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 oitava lista de Estatística Aplicada

Recife-PE, 17 de junho de 2021

Sumário

1 Questão 1

Os novos operários de uma empresa são treinados a operarem uma máquina, cujo tempo X (em horas) de aprendizado é anotado. Observou-se que X segue de perto adistribuição N(25, 100). Uma nova técnica de ensino, que deve melhorar o tempo de aprendizado, foi testada em 16 novos empregados, o quais apresentaram 20.5 horas como tempo médio de aprendizado. Usando o valor-p, você diria que a nova técnica é melhor que a anterior?

1.1 Resolução da questão 1

Passo 1: Definir as hipóteses nula e alternativa:

- H_0 : A média é igual a 25 (μ = 25)
- H_1 : A média é menor que 25 (μ < 25)

Passo 2: Amostras terão média X $\sim N(25, 100)$.

Passo 3: Sob a hipótese de que H_0 é verdadeira, e pelo fato da variância ser conhecida (σ =100), podemos calcular o valor de α para o valor de z

```
#média da distribuição
media_dist = 25

#média da amostra
media_amostral = 20.5

#desvio padrão
desvio = sqrt(100/16)

#Valor z
z = (media_amostral-media_dist)/desvio

#valor de alpha
round(pnorm(z),4)
```

```
## [1] 0.0359
```

Passo 4: Se fixarmos o nível de significância em 5%, então rejeitamos H_0 (média igual a 25) e aceitamos a hipótese alternativa de que a média é menor que 25. Concluímos que sim, a nova técnica é melhor que a anterior.

2 Questão 2

O tempo médio, por operário, para executar uma tarefa, tem sido 100 minutos, com um desvio padrão de 15 minutos. Introduziu-se uma modificação para diminuir esse tempo, e, após certo período, sorteou-se uma amostra de 16 operários, medindo-se o tempo de execução de cada um. O tempo médio da amostra foi 85 minutos, e o desvio padrão foi 12 minutos. Estes resultados trazem evidências estatísticas da melhora desejada? Em caso afirmativo, estime o novo tempo médio de execução. (Apresente as suposições teóricas usadas para resolver o problema.)

2.1 Resolução da questão 2

Passo 1: Definir as hipóteses nula e alternativa:

- H_0 : A média é igual a 100 ($\mu = 100$)
- H_1 : A média é menor que 100 ($\mu < 100$)

Passo 2: Sob H_0, $\dfrac{\overline{X} - 100}{S/4} \sim t(15)$
Passo 3: Definindo $\alpha = 0.05$, vamos encontrar o valor crítico do valor *t*.

```
qt(p=0.05, df=15, lower.tail=T)
```

```
## [1] -1.75305
```

Passo 4: Calcular o valor de t_0

```
#Calculando o valor de t0

#defininfo as variáveis
media_antes = 100
media_depois = 85
dp = 12

#criando uma função para calcular t0
t0 = function(media_antes, media_depois, dp){
  (media_depois-media_antes)/(dp/4)
}

#Calculando t0

t0(media_antes, media_depois, dp)
```

```
## [1] -5
```

Temos que o valor crítico (limite da região crítica de aceitação da hipótese nula) é -1.75305 e o valor de t_0 é -5.

Passo 5: Como t_0 pertence à região crítica, rejeita-se a hipótese nula (H_0) de que a média é igual a 100 . Logo, há evidências de que houve melhorias no tempo de execução .

3 Questão 3

O número médio diário de clientes de um posto de gasolina tem sido 250, com um desvio padrão de 80 clientes. Durante uma campanha de 25 dias, em que os clientes recebiam um brinde, o número médio de clientes foi 280, com um desvio padrão de 50. Você diria que a campanha modificou a distribuição do número de clientes do posto? Descreva as suposições feitas para a resolução do problema.

3.1 Resolução da questão 3

Para sabermos se houve uma mudança na distribuição do número de clientes precisamos saber se as médias e as variâncias são diferentes, então é preciso fazer o teste para a média e o teste para variância.

