Estudo de Caso 01: Comparação do IMC médio de alunos do PPGEE-UFMG ao longo de dois semestres

Homero Castro, Isaac Marinho e Stéfani Verissimo

17 de setembro, 2025

1. Introdução

O presente estudo de caso objetiva comparar o Índice de Massa Corporal (IMC), utilizado como um valor *proxy* para quantificar indicadores de estilo de vida, entre estudantes do Programa de Pós Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais (PPGEE-UFMG) entre os semestres de 2016/2 e 2017/1. O IMC é uma métrica altamente difundida e usada para indicar se ua determinada pessoa está na faixa de peso ideal. O seu cálculo é feito a partir da equação abaixo.

$$IMC = \frac{h}{w^2}, \text{ sendo } \begin{cases} h = \text{altura [cm]} \\ w = \text{peso [kg]} \end{cases}$$

De acordo com a World Health Organization (WHO), a faixa de IMC ideal para adultos situa-se entre 18.5 e 24.9 Kg/m², com valores abaixo de 18.5 Kg/m² indicando subnutrição, valores entre 24.9 e 29.9 Kg/m² indicando sobrepeso e valores acima de 30 Kg/m² indicando obesidade.

Desta forma, espera-se que a qualidade de vida dos alunos seja similar entre os dois semestres e que esta semelhança seja observada na equivalência estatística entre a média dos IMCs dos dois semestres.

Os códigos implementados para realização do presente estudo de caso estão disponíveis no repositório do github do grupo

2. Metodologia

2.1. Seleção e modificação dos dados

Os dados disponíveis foram carregados das bases imc_20162.csv e S01_20172.csv fornecidas, contendo informações acerca do gênero, peso em kilogramas e altura em metros dos estudantes dos semestres 2016.2 e 2017.1 respectivamente. A base do semestre 2016.2 contém dados tanto dos alunos de mestrado quanto dos alunos da graduação, portanto há uma coluna adicional diferenciando os mestrandos dos graduandos. Já na base referente ao semestre 2017.1, uma coluna adicional com a idade dos mestrandos também está presente.

Como o estudo de caso se propõe a analizar apenas os estudantes do PPGEE, os dados da base do semestre 2016.2 foram filtrados de acordo com a coluna que diferencia o curso dos alunos. Além disso, os dados obtidos das duas bases de dados foram separados de acordo com o sexo dos participantes e unidos em dois conjuntos de dados em dois conjuntos de dados, um para todos os participantes do sexo feminino e outro para todos os participantes do sexo masculino, diferenciando entre os semestres por meio da adição de uma nova coluna que indica se o individuo cursou a disciplina no semestre 2016.2 ou 2017.1.

```r
# Loading data

data_2016 <- read.csv('../data/imc_20162.csv')
data_2017 <- read.csv('../data/CS01_20172.csv', sep=';')
```

```r
# Selecting only PPGEE students
data_PPGEE_2016 <- data_2016[data_2016$Course == 'PPGEE',]
data_PPGEE_2017 <- data_2017

# Labeling the years
data_PPGEE_2016$Year <- 2016
data_PPGEE_2017$Year <- 2017

########################
# IMC calc. (Weight[kg]/(Height[m]^2))

data_PPGEE_2016$IMC = data_PPGEE_2016$Weight.kg/(data_PPGEE_2016$Height.m**2)
data_PPGEE_2017$IMC = data_PPGEE_2017$Weight.kg/(data_PPGEE_2017$height.m**2)

########################
# Women DF
data_PPGEE_2016_F <- data_PPGEE_2016[data_PPGEE_2016$Gender == 'F',]
data_PPGEE_2017_F <- data_PPGEE_2017[data_PPGEE_2017$Sex == 'F',]

data_PPGEE_full_F <- data.frame(Year = c(data_PPGEE_2016_F$Year, data_PPGEE_2017_F$Year),
                                IMC = c(data_PPGEE_2016_F$IMC, data_PPGEE_2017_F$IMC))

# Men DF
data_PPGEE_2016_M <- data_PPGEE_2016[data_PPGEE_2016$Gender == 'M',]
data_PPGEE_2017_M <- data_PPGEE_2017[data_PPGEE_2017$Sex == 'M',]

data_PPGEE_full_M <- data.frame(Year = c(data_PPGEE_2016_M$Year, data_PPGEE_2017_M$Year),
                                IMC = c(data_PPGEE_2016_M$IMC, data_PPGEE_2017_M$IMC))
```

