

Universidade Federal de Minas Gerais
Engenharia de Sistemas
Estudo de Caso 1: Inferência Estatística Com Uma Amostra

Gustavo Vieira Costa - 2010022003

Rafael Castro - 2013030210

Thaís Matos Acácio - 2013030287

08/04/2016

1 Introdução

O BMI (*body mass index*, ou índice de massa corporal) é um indicador frequentemente usado em avaliações clínicas de questões relacionadas ao peso de um indivíduo. Este índice é calculado como a razão entre o peso e o quadrado da estatura.

O professor Felipe Campelo, do Departamento de Engenharia Elétrica da UFMG, reporta estar atualmente com um valor de $BMI = 26.3kg/m^2$. Neste estudo de caso vamos buscar responder à pergunta: *Os alunos do curso de Engenharia de Sistemas estão, em média, mais "acima do peso" (de acordo com o BMI) do que este professor?* Para isso, cada um dos alunos da disciplina forneceu seu peso e estatura de forma anonimizada, formando uma base de dados com a qual pretende-se realizar a inferência estatística a respeito da população.

2 Projeto experimental

O arquivo contendo o código do experimento se encontra disponível para download em <https://goo.gl/23J2Sp>. Para executá-lo, basta compilar o Script no RStudio e as informações completas sobre sua execução serão exibidas no console da aplicação.

Segue o código:

```
> mu <- 26.3;  
+ alpha <- 0.05;  
+  
+ #Ler dados de entrada  
+ dados <- read.table("data.csv", header=FALSE, sep=";");
```

```

+ #Calculo do BMI
+ BMI <- dados[1]/(dados[2]^2);
+
+ #Numero de Amostras
+ n <- nrow(BMI);
+
+ #Media
+ x_bar <- mean(as.matrix(BMI));
+
+ #Tamanho de Efeito
+ size_effect <- x_bar - mu;
+
+ #Desvio padrao
+ s <- sqrt(sum((BMI-x_bar)^2)/(n-1));
+
+ #t critico
+ t_alpha <- qt(alpha/2, n-1);
+
+ #Intervalo de confianca
+ inter_min <- x_bar + (s*t_alpha / (sqrt(n)));
+ inter_max <- x_bar - (s*t_alpha / (sqrt(n)));

```

3 Coleta de dados

A Tabela 1 contém a amostra de dados coletados, informados pelos alunos da turma, juntamente com o valor do índice BMI calculado utilizando a seguinte fórmula:

$$bmi = \frac{m}{h^2} \quad (1)$$

onde m é o peso dado em kg e h a altura dada em metros.

3.1 Inferência Estatística

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 μ do índice BMI da população.

Hipóteses de Teste

A hipótese nula H_0 assumida como verdade para construção do teste de hipóteses, é

Peso	Altura	BMI
48.0	1.56	19.72387
61.5	1.67	22.05170
60.0	1.68	21.25850
63.0	1.65	23.14050
57.0	1.69	19.95728
80.0	1.83	23.88844
76.0	1.71	25.99090
70.0	1.71	23.93899
70.0	1.65	25.71166
66.0	1.83	19.70796
52.0	1.64	19.33373
68.0	1.78	21.46194
82.5	1.76	26.63352

Tabela 1: Tabela de Amostras

que os alunos do curso de Engenharia de Sistemas(ES) estão, em média, mais "acima do peso" que o professor (de acordo com o BMI). Portanto, $H_0 : \mu = \mu_0$ e, em contrapartida, a hipótese alternativa é $H_1 : \mu < \mu_0$. Sendo μ é a média amostral do experimento, e μ_0 é o valor de referência, ou seja, BMI do professor.

Sempre que se seleciona uma amostra existe uma discrepância entre a média dessa amostra e a média da população, esse fato é conhecido como erro padrão da média ou erro amostral. Com base nisso, considerando o BMI do professor como a média populacional, é possível estudar o erro amostral utilizando o teste estatístico t com n-1 graus de liberdade.

O teste t foi escolhido pois não se sabe previamente qual o desvio padrão da população, o qual consideramos possuir distribuição normal. O termo "graus de liberdade"(df) se refere ao número de observações que são completamente livres para variar: Como estamos estudando uma única amostra: $df = n-1$.

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 $\alpha = 0.05$. Como consequência desse valor, teremos um nível de confiança $(1 - \alpha) = 0.95$, que representa a taxa de sucesso do método.

A potência 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.

4 Análise dos Resultados

Executando o algoritmo apresentado na seção 2 (Projeto Experimental), obtemos os seguintes valores para os parâmetros calculados:

Valor **de** mu: 26.3

Valor **de** alpha: 0.05

Numero **de** amostras: 13

Media amostral: 22.5230002789605

Desvio padrao: 2.56258202347936

Tamanho **de** efeito: -3.77699972103948

T critico: -2.17881282966723

Intervalo **de** confianca: 20.9744474604701 a 24.0715530974509

5 Conclusão

6 Bibliografia