Resolução de Exercícios

Profª Julienne Borges

PUC Minas Virtual

Introdução aos testes de hipóteses

Tipos de Hipóteses Estatísticas

<u>Para uma média:</u>

$H_0: \mu = \mu_0$ vs $H_a: \mu \neq \mu_0$ → Hipótese **bilateral ou bicaudal**

$H_0: \mu \geq \mu_0$ vs $H_a: \mu < \mu_0$ → Hipótese **unilateral à esquerda**

$H_0: \mu \leq \mu_0$ vs $H_a: \mu > \mu_0$ → Hipótese **unilateral à direita**

<u>Para uma proporção:</u>

$H_0: p = p_0$ vs $H_a: p \neq p_0$ → Hipótese **bilateral ou bicaudal**

$H_0: p \geq p_0$ vs $H_a: p < p_0$ → Hipótese **unilateral à esquerda**

$H_0: p \leq p_0$ vs $H_a: p > p_0$ → Hipótese **unilateral à direita**

Nível de significância (α)

Determina o erro máximo tolerado no teste de hipóteses!

O **nível de significância (α)** de um teste é a probabilidade de uma hipótese nula ser rejeitada, quando verdadeira.

Nível de confiança	x	Nível de significância
(1- α)		(α)
95%		5%
98%		2%

Estatística de teste

A **estatística de teste** é uma estatística amostral, ou um valor baseado nos dados amostrais. Utiliza-se uma estatística de teste para tomar uma decisão sobre a rejeição ou não da hipótese nula.

Valor p

<u>O valor p **quantifica o erro cometido ao rejeitar a hipótese nula**</u>. Um valor p muito pequeno sugere que os resultados amostrais são muito improváveis sob a hipótese nula, ou seja, constitui evidência contra a hipótese nula.

O critério de decisão baseado no valor p é feito da seguinte maneira:

- **Rejeitar a hipótese nula (H_0)** se o valor p é no máximo igual ao nível de significância (α).
- **Não rejeitar a hipótese nula (H_0)** se o valor p é maior do que o nível de significância (α).

Utilizando o Microsoft Excel

Teste de hipótese para média populacional

σ é conhecido

n =	**25**	digite o tamanho da amostra
$\bar{x}$=	**15,50**	digite a média amostral
σ=	**3,78**	digite o desvio-padrão populacional (sigma)
α=	**0,05**	digite o nível de significância do teste
M_0 =	**15**	digite o valor da média hipótetica
sinal em HA	**<>**	digite o sinal na hipótese H1 (<, > ou < >)
H0 :	M = 15	hipótese nula
HA :	M < > 15	hipótese alternativa
tipo do teste	bicaudal	tipo do teste (unicaudal ou bicaudal)
Z =	0,66138	estatística de teste
valor-p	0,50837	valor-p do teste

Observações importantes: **Utilize <> quando a o sinal de HA for de ≠**

Bicaudal é sinônimo de bilateral

Teste_Z_Média | Teste_t_Média | Teste_Z_Proporção

Teste de hipóteses para uma média utilizando o R

Teste t para uma média utilizando o R

```
x<-c(x1,x2,...,xn)
t.test(x, alternative = c("two.sided", "less", "greater"), mu = m0, conf.level = 0.95)
```

Onde:

x → representa o vetor de valores observados na amostra.

alternative = c("two.sided", "less", "greater") → representa o tipo de hipótese que será testado:

- "two.sided“ → bilateral
- “less” → Unilateral à esquerda
- “greater” → Unilateral à direita

m0 → você deve indica a média que está sendo testada por meio das hipóteses

conf.level → você deve especificar o nível de confiança para o teste de hipóteses

Teste de hipóteses para uma proporção utilizando o R

```
prop.test(x, n, p = NULL, alternative = c("two.sided", "less",
"greater"), conf.level = 0.95, correct = FALSE)
```

Onde:

x → número de itens que apresentam a característica de interesse.

n → tamanho da amostra

NULL → você deve indica a proporção que está sendo testada por meio das hipóteses

alternative = c("two.sided", "less", "greater") → representa o tipo de hipótese que será testado:

"two.sided“ → bilateral

“less” → Unilateral à esquerda

“greater” → Unilateral à direita

conf.level → você deve especificar o nível de confiança para o teste de hipóteses

correct → FALSE não utiliza a correção de continuidade de Yates ou TRUE utiliza a correção de continuidade de Yates

IMPORTANTE: O R utiliza a distribuição qui-quadrado e não a distribuição Normal para a realização do teste!

