym, q, rd)
  {
    Phy <- apply(rd, 1, loopg, ym, q) # loopg
    phy <- apply(rd, 1, loopi, ym, q) # loopi
    sPhy <- sum(Phy) # soma das diferenças dos Phi's
    sTi <- sum(exp(phy + sPhy - Phy)) * dnorm(ym)
    return(sTi)
  }
  I <- apply(y, 1, loopy, q, rgdgi)
  I <- I * ((1+x$nodes^2)/(1-x$nodes^2)^2)
  I <- sum(I * x$weights); return(I)
}
```

```
#####
# Algoritmo implementado em R para obter a inversa da função de #
# distribuição para o caso unilateral não-balanceado e não-cen- #
# tral , com nu = infty #
#####

qNDUD <- function(p, r, delta, n = 32)
{
  k <- length(r)
  alpha <- 1 - p
  alpha_k <- 1 - (1-alpha)^(1/k)
  q0 <- qnorm(1 - alpha_k) # valor inicial
  maxIt <- 5000
  tol <- 1e-13
  conv <- F
  it <- 0
  while ((conv == F) & (it <= maxIt))
  {
    q1 <- q0 - (pNDUD(q0, r, delta, n) - p)/dNDUD(q0, r, delta, n)
    if (abs(q1-q0) <= tol) conv <- T
    q0 <- q1
    it <- it + 1
  }
  return(q1)
}
```

```
#####
# Algoritmo implementado em R para obter a inversa da função de #
# distribuição para o caso bilateral não-balanceado e não-cen- #
# tral , com nu = infty #
#####

qNDBD <- function(p, r, delta, n = 32)
{
  k <- length(r)
  alpha <- 1 - p
  alpha_k <- 1 - (1-alpha/2)^(1/k)
  q0 <- qnorm(1 - alpha_k) # valor inicial
  if (q0 < 0) q0 <- -q0
  maxIt <- 5000
  tol <- 1e-13
  conv <- F
  it <- 0
  while ((conv == F) & (it <= maxIt))
  {
    q1 <- q0 - (pNDBD(q0, r, delta, n) - p)/dNDBD(q0, r, delta, n)
    if (abs(q1-q0) <= tol) conv <- T
    q0 <- q1
    it <- it + 1
  }
  return(q1)
}
```

```
#####
# Algoritmo implementado em R para obter a função de distribui- #
# ção para o caso unilateral não-balanceado e não-central, com #
# nu < Inf #
#####

pNDUDF <- function(q, r, nu, delta, n = 32)
{
  k <- length(r)
  xx <- gauss.quad(n, kind="legendre")
  x <- xx$nodes
  w <- xx$weights
  # intervalo de integração de 0 a 1
  y <- 0.5 * x + 0.5 # de -1 a 1 para 0 a 1
  fx <- nu/2 * log(nu) - lgamma(nu/2) - (nu/2-1)*log(2) + (nu-1)*
    log(y) - nu * y * y / 2
  fx <- exp(fx)
  y <- matrix(y, n, 1) * q
  fy <- 0.5 * apply(y, 1, pNDUD, r, delta, n) * fx
  I <- sum(fy * w)
  # intervalo de integração de 1 a Infy - transformação
  # y <- -ln(x), 0 < x < exp(-1)
  b <- exp(-1)
  a <- 0
  x <- (b - a) / 2 * x + (b + a) / 2 # de -1 a 1 para 0 a exp(-1)
  y <- -log(x) # de 0 a exp(-1) para 1 a +infty
  fx <- nu/2*log(nu) - lgamma(nu/2) - (nu/2-1) * log(2) + (nu-1)*
    log(y) - nu * y * y / 2
  fx <- exp(fx)
  y <- matrix(y, n, 1) * q
  fy <- (b - a) / 2 * apply(y, 1, pNDUD, r, delta, n) / (x)*fx
  I <- I + sum(fy * w)
  return(I)
}
```

