

Desenvolvimento de uma nova função no MS Excel para avaliação de diferenças entre grupos amostrais

D.C.R. Velasco^{1*}; L.F.F. Freitas¹; F.P.D. Lopes; C.M.F. Vieira¹

¹Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF)

*davidc.r.v2014@gmail.com

Resumo

A análise estatística desempenha um papel crucial em diversas áreas, incluindo a engenharia, onde auxilia na tomada de decisões imparciais para otimizar processos produtivos. A Análise de Variância (ANOVA) é frequentemente utilizada para comparar grupos amostrais, permitindo a avaliação simultânea de várias médias, embora não identifique quais grupos diferem entre si. Testes pós-hoc, como o Teste de Tukey, resolvem esse problema, mas sua confiabilidade diminui com grupos maiores. Neste trabalho, foi desenvolvida uma nova função para o Microsoft Excel, aproveitando de integração com o Virtual Basic for Applications (VBA), para simplificar as comparações diretas entre grupos amostrais. Esta função possibilita realizar uma ANOVA para dois grupos amostrais com base no tamanho, média e desvio padrão fornecidos, gerando como resultado o valor P, que pode ser comparado diretamente com o nível de significância.

Palavras-chave: Análise de Variância; Teste de Tukey; Microsoft Excel.

1. Introdução

A análise estatística desempenha um papel essencial em diversas áreas. Na engenharia, ela possibilita realizar avaliação sem juízos de valor. Essa análise pode ser utilizada, por exemplo, para avaliar a viabilidade de etapas de processamento ou o uso de aditivos, permitindo a otimização dos processos produtivos. Assim, análises estatísticas, como a Análise de Variância (ANOVA), são amplamente empregadas para a comparação de grupos amostrais^[1].

Esta tem vantagens em relação a métodos como Teste t de Student, pois possibilita a comparação de mais de um grupo amostral ao mesmo tempo. Esta realiza uma avaliação entre as médias dos grupos (variância entre grupos) em comparação com a variância média dentro dos grupos^[2].

Contudo, a ANOVA não fornece detalhes sobre quais grupos específicos diferem entre si. Para esse fim, os pesquisadores recorrem a testes de comparação múltipla pós-hoc para mais de dois grupos, como o Teste de Tukey^[3]. Este é uma técnica que se destaca por permitir que em uma única etapa seja realizada uma análise comparativa entre inúmeros grupos amostrais, verificando se há a existência de uma diferença mínima significativa entre estes^[4].

Contudo, embora estes testes sejam normalmente usados de forma complementar, os resultados deste podem ser diferentes caso sejam utilizadas comparações em grupos amostrais superiores a dois. O que indica que uma comparação par a par tem maior confiabilidade^[5].

No âmbito do desenvolvimento de ferramentas sociais, o Excel, em conjunto com o Visual Basic for Applications (VBA), sobressai devido à sua habilidade de executar ações e à sua integração com as funções nativas do primeiro, proporcionando uma simplificação notável no

processo de programação. Importante destacar que, apesar de ser uma ferramenta licenciada, sua ampla disponibilidade abrange bilhões de computadores [6].

Neste sentido, este trabalho tem por objetivo desenvolver uma nova função para o Excel que possibilita realizar comparações pares a pares de grupos amostrais de uma forma simplificada.

2. Materiais e Métodos

2.1. Materiais

Neste estudo foi utilizado o Microsoft (MS) Excel, versão *Home and Student* 2016. A Linguagem de programação adotada foi VBA, utilizando as referências padrões do sistema. Isto é, “*Visual Basic for Applications*”; “*Microsoft Excel 16.0 Object Library*”; “*OLE Automation*” e “*Microsoft Office 16.0 Object Library*”.

2.2. Metodologia

O algoritmo foi desenvolvido de acordo com a simplificação e reorganização das equações de ANOVA presentes na literatura. Uma vez calculado o valor F, foi calculado o valor P com a função nativa do MS Excel, conforme a Figura 1.

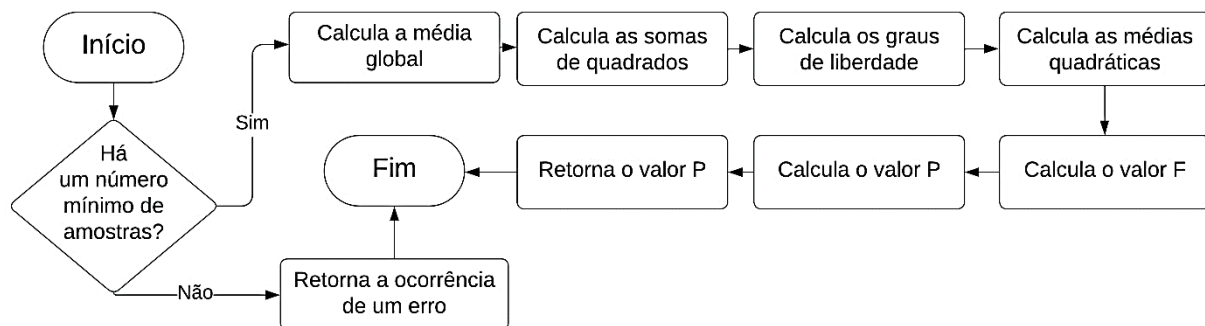


Figura 1: Fluxograma do algoritmo elaborado.