Teste para variância

Passo 1: Definir as hipóteses nula e alternativa.

$$\begin{cases} H_0 : \sigma^2 = 80 \\ H_1 : \sigma^2 \neq 80 \end{cases}$$

Passo 2 : Definir a estatística do teste

$$(n-1)\frac{S^2}{\sigma^2} \sim \chi^2(24)$$

Passo 3: Com α = 0.05, calculamos os valores críticos. A região crítica vai de 0 a 12.40115 e de 39.36408 até $+\infty$. Ou seja, aceitamos a hipótese nula se a estatística que calcularemos for menor que 12.40115 ou maior que 39.36408.

```
c(qchisq(p=.025, df=24, lower.tail=T),
  qchisq(p=.025, df=24, lower.tail=F))
```

```
## [1] 12.40115 39.36408
```

Passo 4: Calculando o valor observado

Abaixo criamos uma função para calcular a estatística e inserimos os valores da questão para obtermos o valor

```
stat_cs = function(n,S,sigma){
  (n-1)*(S^2/sigma^2)
}
```

```
stat_cs(25,50,80)
```

```
## [1] 9.375
```

Passo 5: Por fim vamos analisar os resultado. O valor da estatística é de 9.375 e essa valor enocntra-se dentro da região crítica, logo concluímos que há **evidência de mudança na variância**.

Teste para média

Passo 1: Definir as hipóteses nula e alternativa.

$$\begin{cases} H_0 : \mu = 250 \\ H_1 : \mu > 250 \end{cases}$$

Passo 2 : Definir a estatística do teste, em que sob H_0 temos que

$$\frac{\overline{X} - 250}{S/\sqrt{(25)}} \sim t(24)$$

Passo 3: Com α = 0.05 e sendo o teste unilateral, calculamos o valor crítico abaixo.

```
qt(p=0.05, df=24,lower.tail=F)
```

```
## [1] 1.710882
```

Passo 4: Calculando t_0, temos o valor de 3.

```
to = function(M1,M2,sigma,n){
  (M2-M1)/(sigma/sqrt(n))
}
```

```
to(250,280,50,25)
```

```
## [1] 3
```

Passo 5: Como o valor de t_0 (3) está fora da região de aceitação da hipótese nula, rejeitamos a hipótese de igualdade das média. Então, concluímos que houve uma mudança na média.

Conclusão Final: Como os testes apresentaram que há evidência de que as variâncias e as médias são diferentes, então concluímos que houve uma mudança na distribuição do número de clientes.

4 Questão 4

A receita média, em porcentagem, dos quase 600 municípios de um estado tem sido 7%. O governo pretende melhorar esse índice e, para isso, está estudando alguns incentivos. Para verificar os efeitos desses incentivos, sorteou 10 cidades e estudou quais seriam as porcentagens investidas neles. Os resultados foram, em porcentagem, 8, 10, 9, 11, 8, 12, 16, 9, 12, 13. Admitindo-se que esses números realmente venham a ocorrer, os dados trazem evidência de melhoria? Caso altere a média do estado, dê um intervalo de confiança para a nova média.

4.1 Resolução da questão 4

Definindo as estatísticas:
dados e número de observações:

```
dados <- c(8, 10, 9, 11, 8, 12, 16, 9, 12, 13)
n <- length(dados)
print(n)
```

```
## [1] 10
```

Desvio-padrão e média:

```
sd <- sd(dados)
print(sd)
```

```
## [1] 2.529822
```

```
media <- mean(dados)
print(media)
```

```
## [1] 10.8
```

Valor do α que será usado :

```
alpha <- 0.05
print(alpha)
```

```
## [1] 0.05
```

Queremos testar:

$$\begin{cases} H_0 : \mu = 7 \\ H_1 : \mu > 7 \end{cases}$$

Como σ é desconhecido, então $\overline{X} \sim t$. Para calcular o valor crítico t_c com α = 0.05, utilizamos a função abaixo.

```
t_c=qt(p=1-0.05,df=n-1)
round(t_c,3)
```

```
## [1] 1.833
```

Logo, RC = {t: t > 1.833}. Calculando t_0, temos:

```
t_o=(media-7)*sqrt(n)/sd
round(t_o,3)
```

```
## [1] 4.75
```

Como t_0 pertence à região crítica, rejeita-se H_0. Logo, há evidências de melhorias.