```
#####
# Algoritmo implementado em R para obter a função de distribui- #
# ção para o caso bilateral não-balanceado e não-central, com #
# nu < Inf #
#####

pNDBDF <- function(q, r, nu, delta, n = 32)
{
  k <- length(r)
  xx <- gauss.quad(n, kind="legendre")
  x <- xx$nodes
  w <- xx$weights
  # intervalo de integração de 0 a 1
  y <- 0.5 * x + 0.5 # de -1 a 1 para 0 a 1
  fx <- nu/2*log(nu) - lgamma(nu/2) - (nu/2-1) * log(2) + (nu-1)*
    log(y) - nu * y * y / 2
  fx <- exp(fx)
  y <- matrix(y, n, 1) * q
  fy <- 0.5 * apply(y, 1, pNDBD, r, delta, n) * fx
  l <- sum(fy * w)
  # intervalo de integração de 1 a Infty - transformação
  # y <- -ln(x), 0 < x < exp(-1)
  b <- exp(-1)
  a <- 0
  x <- (b - a) / 2 * x + (b + a) / 2 # de -1 a 1 para 0 a exp(-1)
  y <- -log(x) # de 0 a exp(-1) para 1 a +infy
  fx <- nu/2*log(nu) - lgamma(nu/2) - (nu/2-1)*log(2) + (nu-1)*
    log(y) - nu * y * y / 2
  fx <- exp(fx)
  y <- matrix(y, n, 1) * q
  fy <- (b - a) / 2 * apply(y, 1, pNDBD, r, delta, n) / x * fx
  l <- l + sum(fy * w)
  return(l)
}
```

```
#####
# Algoritmo implementado em R para obter a função densidade para#
# o caso unilateral não-balanceado e não-central, com nu < Inf #
#####

dNDUDF <- function(q, r, nu, delta, n = 32)
{
  k <- length(r)
  xx <- gauss.quad(n, kind="legendre")
  x <- xx$nodes
  w <- xx$weights
  # intervalo de integração de 0 a 1
  y <- 0.5 * x + 0.5 # de -1 a 1 para 0 a 1
  fx <- nu/2 * log(nu) - lgamma(nu/2) - (nu/2-1) * log(2) +
    (nu-1) * log(y) - nu * y * y / 2 + log(y) # + jacobiano
  fx <- exp(fx)
  y <- matrix(y, n, 1) * q
  fy <- 0.5 * apply(y, 1, dNDUD, r, delta, n) * fx
  l <- sum(fy * w)
  # intervalo de integração de 1 a Inf - transformação
  # y <- -ln(x), 0 < x < exp(-1)
  b <- exp(-1)
  a <- 0
  x <- (b - a) / 2 * x + (b + a) / 2
  # de -1 a 1 para 0 a exp(-1)
  y <- -log(x) # de 0 a exp(-1) para 1 a +infy
  fx <- nu/2 * log(nu) - lgamma(nu/2) - (nu/2-1) * log(2) +
    (nu-1) * log(y) - nu * y * y / 2 + log(y) # + jacobiano
  fx <- exp(fx)
  y <- matrix(y, n, 1) * q
  fy <- (b - a) / 2 * apply(y, 1, dNDUD, r, delta, n) / x * fx
  l <- l + sum(fy * w)
  return(l)
}
```

```
#####
# Algoritmo implementado em R para obter a função densidade para#
# o caso bilateral não-balanceado e não-central, com nu < Inf #
#####

dNDBDF <- function(q, r, nu, delta, n = 32)
{
  k <- length(r)
  xx <- gauss.quad(n, kind="legendre")
  x <- xx$nodes
  w <- xx$weights
  # intervalo de integração de 0 a 1
  y <- 0.5 * x + 0.5 # de -1 a 1 para 0 a 1
  fx <- nu/2 * log(nu) - lgamma(nu/2) - (nu/2-1) * log(2) +
  (nu-1) * log(y) - nu * y * y / 2 + log(y) # + jacobiano
  fx <- exp(fx)
  y <- matrix(y, n, 1) * q
  fy <- 0.5 * apply(y, 1, dNDBD, r, delta, n) * fx
  I <- sum(fy * w)
  # intervalo de integração de 1 a Infy - transformação
  # y <- -ln(x), 0 < x < exp(-1)
  b <- exp(-1)
  a <- 0
  x <- (b - a) / 2 * x + (b + a) / 2 # de -1 a 1 para 0 a exp(-1)
  y <- -log(x) # de 0 a exp(-1) para 1 a +infy
  fx <- nu/2 * log(nu) - lgamma(nu/2) - (nu/2-1) * log(2) +
  (nu-1) * log(y) - nu * y * y / 2 + log(y) # + jacobiano
  fx <- exp(fx)
  y <- matrix(y, n, 1) * q
  fy <- (b - a) / 2 * apply(y, 1, dNDBD, r, delta, n) / x * fx
  I <- I + sum(fy * w)
  return(I)
}
```

