UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA

Alunos
Gleyce Alves Pereira da Silva
Ivanildo Batista da Silva Júnior
Jaine de Moura Carvalho
Taciana Araújo da Silva

Professor
Dr. Lucian Bogdan Bejan

Resolução da nona lista de Estatística Aplicada

Recife-PE, 17 de junho de 2021

Sumário

1 Questão 1

Uma empresa deseja estudar o efeito de uma pausa de dez minutos para um cafezinho sobre a produtividade de seus trabalhadores. Para isso, sorteou seis operários, e contou o número de peças produzidas durante uma semana sem intervalo e uma semana com intervalo. Os resultados sugerem se há ou não melhora na produtividade? Caso haja melhora, qual deve ser o acréscimo médio de produção para todos os trabalhadores da fábrica ?

Operário	1	2	3	4	5	6
Sem Intervalo	23	35	29	33	43	32
Com Intervalo	28	38	29	37	42	30

1.1 Resolução da questão 1

Para resolver essa questão é necessário usar o teste para saber as médias são ou não estatísticamente iguais. A hipóteses nula e alternativa são dadas por:

$$\begin{cases} \mu_{sem} = \mu_{com} \\ \mu_{sem} \neq \mu_{com} \end{cases}$$

Vamos criar dois vetores com os resultados dos operário e usar o teste de igualdade de média definindo o nível de significância em 95% (α = 0.05).

```
t.test(sem_intervalo, com_intervalo, alternative = "less",
       paired = TRUE, conf.level = 0.95)
```

```
##
##  Paired t-test
##
## data:  sem_intervalo and com_intervalo
## t = -1.2753, df = 5, p-value = 0.1291
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##       -Inf 0.870003
## sample estimates:
## mean of the differences
##                    -1.5
```

O resultado acima mostra um *p-valor* igual a 0.1291, acima do α de 0.05, logo não podemos rejeitar a hipótese nula de que as duas médias são iguais, portanto concluímos que não houve melhoria com a implementação do intervalo para os operários.

2 Questão 2

Num levantamento feito com os operários da indústria mecânica, chegou-se aos seguintes números: salário médio = 3.64 salários mínimos e desvio padrão = 0.85 salário mínimo. Suspeita-se que os salários da subclasse formada pelos torneiros mecânicos são diferentes dos salários do conjunto todo, tanto na média como na variância. Que conclusões você obteria se uma amostra de 25 torneiros apresentasse salário médio igual a 4.22 salários mínimos e desvio padrão igual a 1.25 salário mínimo ?

2.1 Resolução da questão 2

Primeiro vamos testar se as médias são iguais: A suspeita é que os salários médio são diferentes, por isso vamos gerar um intervalo de confiança.

```
# Intervalo de confianca para a media
desvio <-0.85
nobs <-25
x_media <-4.22

#z-score
z<-qnorm(1-(0.05/2))
me<-z*desvio/sqrt(nobs)
x_media+c(-1,1)*me
```

```
## [1] 3.886806 4.553194
```

O valor de 3.64 não está dentro do intervalo de confiança gerado, por isso rejeitamos que o salário médio é diferente de 3.64.

Agora é preciso realizar um teste de igualdade de variância

```
# Intervalo de confianca para a varianca
S<-1.25
S2<-S^2
nobs<-25
sup<-qchisq(1-(0.05/2),nobs-1)
inf<-qchisq(0.05/2,nobs-1)
lim_inf<-(nobs-1)*S2/sup
lim_sup<-(nobs-1)*S2/inf

#intervalo de confiança
c(lim_inf,lim_sup)
```

```
## [1] 0.9526452 3.0239131
```

O valor da variância que é $\sigma^2 = 0.85^2 = 0.7225$, não encontra-se dentro do intervalo de confiança, logo rejeitamos que a variância é diferente de 0.7225.

3 Questão 3

Uma amostra de 100 trabalhadores de uma fábrica grande demora, em média, 12 minutos para completar uma tarefa, com um desvio padrão de dois minutos. Uma amostra de 50 trabalhadores de uma outra fábrica demora, em média, 11 minutos para completar a mesma tarefa, com desvio padrão igual a três minutos.

(a) Construa um IC de 95% para a diferença entre as duas médias populacionais.

(b) Deixe bem claro quais as suposições feitas para a solução apresentada.

3.1 Resolução da questão 3

Definindo os dados.

```
nobs_x = 100
x_media = 12
desvio_x = 2

nobs_y = 50
y_media = 11
desvio_y = 3
```

Passo 1: Definindo as hipóteses

$$\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

Passo 2: Estatística do teste $W = \frac{S_2^2}{S_1^2}$. Sob H_0, $W \sim F(49, 99)$. Definindo os valores críticos.

```
c(qf(p=0.025, df1=49, df2=99, lower.tail=T),
  qf(p=0.025, df1=49, df2=99, lower.tail=F))
```

```
## [1] 0.6010322 1.5973853
```

No caso a região crítica engloba de 0 a 0.6010322 e de 1.5973853 até $+\infty$.