Calculando **intervalo de confiança** para a nova média:

```
lim_inf=media + qt(alpha/2, df=n-1)*(sd/sqrt(n))
lim_sup=media + qt(1-alpha/2, df=n-1)*(sd/sqrt(n))
print(paste("IC(mu;", 1-alpha,")=[", round(lim_inf,2), ";", round(lim_sup,2),"]"))
```

```
## [1] "IC(mu; 0.95 )=[ 8.99 ; 12.61 ]"
```

5 Questão 5

Um escritório de investimento acredita que o rendimento das diversas ações movimentadas por ele foi de 24%. Mais ainda, a nova estratégia definida deve garantir uma maior uniformidade nos rendimentos das diversas ações. No passado, o desvio padrão do rendimento era da ordem de 5%. Para verificar as duas hipóteses, tomaram-se 8 empresas ao acaso, obtendo-se os seguintes rendimentos (dados em %): 23.6, 22.8, 25.7, 24.8, 26.4, 24.3, 23.9 e 25. Quais seriam as conclusões?

5.1 Resolução da questão 5

Para sabermos se houve uma mudança na distribuição do número de clientes precisamos saber se as médias e as variâncias são diferentes, então é preciso fazer o teste para a média e o teste para variância.

Teste para variância

Passo 1: Definir as hipóteses nula e alternativa.

$$\begin{cases} H_0 : \sigma^2 = 25 \\ H_1 : \sigma^2 < 25 \end{cases}$$

Passo 2 : Definir a estatística do teste

$$(n-1)\frac{S^2}{\sigma^2} \sim \chi^2(7)$$

Passo 3: Com α = 0.05, calcularemos o valor crítico. A região crítica vai até o valor de 2.167, conforme cálculo realizado abaixo. Ou seja, aceitamos a hipótese nula se a estatística que calcularemos for menor que 2.167.

```
round(qchisq(p=.05, df=7, lower.tail=T),3)
```

```
## [1] 2.167
```

Passo 4: Calculando o valor observado. Porém, antes iremos calcular a variância S^2 para ser usada no cálculo do valor observado.

```
var(c(23.6,22.8,25.7, 24.8, 26.4, 24.3, 23.9, 25))
```

```
## [1] 1.35125
```

Obtido o valor da variância S^2, vamos criar uma função para calcular a estatística e inserimos os valores da questão para obtermos o valor e usaremos para realizar o cálculo.

```
chi_obs = function(obs,S2,sigma2){
  (sigma2/S2^2)*(obs-1)
}
```

```
chi_obs(8,5,1.35125)
```

```
## [1] 0.37835
```

Passo 5: Como o valor observado está dentro da região crítica, há indícios da diminuição da variância

Teste para média

Passo 1: Definir as hipóteses nula e alternativa.

$$\begin{cases} H_0 : \mu = 24 \\ H_1 : \mu \neq 24 \end{cases}$$

Passo 2 : Definir a estatística do teste, em que sob H_0 temos que

$$\frac{\overline{X} - 24}{S/\sqrt{(8)}} \sim t(7)$$

Passo 3: Com α = 0.05 e sendo o teste bicaudal, calculamos o valor crítico abaixo.

```
qt(p=0.025, df=7,lower.tail=F)
```

```
## [1] 2.364624
```

Além disso vamos calcular a média dos valores coletados das 8 empresas.

```
mean(c(23.6,22.8,25.7, 24.8, 26.4, 24.3, 23.9, 25))
```

```
## [1] 24.5625
```

e o desvio padrão

```
sd(c(23.6,22.8,25.7, 24.8, 26.4, 24.3, 23.9, 25))
```

```
## [1] 1.162433
```

Passo 4:Usando a função criada da questão 3, calculamos o valor de t_0:

```
to(24,24.5625,1.162433,8)
```

```
## [1] 1.368673
```

Passo 5: Como o valor de t_0 é menor do que o valor crítico, então rejeitamos a hipótese nula de que as médias são iguais e aceitamos que há evidência de que os rendimentos médio são diferentes de 24%