

Título:

Introdução ao Planejamento de Experimentos e Inferência em modelo de ANOVA.

Área: Estatística

Tipo de Iniciação Científica: PIBIC – Ensino Médio.

Fomento: CNPq

Autora: Alanie Leite Orcino - Colégio Estadual Professor Pedro Gomes, Goiânia/GO.

Prof^ª. Orientadora: Marta Cristina Colozza Bianchi – IME/UFG.

e-mail: marta_bianchi@mat.ufg.br; mccb_mat@yahoo.com.br.

Palavras-Chave: ANOVA, Delineamento Experimental, Inferência.

1. Introdução

Neste trabalho é apresentado um projeto de pesquisa de iniciação científica PIBIC – EM , em desenvolvimento, vinculado ao Instituto de Matemática e Estatística da Universidade Federal de Goiás. Tal projeto tem o objetivo de fortalecer o conhecimento científico, bem como desenvolver as atitudes, habilidades e valores necessários a educação científica de um aluno do ensino médio por meio do estudo da Estatística. O projeto contempla o estudo introdutório de Estatística com ênfase em Planejamento de Experimentos e ANOVA e, por fim, uma aplicação da teoria estudada a dados reais.

Uma das preocupações de um estatístico ao analisar um conjunto de dados é criar modelos que explicitem estruturas do fenômeno sob observação, as quais frequentemente estão misturadas a variações acidentais ou aleatórias. A identificação dessas estruturas permite conhecer melhor o fenômeno, bem como fazer afirmações sobre possíveis comportamentos.

Uma estratégia de análise é supor que cada observação seja formada por duas partes:

$$\textit{observação} = \textit{previsível} + \textit{aleatório}$$

A primeira parte incorpora o conhecimento que o pesquisador tem sobre o fenômeno é usualmente expressa por uma função matemática, com parâmetros desconhecidos. A segunda parte, a aleatória (não previsível), representa o que o

pesquisador não pode controlar e para a qual são impostas algumas suposições, como por exemplo que ela obedeça a algum modelo probabilístico específico.

Dentro desse cenário o trabalho do estatístico passa a ser o de estimar os parâmetros desconhecidos do modelo, baseado em um conjunto de dados, para posteriormente poder fazer previsões e tomar decisões relativas ao fenômeno estudado.

Um modelo simples particular é o modelo de análise de variância com um fator (ANOVA). A situação geral é descrita como segue. Temos uma população de unidades experimentais (indivíduos, animais, empresas etc.), para a qual temos uma variável aleatória Y de interesse. Tais unidades experimentais podem ser classificadas segundo níveis de um fator. Por exemplo, consideremos a população de alunos em uma escola. O fator pode ser o sexo, com dois níveis, arbitrariamente denotados por 1: sexo masculino e 2: sexo feminino. A variável aleatória Y pode ser a altura de cada indivíduo. Cada nível do fator representa uma subpopulação, no exemplo temos duas subpopulações: alunos do sexo masculino e alunos do sexo feminino. Em geral, um fator pode ser classificado em I níveis subpopulações.

O modelo de análise de variância que descreve essa população é definido por:

$$y_{ij} = \mu_i + e_{ij}, \quad i = 1, \dots, I; \quad j = 1, \dots, n_i$$

com n_i o número de observações do nível i , y_{ij} o valor da observação j do nível i , μ_i é o parâmetro desconhecido que representa a média do nível i , e e_{ij} é a parte aleatória (não previsível) do modelo. No exemplo temos dois níveis portanto $I=2$, μ_1 representa a altura média dos alunos e μ_2 a altura média das alunas. A estimação dos parâmetros desconhecidos é feita com base em amostras de todos os níveis: $(y_{11}, y_{12}, \dots, y_{1n_1})$ amostra de altura dos alunos e $(y_{21}, y_{22}, \dots, y_{2n_2})$ amostra de altura das alunas.

O objetivo principal da ANOVA é realizar comparações entre a média da variável Y para a população toda e suas médias em cada nível, permitindo assim verificar se os níveis causam algum efeito sobre a população, ou mesmo se existem diferenças entre os níveis. No exemplo, podemos nos perguntar: a altura dos alunos e das alunas é realmente diferente da altura global dos alunos? Ou, a altura de meninos e meninas é diferente? Quão significativamente diferente? Essas e outras perguntas, são respondidas pelo modelo de ANOVA.

O modelo de ANOVA é freqüentemente utilizado em Agronomia, juntamente empregado com delineamentos experimentais. Nesta área surge o termo *tratamento* equivalente ao nível de um fator. Por exemplo, em uma plantação há o interesse de se verificar se algum dos tipos de fertilizante entre os tipos A, B e C verte em algum efeito de produtividade para a plantação, ou se os fertilizantes não surtem efeito algum. Cada tipo de fertilizante é denominado como um *tratamento*. A saúde também é outra área de aplicação, confrontando eficácia de medicamentos por exemplo.

O estudo introdutório de estatística descritiva, probabilidade, estatística inferencial e delineamento de experimentos é de extrema importância para a compreensão do objeto de pesquisa ANOVA. O estudo de estatística descritiva, permiti compreender e manejar alguns métodos úteis para a descrição e comparação de dados. Conceitos de probabilidade e inferência, permitem tomar contato com o cálculo de probabilidades de eventos ou variável aleatória e, previsão e tomada de conclusões com base nos dados. Por fim, a utilização e manejo de planejamento de experimento e modelo de ANOVA permite realizar um experimento bem delineado e verificar se existe, e quão significativo é, o efeito produzido por uma subpopulação na população do experimento em estudo.

2. Metodologia

Uma vez que o aluno esta cursando o ensino médio, o mesmo não tem conhecimento sobre as disciplinas relacionados com o tópico de pesquisa. Assim, para atingir os objetivos do trabalho desenvolve-se um estudo dirigido dos temas: estatística descritiva, noções de probabilidade, noções de inferência estatística, delineamento experimental e modelo de ANOVA. Neste sentido, a professora orientadora fornece ao aluno um tema quinzenal a ser estudado e que no final da quinzena é apresentado pelo aluno e avaliado pela orientadora. Por fim, o aluno aplica a teoria estudada a dados reais.

4. Resultados

O objetivo principal do trabalho é desenvolver a capacidade de pesquisa do aluno. O aluno deve ser capaz de expor e exemplificar a teoria estudada por meio de apresentação em congresso científico. E, ao final do projeto, escrever sobre os tópicos estudados em forma de um pequeno artigo contendo introdução, fundamentação teórica e uma aplicação simples com dados reais usando o software estatístico gratuito *R* (www.r-project.org).

5. Referências Bibliográficas

1. Bussab, W.O. *Análise de Variância e de Regressão: uma introdução.*, 2a ed., Atual, 1988.
2. Bussab, W.O., Morettin, P. A. *Estatística Básica.*, 5a ed., Saraiva, 2004.
3. Callegare A.J.A. *Introdução ao delineamento de experimentos.*, 2a ed., Edgard Blucher, 2001.
4. Triola, M. F. *Introdução à Estatística.*, LTC, 2008.
5. Vieira, S. *Estatística Experimental.*, 2a ed., Atlas, 1999.
6. Walpole, R. E.; Myers, R. H.; Myers, S. L.; Ye K. *Probabilidade e Estatística para engenharia e ciências*, Pearson, 2009.