Passo 3: Valor observado é dado por $w_0 = s_1^2/s_2^2$

```
w = desvio_y^2/desvio_x^2
w
```

```
## [1] 2.25
```

Vemos que a estatística w_0 encontra-se dentro da região crítica, logo concluímos que podemos rejeitamos a hipótese nula de igualdade das variâncias.

3.1.1 letra a)

Definindo os valores de s_1^2 e de s_2^2:

```
A = desvio_x^2/nobs_x
B = desvio_y^2/nobs_y
```

Graus de liberdade para distribuição *t-student*.

```
v=round((A+B)^2/(((A^2)/(nobs_x-1))+((B^2)/(nobs_y-1))),0)
v
```

```
## [1] 71
```

Gerando o intervalo de confiança a 95%.

```
round((x_media-y_media)+c(-1,1)*qt(0.975,v)*sqrt(A+B),3)
```

```
## [1] 0.065 1.935
```

3.1.2 letra b)

O tempo necessário para a realização das tarefa é uma variável com uma distribuição normal e as variâncias populacionais das fábricas são diferentes.

4 Questão 4

A Torrefação Guarany está querendo comprar uma nova ensacadora de café. Após consultar o mercado, ficou indecisa entre comprar a de marca A ou a de marca B. Quanto ao custo, facilidade de pagamento, tamanho etc. elas são equivalentes. O fator que decidirá a compra será a precisão em encher os pacotes (medido pela variância). Deseja-se, na realidade, testar hipótese $\sigma_A^2 = \sigma_B^2$, através da estatística $F = S_A^2/S_B^2$. Podem-se construir regiões críticas bilaterais, unilaterais à direita ou à esquerda, dependendo do objetivo. Indique qual seria a região crítica mais favorável às seguintes pessoas: (Justifique.)

(a) proprietário da torrefação;

(b) fabricante de A;

(c) fabricante de B.

4.1 Resolução da questão 4

4.1.1 letra a)

a) Para favorecer o proprietário da Torrefação, o deve-se realizar um teste de bilateral, os valores críticos são

$$F_{\alpha/2}(n-1, m-1)$$

e

$$\frac{1}{F_{\alpha/2}(n-1, m-1)}$$

Se a estatística de teste for menor que o primeiro valor ou maior do que o segundo valor, você deve rejeitar a hipótese nula. Se a estatística de teste estiver entre o primeiro e segundo valores, você não deve rejeitar a hipótese nula. Nesse caso, o dono tem apenas o interesse no melhor caso (menor variância), sem desejar favorecer uma ou outra marca. Um nível de significância de 0,05, por exemplo, indica que o risco de se concluir que existe uma diferença, quando, na verdade, não existe nenhuma diferença real, é de 5%.

4.1.2 letra b)

b) Para favorecer o fabricante A deve-se aplicar um teste unilateral à esquerda, com uma hipótese alternativa de "menor que", o valor crítico é

$$\frac{1}{F_{\alpha/2}(n-1, m-1)}$$

Se a estatística de teste for inferior ao valor crítico, você rejeita a hipótese nula. Caso contrário, você não deve rejeitar a hipótese nula. Pois a chance de rejeitar a Hipótese nula é pequena.

4.1.3 letra c)

c) Para favorecer o fabricante B deve-se aplicar um teste unilateral à direita, com hipótese de "maior que", o valor crítico é $F_{\alpha/2}(n-1, m-1)$. Se a estatística de teste for maior que o valor crítico, você rejeita a hipótese nula. Caso contrário, você não deve rejeitar a hipótese nula. Pois a chance de rejeitar a Hipótese nula é pequena.

5 Questão 5

Na região sul da cidade, 60 entre 400 pessoas preferem a bebida Meca-Mela entre as demais similares. Na região norte, a proporção é de 40 entre 225 entrevistados. Baseado no resultado dessa amostra, você diria que a proporção de todos os moradores nas duas regiões é a mesma? Use $\alpha =$ 0.05.

5.1 Resolução da questão 5

Gerando o teste para igualdade de proporção entre as duas amostras:

```
prop.test(x = c(60, 40), n = c(400,225))
```

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  c(60, 40) out of c(400, 225)
## X-squared = 0.63296, df = 1, p-value = 0.4263
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.09224259  0.03668703
## sample estimates:
##    prop 1    prop 2
## 0.1500000 0.1777778
```

O valor p do teste é 0.4263, que é maior que o nível de significância $\alpha = 0.05$. Podemos concluir que a proporção de fumantes é significativamente igual nos dois grupos com um valor de p = 0.4263. Então, concluímos que, a proporção de moradores nas duas regiões é a mesma.