

Universidade Federal de Minas Gerais

Engenharia de Sistemas

Estudo de Caso 03: Emparelhamento de dados

Gustavo Vieira Costa - 2010022003

Marcus Vinícius - 2013030147

Rafael Castro - 2013030210

Thaís Matos Acácio - 2013030287

25/04/2016

1 Introdução

Uma percepção comum é a de que pessoas tendem a sistematicamente declarar um peso corporal inferior ao valor real. Com base nessa afirmação, iremos investigar o viés de relato de peso dos alunos de Engenharia de Sistemas, mediante comparação dos valores estimados pelos alunos da disciplina com valores aferidos por uma balança digital de uso doméstico.

O experimento realizado utiliza o conceito de emparelhamento de dados, no qual a hipótese a ser testada passa a ser relacionada com a diferença entre as médias da amostra.

2 Coleta de dados

A Tabela 1 contém a amostra de dados coletada, ou seja, o peso informado pelos alunos da disciplina e aferidos pela balança, juntamente com um identificador aleatório e único para cada um dos alunos:

Os valores dos pesos declarados e medidos foram exibidos no gráfico de blocos 1 para possibilitar uma melhor visualização do comportamento da amostra. O diagrama 2 retrata a diferença entre esses valores.

ID	Peso Estimado(kg)	Peso Medido(kg)
EngSis.A	78.4	79.6
EngSis.B	69.5	69.3
EngSis.C	66.2	66.4
EngSis.D	61.0	60.4
EngSis.E	59.0	57.7
EngSis.F	70.5	71.3

ID	Peso Estimado(kg)	Peso Medido(kg)
EngSis.G	63.5	63.3
EngSis.H	81.0	79.3
EngSis.I	57.7	53.5
EngSis.J	49.0	52.0
EngSis.K	68	71.7