```
#####
# Algoritmo implementado em R para obter a inversa da função de #
# distribuição para o caso unilateral não-balanceado e não-cen- #
# tral, com nu < infty #
#####

qNDUDF <- function(p, r, nu, delta, n = 32)
{
  k <- length(r)
  alpha <- 1 - p
  alpha_k <- 1 - (1-alpha)^(1/k)
  q0 <- qt(1 - alpha_k, nu) # valor inicial
  maxIt <- 5000
  tol <- 1e-13
  conv <- F
  it <- 0
  while ((conv == F) & (it <= maxIt))
  {
    q1 <- q0 - (pNDUDF(q0, r, nu, delta, n) - p) / dNDUDF(q0, r, nu, delta, n)
    if (abs(q1 - q0) <= tol) conv <- T
    q0 <- q1
    it <- it + 1
  }
  return(q1)
}
```

```
#####
# Algoritmo implementado em R para obter a inversa da função de #
# distribuição para o caso bilateral não-balanceado e não-cen- #
# tral, com nu < infty #
#####

qNDBDF <- function(p, r, nu, delta, n = 32)
{
  k <- length(r)
  alpha <- 1 - p
  alphak <- 1 - (1-alpha/2)^(1/k)
  q0 <- qt(1 - alphak, nu) # valor
  if (q0 < 0) q0 <- -q0
  maxIt <- 5000
  tol <- 1e-13
  conv <- F
  it <- 0
  while ((conv == F) & (it <= maxIt))
  {
    q1 <- q0 - (pNDBDF(q0, r, nu, delta, n) - p) / dNDBDF(q0, r, nu, delta, n)
    if (abs(q1 - q0) <= tol) conv <- T
    q0 <- q1
    it <- it + 1
  }
  return(q1)
}
```

ANEXO A - Inferência básica

Probabilidade

Tomando Magalhães (2006) por referência, pode-se calcular a probabilidade de uma variável estar num certo intervalo $(a, b]$, pela integração da função de densidade da variável naquele intervalo ou então pela diferença da função de distribuição da variável²⁸ calculada no limite superior do intervalo b e no limite inferior do intervalo a , ou seja:

$$P[a < X \leq b] = \int_a^b f(x)dx = F(b) - F(a).$$

A inclusão ou não dos valores a e b não altera o valor da integral e, portanto, a probabilidade, o que acarreta que para as variáveis aleatórias contínuas a probabilidade da variável ser igual a um particular valor é zero.

Esperança condicional

Kolmogorov (1956), Mood, Graybill e Boes (1974), Magalhães (2006), dentre outros, definem a probabilidade condicional de um evento A qualquer, por exemplo $Y \leq y$, dada uma variável aleatória $X = x$ discreta como:

$$P[A|X = x] = \frac{P[A; X = x]}{P[X = x]},$$

com o evento A e a variável aleatória X bem definidos no mesmo espaço de probabilidades.

Mas se X é contínua, $P[A|X = x]$ não pode ser analogamente definida pois $P[X = x] = 0$. No entanto, se x é tal que o evento $(x - h < X < x + h)$ tem probabilidade positiva para algum $h > 0$, então, pode-se definir a probabilidade condicional do evento A dada uma variável aleatória contínua $X = x$ como o

²⁸Derivado do teorema fundamental do cálculo para função $f(x)$ que seja contínua em $[a, b]$ e sua primitiva $F(x)$ seja conhecida.

limite de probabilidades condicionadas calculadas na vizinhança de x , ou seja:

$$P[A|X = x] = \lim_{0 < h \rightarrow 0} P[A|x - h < X < x + h]$$

se este limite existir.