2.2. Hipóteses levantadas

Para ambos os casos, sexo feminino e masculino, a hipótese nula representa a igualdade entre as médias dos IMCs nos dois períodos enquanto a hipótese alternativa defende que as médias são diferentes entre os semestres, ou seja, para o caso feminino, têm-se:

$$\begin{cases} H_0 : \mu_{F_{2016}} = \mu_{F_{2017}} \\ H_1 : \mu_{F_{2016}} \neq \mu_{F_{2017}} \end{cases}$$

já para o masculino, a representação matemática é:

$$\begin{cases} H_0 : \mu_{M_{2016}} = \mu_{M_{2017}} \\ H_1 : \mu_{M_{2016}} \neq \mu_{M_{2017}} \end{cases}$$

2.3. Testes realizados

Com o objetivo de estudar a hipótese nula e confirmar que não se pode rejeitá-la, é necessário realizar um teste estatístico pertinente. Para decidir qual teste estatístico utilizar, se faz necessário comfirmar uma série de suposições acerca dos dados utilizados, cada uma confirmada por um teste estatístico especifico.

Num primeiro momento, é necessário confirmar a normalidade da distribuição dos dados, que no presente estudo é averiguada por meio do teste de ***Shapiro-Wilk***. Num segundo momento, é necessário analisar se as varIâncias entre os grupos de dados são iguais que neste estudo ocorre por meio da aplicação do teste de ***Fligner-Killen***. Após isto, é interessante confirmar a independência entre os dados dos conjuntos utilizados para garantir que os resultados obtidos não possam ser explicados por meio de correlação interna dos dados, verificação esta que ocorre por meio da aplicação do teste de ***Durbin-Watson***.

Como foi observado que os dados utilizados são normalmente distribuidos, as variâncias são iguais e que os conjuntos apresentam grau satisfatório de independência, o teste estatístico aplicado para averiguar a hipótese nula foi o **teste T**. Por fim, para garantir que os resultados obtidos sejam relevantes, convém também averiguar o tamanho do efeito, avaliado pelo cálculo do poder do **teste t** realizado, considerando um delta de 1 desvio padrão.

3. Resultados

3.1 Avaliação da normalidade dos dados

```
# Resultado Shapiro feminino 2016
shapiro.test(data_PPGEE_full_F$IMC[data_PPGEE_full_F$Year == 2016])
```

```
##
##  Shapiro-Wilk normality test
##
## data:  data_PPGEE_full_F$IMC[data_PPGEE_full_F$Year == 2016]
## W = 0.91974, p-value = 0.4674
```

```
# Resultado Shapiro feminino 2017
shapiro.test(data_PPGEE_full_F$IMC[data_PPGEE_full_F$Year == 2017])
```

```
##
##  Shapiro-Wilk normality test
##
## data:  data_PPGEE_full_F$IMC[data_PPGEE_full_F$Year == 2017]
## W = 0.7475, p-value = 0.03659
```

```
# Resultado Shapiro masculino 2016
shapiro.test(data_PPGEE_full_M$IMC[data_PPGEE_full_M$Year == 2016])
```

```
##
##  Shapiro-Wilk normality test
##
## data:  data_PPGEE_full_M$IMC[data_PPGEE_full_M$Year == 2016]
## W = 0.92833, p-value = 0.1275
```

```
# Resultado Shapiro masculino 2017
shapiro.test(data_PPGEE_full_M$IMC[data_PPGEE_full_M$Year == 2017])
```

```
##
##  Shapiro-Wilk normality test
##
## data:  data_PPGEE_full_M$IMC[data_PPGEE_full_M$Year == 2017]
## W = 0.96494, p-value = 0.6206
```

Ao utilizar o Teste de Shapiro-Wilk para avaliar a suposição de normalidade dos dados de IMC em cada grupo, foi observado que para a amostra feminina de 2017, a hipótese de normalidade foi rejeitada ($W = 0.748$, $p = 0.037$), indicando que os dados não seguem uma distribuição normal, o que pode ser explicado pelo baixo espaço amostral (apenas 4 alunas).

Para os demais grupos (feminino 2016, masculino 2016 e masculino 2017), não foram encontradas evidências estatísticas para rejeitar a hipótese de normalidade ($p > 0.05$), sustentando a aplicação de testes paramétricos para estas amostras, de modo com que é possível concluir que é seguro seguir com a suposição da distribuição normal dos dados.