Teste de hipóteses para uma proporção utilizando o R

```
binom.test(x, n, p = NULL, alternative = c("two.sided", "less",
"greater"), conf.level = 0.95)
```

Onde:

x → número de itens que apresentam a característica de interesse.

n → tamanho da amostra

NULL → você deve indica a proporção que está sendo testada por meio das hipóteses

alternative = c("two.sided", "less", "greater") → representa o tipo de hipótese que será testado:

"two.sided“ → bilateral

“less” → Unilateral à esquerda

“greater” → Unilateral à direita

conf.level → você deve especificar o nível de confiança para o teste de hipóteses

A função *binom.test* utiliza a distribuição exata, a distribuição binomial, para construção do intervalo de confiança para uma proporção.

https://www.est.ufmg.br/~monitoria/Material/ApostilaR/InferenciaProporcoes.html

Exercício 1

Um pesquisador afirma que uma determinada dieta pode proporcionar uma perda média de peso de 12 kg no período de acompanhamento. Ele quer verificar se uma nova dieta é mais eficaz, ou seja, proporciona maior emagrecimento. Para isso, ele acompanhou um grupo de 17 pacientes que seguiu a nova dieta e, decorrido certo tempo, observou a perda de peso de cada paciente do grupo. Os dados estão apresentados a seguir: 12, 8, 15, 13, 10, 12, 14, 11, 12, 13, 15, 19, 15, 12, 13, 16, 15.

Utilizando um nível de 5% de significância, o que você pode concluir?

Exercício 1

Um pesquisador afirma que uma determinada dieta pode proporcionar uma perda média de peso de 12 kg no período de acompanhamento. Ele quer verificar se uma nova dieta é mais eficaz, ou seja, proporciona maior emagrecimento. Para isso, ele acompanhou um grupo de 17 pacientes que seguiu a nova dieta e, decorrido certo tempo, observou a perda de peso de cada paciente do grupo. Os dados estão apresentados a seguir: 12, 8, 15, 13, 10, 12, 14, 11, 12, 13, 15, 19, 15, 12, 13, 16, 15.

Utilizando um nível de 5% de significância, o que você pode concluir?

Peso médio perdido com a nova dieta (em kg)

↓

σ NÃO é conhecido

↓

Distribuição t-Student

↓

Dados brutos

↓

$H_0: \mu<12$

$H_a: \mu>12$

Exercício 1

Utilizando a função t.test do R, temos:

> peso<-c(12, 8, 15, 13, 10, 12, 14, 11, 12, 13, 15, 19, 15, 12, 13, 16, 15)
> t.test(peso, alternative=c("greater"), mu=12, conf.level=0.95)

```
One Sample t-test

data:  peso
t = 2.0068, df = 16, p-value =
0.03099
alternative hypothesis: true mean
is greater than 12
95 percent confidence interval:
 12.16063      Inf
sample estimates:
mean of x
 13.23529
```

Como interpretar o valor p?

SE rejeitarmos H_0, cometemos um erro de 0,03099 ou 3,099%.

Exercício 1

Utilizando a função t.test do R, temos:

> peso<-c(12, 8, 15, 13, 10, 12, 14, 11, 12, 13, 15, 19, 15, 12, 13, 16, 15)
> t.test(peso, alternative=c("greater"), mu=12, conf.level=0.95)

Como concluir o teste?
Podemos dizer, com um nível de 5% de significância, que a nova dieta é mais eficaz, ou seja, proporciona maior emagrecimento.

```
One Sample t-test

data:  peso
t = 2.0068, df = 16, p-value =
0.03099
alternative hypothesis: true mean
is greater than 12
95 percent confidence interval:
 12.16063       Inf
sample estimates:
mean of x
 13.23529
```

Exercício 2

Uma companhia de serviços de ônibus intermunicipais planejou uma nova rota para servir vários locais situados entre duas cidades importantes. Um estudo preliminar afirma que a duração das viagens pode ser considerada uma variável aleatória com distribuição normal, com média igual a 300 minutos e desvio padrão 30 minutos. Uma amostra de dez viagens realizadas nessa nova rota apresentou média igual a 314 minutos. Esse resultado comprova ou não o tempo médio determinado nos estudos preliminares? Considere α=0,1.

Duração média das viagens na nova rota (em minutos).

↓

σ é conhecido

↓

Distribuição Normal

↓

Dados resumidos

n=10; $\bar{x}$=314 ; σ=30

↓

H_0: $\mu = 300$

H_a: $\mu \neq 300$

Exercício 2

Duração média das viagens na nova rota (em minutos).

