Estudo de caso: Grupo D 3

Gilmar and Maressa Nunes R. Tavares and Victor

3 de Setembro, 2019

1 Summary

2 Planejamento do experimento

2.1 Objetivo do experimento

$$\begin{cases} H_0 : \mu = 50 \\ H_1 : \mu < 50 \end{cases}$$

$$\begin{cases} H_0 : \sigma^2 = 100 \\ H_1 : \sigma^2 < 100 \end{cases}$$

2.1.1 Descrição da coleta de dados

2.1.2 Descrição da coleta de dados

```r
imc <- function(height, weight){
  return(weight/(height^2))
}

data.2016.2 <- read.csv('imc_20162.csv', header = T, sep = ",")
data.2017.2 <- read.csv('CS01_20172.csv', header = T, sep = ";")
data.2017.2 <- data.2017.2 %>% rename(Gender = Sex)

# calculate the IMC
data.2016.2$imc <- imc(data.2016.2$Height.m, data.2016.2$Weight.kg)
data.2017.2$imc <- imc(data.2017.2$height.m, data.2017.2$Weight.kg)

# remove undergraduate students for 2016.2 data
data.2016.2 <- data.2016.2 %>% filter(Course == 'PPGEE')

# remove unecessary columns and rename
data.2016.2 <- data.2016.2 %>% select(Gender, imc)
data.2017.2 <- data.2017.2 %>% select(Gender, imc)

# split the 2016.2 data by gender and remove undergraduate students
data.2016.2.Females <- data.2016.2 %>% filter(Gender == 'F') %>% select(imc)
data.2016.2.Males <- data.2016.2 %>% filter(Gender == 'M') %>% select(imc)

# split the 2017.2 data by gender
data.2017.2.Females <- data.2017.2 %>% filter(Gender == 'F') %>% select(imc)
data.2017.2.Males <- data.2017.2 %>% filter(Gender == 'M') %>% select(imc)
```

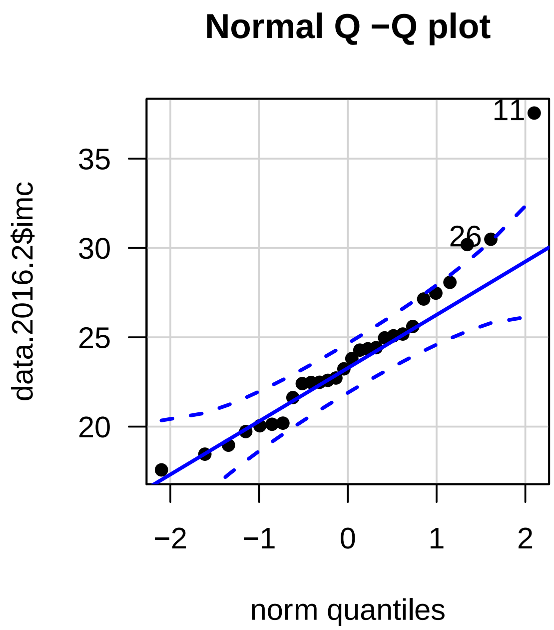

Figure 1: Gráfico quantil-quantil das populações

3 Análise Exploratória dos Dados

4 Resultados

4.1 Validação das premissas

Para realizar as inferências estatísticas sobre o IMC das duas populações é necessário validar as premissas antes de executar o teste. Neste caso, como as variâncias das duas populações é desconhecida, utiliza-se a distribuição t para o teste de hipóteses e para os intervalos de confiança [1]. A seguir são apresentados os testes realizados para validar as premissas exigidas pelo teste t.