Segundo Magalhães (2006), usando a ideia de probabilidade condicional, o conceito de função de distribuição pode ser estendido, considerando o condicionamento a um evento. Assim, a definição de função de distribuição condicional de A , dado que ocorreu X , com $P(X) > 0$, é:

$$F(a|X) = P(A \leq a|X) = \frac{P([A \leq a] \cap X)}{P(X)}.$$

Assim, se A e X são ambas variáveis contínuas define-se a densidade condicional de A dado X , como:

$$f_{A|X}(a|x) = \frac{f_{A,X}(a,x)}{f_X(x)},$$

cujas expressão é a densidade conjunta de A e X dividida pela densidade marginal de X .

O condicionamento é preservado também na relação entre função densidade e função de distribuição, resultando como consequência, dentre outras, a expressão:

$$F_A(a) = \int_{-\infty}^{\infty} f_X(x) F_{A|X}(a|x) dx, \text{ com } -\infty < x < \infty.$$

Mas $F(a|X) = P(A \leq a|X)$, então:

$$F_A(a) = \int_{-\infty}^{\infty} P(A \leq a|X) f_X(x) dx.$$

Em concordância com Mood, Graybill e Boes (1974), Magalhães (2006) também propõe usar a probabilidade condicional $P[A|X = x]$ para calcular outras probabilidades, como a probabilidade total de A (considerando que $F_A(a) = P(A \leq$

$a) = P[A]$), ou seja:

$$P[A] = \int_{-\infty}^{\infty} P[A|X = x]f(x)dx,$$

com X uma variável aleatória contínua.

Magalhães (2006) também define a esperança condicional como a esperança calculada com a probabilidade condicional. Assim, com A e X variáveis aleatórias no mesmo espaço de probabilidade, a esperança condicional de A dado que $X = x$, se existir, é dada por:

$$E(A|X = x) = \int_{-\infty}^{\infty} a \partial F_{A|X}(a|X = x).$$

Ross (2007) lembra que esta esperança condicional de A é definida para todos os valores de x tais que $f_x(x) > 0$.

Magalhães (2006) ainda diz que se $E(A)$ é finita, o Teorema de Radon-Nikodiyim garante a existência da esperança condicionada à X . Ressalta que a expressão $E(A|X = x)$ é uma função de x , frequentemente denominada de *regressão de A em relação à X*, e corresponde a média de valores de A , restrita aos valores de X iguais a x . O autor ainda complementa que caso não seja particularizado o valor de X , escreve-se $E(A|X)$, denominando-a de *esperança condicional de A dado X*, que é uma função da variável aleatória X e, portanto, é também uma variável aleatória, no mesmo espaço de probabilidade.

Outro autor, Meyer (1984) diz que só tem sentido falar de esperança condicional porque $E(A|X)$ é uma variável aleatória. Ele interpreta a esperança condicional $E(A|X = x)$ como pelo fato de que $\partial F_{A|X}(a|X = x)$ representa a função de densidade de probabilidade de A para um dado $X = x$, $E(A|x)$ é a esperança de A , condicionado ao evento $X = x$. Ross (2007) reforça que $E(A|X = x)$ é uma função da variável aleatória X cujo valor para $X = x$ é a $E(A|X = x)$. Ele também afirma e prova para o caso de X ser discreta, que:

$$E(A) = \int_{-\infty}^{\infty} E(A|X = x)f_X(x)dx.$$