3. Resultados

O algoritmo desenvolvido pode ser observado na Figura 2. Nesta, pode-se verificar as etapas do algoritmo de forma correlacionada com a Figura 1. A função desenvolvida foi denominada como ANOVA_2V e recebe como parâmetros as seguintes variáveis: quantidade de amostras em cada grupo amostral (CNTS); média do grupo amostral 1 (M1); média do grupo amostral 1 (M2); desvio padrão do grupo amostral 1 (STD1); desvio padrão do grupo amostral 1 (STD2).

```

Function ANOVA_2V(ByVal Cnts As Integer, m1 As Double, m2 As Double, STD1 As Double, STD2 As Double)

'Verifica se há pelo menos 4 amostras em cada grupo
If Cnts < 4 Then
    ANOVA_2V = CVErr(xlErrNA)
    Exit Function
End If

'Declara as variáveis
Dim GlobalMean, SSw, SSb, DFw, MSb, MSw, F As Double

'Calcula a média global
GlobalMean = (m1 + m2) / 2

'Calcula as somas de quadrados
SSw = (Cnts - 1) * (STD1 * STD1 + STD2 * STD2)
SSb = Cnts * ((m1 - GlobalMean) * (m1 - GlobalMean) + (m2 - GlobalMean) * (m2 - GlobalMean))

'Calcula os graus de liberdade
DFw = (Cnts - 1) * 2

'Calcula as médias quadráticas
MSb = SSb
MSw = SSw / DFw

'Calcula o valor F
F = MSb / MSw

'Calcula e retorna o valor P
ANOVA_2V = Application.FDist(F, 1, DFw)

End Function

```

Figura 2: Função produzida neste trabalho.

O desenvolvimento de tal algoritmo, permitiu a utilização conjunta com funções nativas do MS Excel. A Figura 3 (a), apresenta uma tabela de resultados que foram obtidos em por meio de ensaios laboratoriais, enquanto a Figura 3 (b) apresenta a análise que fora realizada usando a função criada em conjunto com outras recursos nativos do MS Excel (SE, PROCV e formatação condicional). A última, foi resultado de uma única formula que fora elaborada em uma célula e preenchida automaticamente para as demais células do Excel.

ID	Média	Desvio		A	B	C	D	E
A	2,037	0,075	A		0%	0%	1%	80%
B	1,554	0,194	B	0%		31%	7%	0%
C	1,655	0,071	C	0%	31%		11%	0%
D	1,777	0,134	D	1%	7%	11%		2%
E	2,020	0,129	E	80%	0%	0%	2%	

(a)

(b)

Figura 3: Experimentação da função criada.

4. Conclusões

Por meio deste trabalho pode-se desenvolver uma função intitulada "ANOVA_2V". Esta não só possibilitou executar com sucesso a comparação de grupos amostrais, mas também demonstrou capacidade de reforçar o ecossistema do MS Excel e atuar de forma conjunta com os seus recursos nativos.

A função desenvolvida apresenta como vantagem em relação às ferramentas nativas do MS Excel a simplificação dos processos de análise, reduzindo a quantidade de etapas necessárias. Além disso, sua capacidade de integração com outras fórmulas do Excel permite a execução eficiente de análises par-a-par, como evidenciado na Figura 3. O que implica na possibilidade de realizar em uma etapa de diversos grupos amostrais sem prejuízo de confiabilidade das análises.

Agradecimentos

A FAPERJ (proc. n. E-26/202.773/2017 e E-26/203.874/2022) e CNPQ (proc. n. 301634/2018-1) pelo apoio financeiro.

Referências

- [1] VELASCO, D. C. R.; LOPES, F. P. D.; LOIOLA, R. L.; SIMONASSI, N. T.; VIEIRA, C. M. F. **Influência do aditivo anti-bolha na resistência à compressão de compósitos particulados de matriz epóxi**, p. 2199-2212. In: 76º Congresso Anual da ABM - Internacional, São Paulo, 2023.
- [2] KIM, H. Y. Analysis of variance (ANOVA) comparing means of more than two groups. **Restorative dentistry & endodontics**, v. 39, n. 1, p. 74-77, 2014.
- [3] TUKEY, J. W. Comparing individual means in the analysis of variance. **Biometrics**, p. 99-114, 1949.
- [4] BENJAMINI, Y.; BRAUN, H. John W. Tukey's Contributions to Multiple Comparisons. **Annals of Statistics**, p. 1576-1594, 2002.
- [5] GURVICH, V.; NAUMOVA, M. Logical Contradictions in the One-Way ANOVA and Tukey–Kramer Multiple Comparisons Tests with More Than Two Groups of Observations. **Symmetry**, v. 13, n. 8, p. 1387, 2021.
- [6] VELASCO, D. C. R.; CARVALHO, A. L. V.; MATIAS, Í. O. Criação e distribuição de documentos: proposta de utilização do visual basic for applications (vba) para automatização de tarefas. **InterSciencePlace**, v. 16, n. 3, 2021.