1 - Normalidade:

Para avaliar a normalidade dos dados das duas populações utilizou-se o gráfico quantil-quantil e o teste de Shapiro Wilk com $\alpha = 0.05$

```
## [1] 11 26
```

```
shapiro.test(data.2016.2$imc)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  data.2016.2$imc
## W = 0.92185, p-value = 0.03854
```

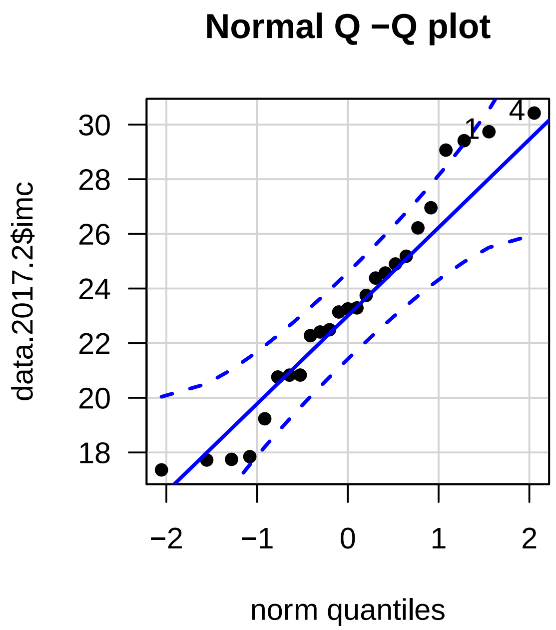

Figure 2: Gráfico quantil-quantil das populações

Pela análise do teste de Shapiro-Wilk e do gráfico concluímos que há evidências de que os dados de 2016.2 não são normalmente distribuídos, pois o pvalordo teste de Shapiro Wilk foi menor que 0.05.

```
## [1] 4 1
```

```
shapiro.test(data.2017.2$imc)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  data.2017.2$imc
## W = 0.95381, p-value = 0.3049
```

Por outro lado, não há evidências para rejeitar a normalidade dos dados de IMC coletados em 2017.2, pois p-valor foi maior que 0.05 (p-value = 0.3049); fato este que pode ser observado também no gráfico.

2- Igualdade de Variâncias:

A segunda premissa a ser avaliada é a igualdade de variâncias das duas populações, homocedasticidade. Para tanto, utilizou-se o teste F com a função var.test e $\alpha = 0.05$, que tem como hipótese nula a igualdade da variância das duas amostras.

```
var.test(data.2016.2$imc,data.2017.2$imc)
```

```
##
##  F test to compare two variances
```

```
##
## data:  data.2016.2$imc and data.2017.2$imc
## F = 1.2076, num df = 27, denom df = 24, p-value = 0.6443
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.5400368 2.6501200
## sample estimates:
## ratio of variances
##           1.207567
```

Pelo teste F conclui-se que não há evidências para rejeitar a hipótese nula de igualdade das variâncias da população duas populações, portanto, as duas amostras são homocedásticas.

3 - Independência

Considerando que as populações referem-se a dois grupos distintos da pós-graduação, sabe-se a priori, que as amostras são independentes. Para concluir em relação à independência realizou-se o teste Qui-quadrado com $\alpha = 0.05$.

```r
data <- c(data.2016.2$imc, data.2017.2$imc)
chisq.test(data)
```

```
##
##  Chi-squared test for given probabilities
##
## data:  data
## X-squared = 35.902, df = 52, p-value = 0.9565
```

Como era esperado, o teste Qui-quadrado reafirma a independência entre as duas amostras, como o pvalor $= 0.9565$.

4.2 Teste do custo médio

4.2.1 Teste da variância do custo

5 Análise Estatística

5.1 Teste sobre a média do custo

5.1.1 Cálculo do tamanho amostral

5.1.2 Teste de Hipoteses

5.1.3 Calculo do intervalo de confianca

5.1.4 Validaçao das premissas

5.2 Teste sobre a variância do custo

5.2.1 Teste de Hipoteses

5.2.2 Calculo do intervalo de confianca

5.2.3 Validaçao das premissas

6 Discussão e Conclusões

7 Divisão das Atividades

Victor - Reporter Maressa - Coordenadora Gilmar - Verificador e Monitor

Referências

[1] D. C. Montgomery and G. C. Runger, *Applied statistics and probability for engineers, (with cd)*. John Wiley & Sons, 2007.