Além disso Mood, Graybill e Boes (1974) afirmam que se $A = \{h(X,Y) \leq$

$z\}$ então:

$$P[A|X = x] = P[h(X, Y) \leq z|X = x] = P[h(x, Y) \leq z|X = x],$$

ou seja,

$$P[h(x, Y) \leq z] = \int_{-\infty}^{\infty} P[h(x, Y) \leq z|X = x]f(x)dx.$$

Assim, essa expressão é a esperança condicional de $P[h(x, Y) \leq z]$.

Por exemplo, tomando X e Y variáveis aleatórias independentes e $Z = X + Y$, então:

$$\begin{aligned} F_Z(z) &= P[Z \leq z] = P[X + Y \leq z] = \int_{-\infty}^{\infty} P[X + Y \leq z|X = x]f(x)dx \\ P[Z \leq z] &= P[X + Y \leq z] = \int_{-\infty}^{\infty} P[x + Y \leq z]f(x)dx \\ P[Z \leq z] &= P[X + Y \leq z] = \int_{-\infty}^{\infty} P[Y \leq (z - x)]f(x)dx \\ P[Z \leq z] &= P[X + Y \leq z] = \int_{-\infty}^{\infty} F_Y(z - x)f(x)dx, \end{aligned}$$

em que $F_Y(y = (z - x))$ é a função de distribuição de probabilidades da variável aleatória Y no ponto $y = z - x$.

ANEXO B - Estrutura correlação produto

Este anexo tem por objetivo esclarecer como se obtém a expressão para determinar o coeficiente de correlação entre as diferenças de médias do tratamento controle e dos tratamentos em teste, ou seja, a $Corr[(\bar{Y}_i - \bar{Y}_{r+1}), (\bar{Y}_{i'} - \bar{Y}_{r+1})]$, representado por $\rho_{ii'}$, para $i, i' = 1, 2, \dots, r$, com $i \neq i'$.

Genericamente, quando a partir de uma população $Y \sim N(\mu_y, \sigma_y^2)$ toma-se uma amostra aleatória de tamanho n , $Y_1, Y_2, \dots, Y_n$, o estimador não viesado da variância populacional $\hat{\sigma}_y^2$ é dado por

$$\hat{\sigma}_y^2 = \frac{\sum_{k=1}^n (Y_k - \bar{Y})^2}{n - 1}.$$

A média amostral $\bar{Y}$ é independente da variância amostral estimada $\hat{\sigma}_y^2$.

Se forem tomadas duas amostras independentes de tamanhos n_A e n_B de duas populações normais A e B , os estimadores da média e da variância para estas populações são, respectivamente, $\bar{Y}_A$ e $\hat{\sigma}_A^2$ e $\bar{Y}_B$ e $\hat{\sigma}_B^2$.

Pelo fato das amostras serem independentes, a combinação linear dada pela diferença de suas médias, $\bar{Y}_A - \bar{Y}_B$, também tem distribuição normal, com média $\mu_{\bar{Y}_A - \bar{Y}_B} = \mu_A - \mu_B$. A variância da diferença das médias amostrais é igual à soma das variâncias de cada média, como mostrado a seguir.

$$\begin{aligned} \sigma_{(\bar{Y}_A - \bar{Y}_B)}^2 &= E \{ [(\bar{Y}_A - \bar{Y}_B) - (\mu_A - \mu_B)]^2 \} \\ &= E \{ [(\bar{Y}_A - \mu_A) - (\bar{Y}_B - \mu_B)]^2 \} \\ &= E \{ (\bar{Y}_A - \mu_A)^2 - 2(\bar{Y}_A - \mu_A)(\bar{Y}_B - \mu_B) + (\bar{Y}_B - \mu_B)^2 \} \\ &= E[(\bar{Y}_A - \mu_A)^2] - 2E[(\bar{Y}_A - \mu_A)(\bar{Y}_B - \mu_B)] + E[(\bar{Y}_B - \mu_B)^2]. \end{aligned}$$

Por definição, a covariância entre as variáveis aleatórias $\bar{Y}_A$ e $\bar{Y}_B$ é dada por $Cov(\bar{Y}_A, \bar{Y}_B) = E[(\bar{Y}_A - \mu_A)(\bar{Y}_B - \mu_B)]$ (MAGALHÃES; LIMA, 2005). Como