3.2 Avaliação da homocedasticidade dos dados

```
####
# Verifica-se se as variâncias são iguais.
####

fligner.test(IMC ~ Year, data = data_PPGEE_full_F)
```

```
##
##  Fligner-Killeen test of homogeneity of variances
##
## data:  IMC by Year
## Fligner-Killeen:med chi-squared = 0.71101, df = 1, p-value = 0.3991
```

```
fligner.test(IMC ~ Year, data = data_PPGEE_full_M)
```

```
##
##  Fligner-Killeen test of homogeneity of variances
##
## data:  IMC by Year
## Fligner-Killeen:med chi-squared = 0.082824, df = 1, p-value = 0.7735
```

A homogeneidade das variâncias do IMC entre os semestres de 2016 e 2017 foi avaliada pelo Teste de Fligner-Killeen. Para ambos os sexos, feminino ($^{2}(1) = 0.711$, p = 0.399) e masculino ($^{2}(1) = 0.083$, p = 0.774), não foi possível rejeitar a hipótese nula de igualdade de variâncias. Desta forma, a premissa de homocedasticidade foi satisfeita, permitindo o uso da versão padrão do teste t.

3.3 Avaliação da independência dos resíduos

```
########################
# Verificação de independência
dw_model_F <- lm(IMC ~ Year, data=data_PPGEE_full_F)
durbinWatsonTest(dw_model_F)
```

```
##  lag Autocorrelation D-W Statistic p-value
##    1    -0.001873537      1.831094   0.512
##  Alternative hypothesis: rho != 0
```

```
dw_model_M <- lm(IMC ~ Year, data=data_PPGEE_full_M)
durbinWatsonTest(dw_model_M)
```

```
##  lag Autocorrelation D-W Statistic p-value
##    1      0.03238024      1.935107   0.662
##  Alternative hypothesis: rho != 0
```

A independência das observações foi verificada por meio do Teste de Durbin-Watson aplicado aos resíduos de um modelo linear. Os resultados para o grupo feminino (p = 0.508) e masculino (p = 0.700) não indicaram a presença de autocorrelação, sustentando a suposição de que as medições de IMC são independentes entre si.

3.4 Teste T para a hipótese levantada

```
################## Resultado t-Test feminino ####################
res_ttest_f <- t.test(data_PPGEE_full_F$IMC ~ data_PPGEE_full_F$Year,
                      alternative = "two.sided",
                      mu = 0,
                      var.equal = TRUE,
```

```r
                      conf.level = 0.95)
print(res_ttest_f)
```

```
##
##  Two Sample t-test
##
## data:  data_PPGEE_full_F$IMC by data_PPGEE_full_F$Year
## t = 1.9308, df = 9, p-value = 0.08556
## alternative hypothesis: true difference in means between group 2016 and group 2017 is not equal to 0
## 95 percent confidence interval:
##  -0.4527037  5.7283762
## sample estimates:
## mean in group 2016 mean in group 2017
##          21.08443           18.44660
```

```r
################## Resultado t-Test masculino ####################
res_ttest_m <- t.test(data_PPGEE_full_M$IMC ~ data_PPGEE_full_M$Year,
                      alternative = "two.sided",
                      mu = 0,
                      var.equal = TRUE,
                      conf.level = 0.95)
print(res_ttest_m)
```

```
##
##  Two Sample t-test
##
## data:  data_PPGEE_full_M$IMC by data_PPGEE_full_M$Year
## t = 0.53979, df = 40, p-value = 0.5923
## alternative hypothesis: true difference in means between group 2016 and group 2017 is not equal to 0
## 95 percent confidence interval:
##  -1.784943  3.085836
## sample estimates:
## mean in group 2016 mean in group 2017
##          24.93595           24.28551
```

Para comparar o IMC médio entre os semestres, foi realizado um Teste t para amostras independentes. No grupo feminino, não foi encontrada diferença estatisticamente significativa entre a média de 2016 (M = 21.08) e 2017 (M = 18.45), com $t(9) = 1.93$, $p = 0.086$. Da mesma forma, para o grupo masculino, a diferença entre as médias de 2016 (M = 24.94) e 2017 (M = 24.29) também não foi significativa, $t(40) = 0.54$, $p = 0.592$. Portanto, o estudo não encontrou evidências de que o IMC médio dos alunos tenha se alterado entre os períodos analisados, corroborando com a hipótese nula de que os estilos de vida dos alunos do PPGEE nos semestres 2016.2 e 2017.1 são estatisticacmente equivalentes.