Tabela 1: Tabela de Amostras

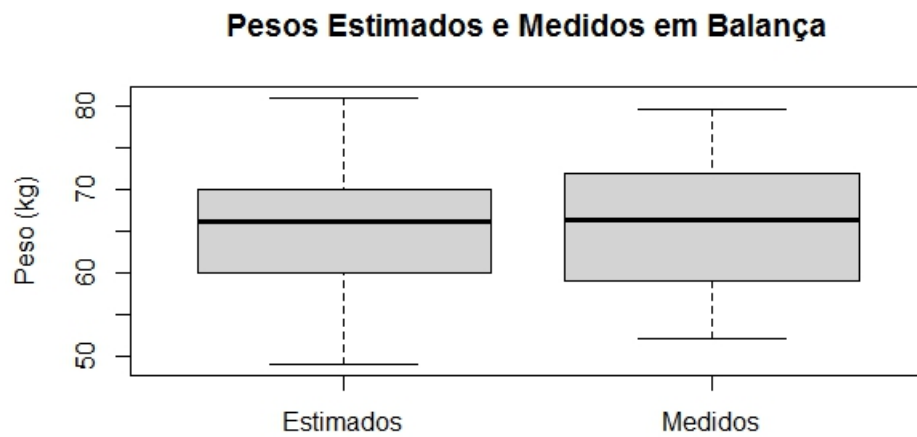


Figura 1: Pesos Estimados e Medidos

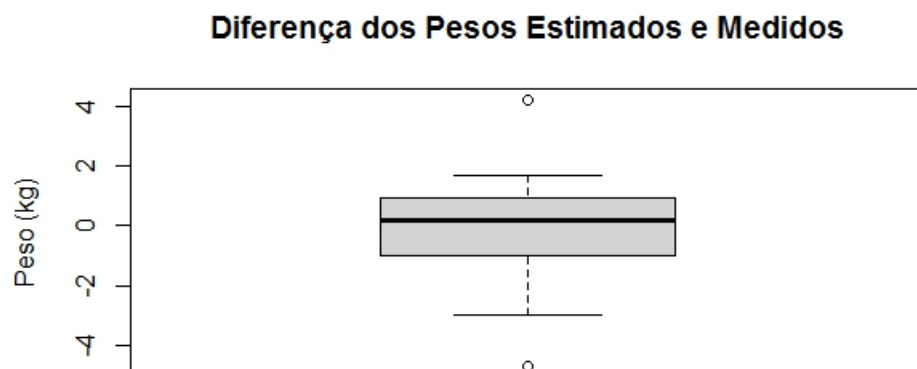


Figura 2: Diferença dos Pesos Estimados e Medidos

De acordo com o *Teorema do Limite Central*, se a amostra tiver tamanho n suficiente, a distribuição amostral de \bar{x} é aproximadamente Normal. Nesse caso, iremos assumir que n é suficiente, conforme o gráfico de normalidade presente na figura ??.

3 Estratégia de Inferência

O processo de inferência estatística consiste em tirar conclusões sobre uma população com base em informações extraídas de amostras da mesma. No presente estudo de caso, o parâmetro sobre o qual temos interesse é a média das diferenças entre os pesos medidos e estimados μ_D dos alunos de graduação em Engenharia de Sistemas.

O método se baseia em calcular a diferença entre cada par de pesos estimado e medido de cada aluno, determinar a média dessas diferenças e informar se essa média é estatisticamente significativa. Podemos utilizar o método de inferência para uma única amostra, visto que temos apenas duas observações de uma mesma amostra, as quais serão pareadas.

Premissas do Teste

- O **nível de significância**(α), representa a probabilidade de erro tipo I, ou seja, a probabilidade de rejeitarmos a hipótese nula quando ela é efetivamente verdadeira. Pensando em uma taxa de erro aceitável para o domínio do problema, fixamos $\alpha = 0.05$.
- β representa a **probabilidade de erro tipo II**, ou seja, aceitarmos a hipótese nula quando ela é efetivamente falsa. Resolvemos permitir um erro tipo II de 20% ($\beta = 0.2$), considerando que esse tipo de erro tem menor impacto negativo do que o erro tipo I.
- **Nível de confiança** ($1 - \alpha$) tem como objetivo conhecer o quanto o teste de hipóteses controla um erro do tipo I, ou qual a probabilidade de aceitar a hipótese nula se realmente for verdadeira.
- O **poder do teste** ($1 - \beta$) tem como objetivo conhecer o quanto o teste de hipóteses controla um erro do tipo II ou, qual a probabilidade de rejeitar a hipótese nula se realmente for falsa.
- O menor **tamanho de efeito** de importância prática(δ^*) como 0.5, considerando o peso médio das roupas e um erro de precisão aceitável para a balança.

Hipóteses de Teste

O teste estatístico é planejado para avaliar a força da evidência *contra* a hipótese nula H_0 . Usualmente, a hipótese nula é uma afirmativa de "nenhum efeito". A afirmativa sobre a população *a favor* da qual estamos tentando achar evidência é a hipótese alternativa H_1 . Logo, as hipóteses são:

$$\begin{cases} H_0 : \mu_D = 0 \\ H_1 : \mu_D \neq 0 \end{cases} \quad (1)$$

sendo μ_D a média das diferenças, entre os pesos estimados e medidos, dos alunos de Engenharia de Sistemas.

4 Projeto experimental e Análise dos Resultados

5 Conclusão

Referências

- [1] <https://github.com/fcampelo/Design-and-Analysis-of-Experiments>
- [2] Estatística Aplicada e Probabilidade para Engenheiros (4ª edição) - Montgomery
- [3] A Estatística Básica e Sua Prática (6ª edição) - David S. Moore, William I. Nortz, Michael A. Fligner
- [4] <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/var.test.html>
- [5] <http://ww2.coastal.edu/kingw/statistics/R-tutorials/independent-t.html>
- [6] <http://www.portalaction.com.br/inferencia/56-teste-para-comparacao-de-duas-variaveis>