$\bar{Y}_A$ e $\bar{Y}_B$ são variáveis aleatórias independentes esta covariância é nula já que:

$$\begin{aligned}
 Cov(\bar{Y}_A, \bar{Y}_B) &= E[(\bar{Y}_A - \mu_A)(\bar{Y}_B - \mu_B)] \\
 &= E[\bar{Y}_A \bar{Y}_B - \bar{Y}_A \mu_B - \mu_A \bar{Y}_B + \mu_A \mu_B] \\
 &= E(\bar{Y}_A \bar{Y}_B) - \mu_B E(\bar{Y}_A) - \mu_A E(\bar{Y}_B) + \mu_A \mu_B \\
 &= \mu_A \mu_B - \mu_B \mu_A - \mu_A \mu_B + \mu_A \mu_B \\
 &= 0.
 \end{aligned}$$

Assim

$$\sigma_{(\bar{Y}_A - \bar{Y}_B)}^2 = E[(\bar{Y}_A - \mu_A)^2] + E[(\bar{Y}_B - \mu_B)^2] = \sigma_{\bar{Y}_A}^2 + \sigma_{\bar{Y}_B}^2,$$

Da teoria das distribuições amostrais, a distribuição amostral da média amostral de uma população normal com média μ e variância finita σ^2 , obtida a partir de amostras de tamanho n , é normal com média $\mu_{\bar{X}} = \mu$ e variância $\hat{\sigma}_{\bar{X}}^2 = \sigma^2/n$ (FERREIRA, 2009). Assim

$$\sigma_{\bar{Y}_A}^2 = \frac{\sigma_A^2}{n_A} \text{ e } \sigma_{\bar{Y}_B}^2 = \frac{\sigma_B^2}{n_B}.$$

No caso de homocedasticidade, ou seja $\sigma_A^2 = \sigma_B^2 = \sigma^2$, tem-se que:

$$\sigma_{\bar{Y}_A - \bar{Y}_B}^2 = \sigma^2 \left(\frac{1}{n_A} + \frac{1}{n_B} \right).$$

Nestas mesmas condições de populações e amostras, tomam-se os estimadores $\bar{Y}_C$ e $\hat{\sigma}_C^2$ de uma população normal C . Para determinar a correlação entre duas variáveis aleatórias dadas pela diferença entre médias, sendo uma das médias $\bar{Y}_C$ comuns em cada diferença, ou seja $Corr[(\bar{Y}_A - \bar{Y}_C), (\bar{Y}_B - \bar{Y}_C)]$, tem-se

$$Corr[(\bar{Y}_A - \bar{Y}_C), (\bar{Y}_B - \bar{Y}_C)] = \frac{Cov[(\bar{Y}_A - \bar{Y}_C), (\bar{Y}_B - \bar{Y}_C)]}{\sqrt{\sigma_{\bar{Y}_A - \bar{Y}_C}^2} \sqrt{\sigma_{\bar{Y}_B - \bar{Y}_C}^2}}.$$

A $Cov[(\bar{Y}_A - \bar{Y}_C), (\bar{Y}_B - \bar{Y}_C)]$ é dada por:

$$\begin{aligned}
 &= E\{[(\bar{Y}_A - \bar{Y}_C) - (\mu_A - \mu_C)][(\bar{Y}_B - \bar{Y}_C) - (\mu_B - \mu_C)]\} \\
 &= E\{[(\bar{Y}_A - \mu_A) - (\bar{Y}_C - \mu_C)][(\bar{Y}_B - \mu_B) - (\bar{Y}_C - \mu_C)]\} \\
 &= E[(\bar{Y}_A - \mu_A)(\bar{Y}_B - \mu_B) - (\bar{Y}_A - \mu_A)(\bar{Y}_C - \mu_C) - (\bar{Y}_C - \mu_C)(\bar{Y}_B - \mu_B) \\
 &\quad + (\bar{Y}_C - \mu_C)^2] \\
 &= E[(\bar{Y}_A - \mu_A)(\bar{Y}_B - \mu_B)] - E[(\bar{Y}_A - \mu_A)(\bar{Y}_C - \mu_C)] - E[(\bar{Y}_C - \mu_C)(\bar{Y}_B - \mu_B)] \\
 &\quad + E[(\bar{Y}_C - \mu_C)^2] \\
 &= Cov(\bar{Y}_A, \bar{Y}_B) - Cov(\bar{Y}_A, \bar{Y}_C) - Cov(\bar{Y}_C, \bar{Y}_B) + E[(\bar{Y}_C - \mu_C)^2].
 \end{aligned}$$

Como $\bar{Y}_A$, $\bar{Y}_B$ e $\bar{Y}_C$ são independentes, a covariância entre cada par de médias é nula. Assim,

$$Cov[(\bar{Y}_A - \bar{Y}_C), (\bar{Y}_B - \bar{Y}_C)] = E[(\bar{Y}_C - \mu_C)^2] = \sigma_{\bar{Y}_C}^2 = \frac{\sigma^2}{n_C},$$

e, portanto, a correlação entre as duas diferenças de médias é dada por