3.5 Tamanho do efeito e intervalo de confiânça

```r
########################
# Verificando o poder dos testes

# Calculando o desvio padrão agregado e o poder do teste feminino
agg_sd_F <- sd_pooled(data_PPGEE_2016_F$IMC, data_PPGEE_2017_F$IMC)

power_f <- power.t.test(n=nrow(data_PPGEE_full_F), delta=sd(data_PPGEE_full_F$IMC), sd=sd(data_PPGEE_fu
                        sig.level=0.05, type="two.sample", alternative = "two.sided")$power

print(paste("Poder teste feminino: ", round(power_f,4)))
```

```
## [1] "Poder teste feminino:  0.6071"
```

```r
# Calculando o desvio padrão agregado e o poder do teste masculino
agg_sd_M <- sd_pooled(data_PPGEE_2016_M$IMC, data_PPGEE_2017_M$IMC)

# Masculino
power_m <- power.t.test(n=nrow(data_PPGEE_full_M), delta=sd(data_PPGEE_full_M$IMC), sd=sd(data_PPGEE_fu
                        sig.level=0.05, type="two.sample", alternative = "two.sided")$power

print(paste("Poder teste masculino: ", round(power_m,4)))
```

```
## [1] "Poder teste masculino:  0.9949"
```

A análise de poder estatístico indicou que o teste para o grupo masculino apresentou um poder muito elevado (99.5%), conferindo alta confiança ao resultado não significativo. Em contrapartida, o teste para o grupo feminino apresentou um poder de 60.7%, valor considerado baixo. Isso sugere que o teste para as mulheres tinha um risco elevado de não detectar uma diferença real entre os grupos (erro tipo II), sendo esta uma limitação do estudo, provavelmente associada ao tamanho reduzido da amostra feminina, especialmente quanto ao semestre 2016.2.

Já a análise dos intervalos de confiança de 95% para a diferença entre as médias de IMC corrobora os resultados dos testes de hipótese. Para o grupo feminino, o intervalo variou de -0.45 a 5.73, enquanto para o masculino, a variação foi de -1.78 a 3.09.

Crucialmente, ambos os intervalos contêm o valor zero, indicando que uma diferença nula entre as médias dos semestres é um resultado plausível. A amplitude dos intervalos, especialmente para o grupo feminino, também sugere um grau considerável de incerteza na estimativa da verdadeira diferença, o que é consistente com o baixo poder estatístico e o tamanho reduzido da amostra.

4. Conclusões

A partir dos resultados obtidos, concluí-se que a hipótese nula não pode ser rejeitada com os testes realizados, principalmente no subconjunto dos indivíduos do sexo masculino, uma vez que apresentam um tamanho do efeito relevante e um intervalo de confiança significativamente mais estreito quando comparado ao subconjunto dos indivíduos do sexo feminino, que não apresentou um tamanho de efeito relevante.

Esta diferença entre os subconjuntos pode ser explicada por meio do tamanho do espaço amostral, que para o subconjunto feminino é significativamente menor, sendo até mesmo refutada a suposição da normalidade dos dados no subconjunto de alunas do semestre 2017.1, o que pode ter influênciado a confiabilidade do **teste t** realizado, uma vez que ele depende da suposição da normalidade dos dados e se beneficie de um maior espaço amostral.

Deste modo, futuras melhoras para o estudo de caso realizado podem partir de duas abordagens. A primeira consiste em adaptar o teste estatístico no subconjunto dos alunos do sexo feminino, de modo com que o teste utilizado não se baseie na suposição da normalidade dos dados, a exemplo do *teste t de welch*.

A segunda abordagem é mais direta e trabalhosa, consistindo em aumentar o espaço amostral ao entrevistar alunos de múltiplas turmas ao invés de apenas turmas da disciplina de projeto e análise de experimentos, o que resolveria o problema do espaço amostral reduzido.

Apendice A - Funções dos integrantes

- Homero: Responsável pelo projeto e implementação do estudo de caso e esboçar o relatório;
- Isaac: Responsável por redigir o texto completo completando e incrementando o esboço;
- Stéfani: Responsável por verificar o relatório completo e redigir a versão final.