$$Corr[(\bar{Y}_A - \bar{Y}_C), (\bar{Y}_B - \bar{Y}_C)] = \frac{\frac{\sigma^2}{n_C}}{\sqrt{\sigma_{\bar{Y}_A - \bar{Y}_C}^2} \sqrt{\sigma_{\bar{Y}_B - \bar{Y}_C}^2}}.$$

Como as populações são homocedásticas, chega-se a

$$Corr[(\bar{Y}_A - \bar{Y}_C), (\bar{Y}_B - \bar{Y}_C)] = \frac{\frac{\sigma^2}{n_C}}{\sqrt{\sigma^2 \left(\frac{1}{n_A} + \frac{1}{n_C} \right)} \sqrt{\sigma^2 \left(\frac{1}{n_B} + \frac{1}{n_C} \right)}}$$

que também pode ser escrita como:

$$\begin{aligned}
 Corr[(\bar{Y}_A - \bar{Y}_C), (\bar{Y}_B - \bar{Y}_C)] &= \frac{\frac{1}{\sqrt{n_C}}}{\sqrt{\left(\frac{1}{n_A} + \frac{1}{n_C} \right)}} \frac{\frac{1}{\sqrt{n_C}}}{\sqrt{\left(\frac{1}{n_B} + \frac{1}{n_C} \right)}} \\
 &= \frac{1}{\sqrt{1 + \frac{n_C}{n_A}}} \frac{1}{\sqrt{1 + \frac{n_C}{n_B}}}.
 \end{aligned}$$

Esta correlação tem uma estrutura especial denominada de correlação-

produto.

Assim, voltando ao contexto da determinação do coeficiente de correlação entre o contraste do i -ésimo tratamento em teste com o controle e o contraste do i' -ésimo tratamento em teste com o controle, com $i, i' = 1, 2, \dots, r$ e $i \neq i'$, tem-se a expressão genérica para a correlação entre cada par de contrastes

$$\text{Corr}[(\bar{Y}_i - \bar{Y}_{r+1}), (\bar{Y}_{i'} - \bar{Y}_{r+1})] = \frac{1}{\sqrt{1 + \frac{n_{r+1}}{n_i}} \sqrt{1 + \frac{n_{r+1}}{n_{i'}}}},$$

em que n_i e $n_{i'}$ são o número de repetições do i -ésimo e do i' -ésimo tratamento em teste, respectivamente, e n_{r+1} o número de repetições do tratamento controle.

Chamando de ρ_i e $\rho_{i'}$ os termos da fatoração desta correlação, pode-se escrever:

$$\text{Corr}[(\bar{Y}_i - \bar{Y}_{r+1}), (\bar{Y}_{i'} - \bar{Y}_{r+1})] = \sqrt{\rho_i} \sqrt{\rho_{i'}}.$$

Assim, genericamente, cada termo ρ_i é dado por

$$\begin{aligned} \rho_i &= \frac{\frac{\sigma^2}{n_{r+1}}}{\frac{\sigma^2}{n_i} + \frac{\sigma^2}{n_{r+1}}} \\ &= \frac{\frac{1}{n_{r+1}}}{\frac{1}{n_i} + \frac{1}{n_{r+1}}} \\ &= \frac{1}{1 + \frac{n_{r+1}}{n_i}} \text{ para } i = 1, 2, \dots, r. \end{aligned}$$

De forma matricial, denotando $\sqrt{\rho_i} = \lambda_i$, obtém-se a matriz de correlações entre as r comparações como

$$\boldsymbol{\rho} = \begin{pmatrix} 1 & \lambda_1 \lambda_2 & \lambda_1 \lambda_3 & \dots & \lambda_1 \lambda_r \\ \lambda_2 \lambda_1 & 1 & \lambda_2 \lambda_3 & \dots & \lambda_2 \lambda_r \\ \lambda_3 \lambda_1 & \lambda_3 \lambda_2 & 1 & \dots & \lambda_3 \lambda_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_r \lambda_1 & \lambda_r \lambda_2 & \lambda_r \lambda_3 & \dots & 1 \end{pmatrix}.$$

Observa-se que a expressão do termo ρ_i está relacionada com média harmônica dos valores dos tamanhos amostrais das duas médias envolvidas (n_i, n_{r+1}) , pois

$$\overline{X}_{H(n_i, n_{r+1})} = \frac{2}{\frac{1}{n_i} + \frac{1}{n_{r+1}}}$$

pode ser escrita como

$$\frac{\overline{X}_{H(n_i, n_{r+1})}}{2} = \frac{1}{\frac{1}{n_i} + \frac{1}{n_{r+1}}}$$

que é igual a $\rho_i \times n_{r+1}$.