↓

σ é conhecido

↓

Distribuição Normal

↓

Dados resumidos

n=10; $\bar{x}$=314 ; σ=30; α=0,1

↓

H_0: μ = 300

H_a: μ ≠ 300

	A	B	C	D	E	F
1		**Teste de hipótese para média populacional**				
2			***σ é conhecido***			
3						
4		**n =**	**10**	digite o tamanho da amostra		
5		$\bar{x}$=	**314,00**	digite a média amostral		
6		σ=	**30,00**	digite o desvio-padrão populacional (sigma)		
7		α=	**0,1**	digite o nível de significância do teste		
8		**M_0 =**	**300**	digite o valor da média hipótetica		
9		**sinal em HA**	**<>**	digite o sinal na hipótese H1 (<, > ou < >)		
10		**H0 :**	M = 300	hipótese nula		
11		**HA :**	M < > 300	hipótese alternativa		
12		**tipo do teste**	bicaudal	tipo do teste (unicaudal ou bicaudal)		
13						
14		**Z =**	1,47573	estatística de teste		
15		**valor-p**	0,14002	valor-p do teste		
16						
17		**Observações importantes:**		**Utilize <> quando a o sinal de HA for de ≠**		
18				**Bicaudal é sinônimo de bilateral**		
19						

Teste_Z_Média | Teste_t_Média | Teste_Z_Proporção

Exercício 2

Como interpretar o valor p?
SE rejeitarmos H_0, cometemos um erro de 0,14 ou 14%.

Como concluir o teste?
Podemos dizer, com um nível de 10% de significância, que a duração média das viagens na nova rota é igual a 300 minutos, ou seja, comprova os resultados nos estudos preliminares.

	A	B	C	D	E	F
1		**Teste de hipótese para média populacional**				
2			***σ é conhecido***			
3						
4		**n =**	**10**	digite o tamanho da amostra		
5		$\bar{x}$=	**314,00**	digite a média amostral		
6		σ=	**30,00**	digite o desvio-padrão populacional (sigma)		
7		α=	**0,1**	digite o nível de significância do teste		
8		**M_0 =**	**300**	digite o valor da média hipótetica		
9		**sinal em HA**	**<>**	digite o sinal na hipótese H1 (<, > ou < >)		
10		**H0 :**	M = 300	hipótese nula		
11		**HA :**	M < > 300	hipótese alternativa		
12		**tipo do teste**	bicaudal	tipo do teste (unicaudal ou bicaudal)		
13						
14		**Z =**	1,47573	estatística de teste		
15		**valor-p**	0,14002	valor-p do teste		
16						
17		**Observações importantes:**		**Utilize <> quando a o sinal de HA for de ≠**		
18				**Bicaudal é sinônimo de bilateral**		
19						

Teste_Z_Média | Teste_t_Média | Teste_Z_Proporção | +

Exercício 3

Um fabricante de *pen drives* garante que, no mínimo, 90% de seus produtos são perfeitos. Uma amostra de 200 unidades desse produto foi inspecionada e verificou que 30 delas eram defeituosas. Teste a afirmação do fabricante ao nível de significância de 2%. O que você pode concluir?

Proporção de *pen drives* perfeitos

↓

Dados resumidos

$n=200$; $x=170$; $\hat{p}=170/200=0{,}85$

↓

H_0: $p \geq 0{,}90$

H_a: $p < 0{,}90$

↓

$n.p_0=200.0{,}9=180$

$n.(1-p_0)=200.0{,}9=20$

↓

Distribuição Normal

Exercício 3

Proporção de *pen drives* perfeitos

↓

Dados resumidos

n=200; x=170; $\hat{p}$=170/200=0,85

↓

H_0: p $\geq$ 0,90

H_a: p < 0,90

↓

$n.p_0=200.0,9=180$

$n.(1-p_0)=200.0,9=20$

↓

Distribuição Normal

	A	B	C	D	E	F
1		**Teste de hipótese para uma proporção populacional**				
2						
3		n =	**200**	digite o tamanho da amostra		
4		$\hat{p}$=	**0,85**	digite a proporção de sucessos na amostra [0 ; 1]		
5		α=	**0,02**	digite o nível de significância do teste		
6		p_0 =	**0,9**	digite o valor da proporção hipótetica		
7		sinal em HA	**<**	digite o sinal na hipótese HA (<, > ou < >)		
8		H0 :	p >= 0,9	hipótese nula		
9		HA :	p < 0,9	hipótese alternativa		
10		tipo do teste	unicaudal esquerdo	tipo do teste (unicaudal ou bicaudal)		
11		Z =	-2,35702	estatística de teste		
12		valor-p	0,00921	valor-p do teste		
13						
14						
15		**Observações importantes:**		**Utilize <> quando a o sinal de HA for de ≠**		
16				**Bicaudal é sinônimo de bilateral**		
17						
18						
19						
20						
21						

Teste_Z_Média | Teste_t_Média | **Teste_Z_Proporção**

Exercício 3

Interpretação do valor p:

SE rejeitarmos H_0, cometemos um erro de 0,00921 ou 0,921%.

Conclusão do teste:

Considerando um nível de 2% de significância, podemos dizer que a proporção de *pen drives* perfeitos NÃO é de no mínimo 90%, ou seja, não podemos concordar com a afirmação do fabricante.

	A	B	C	D	E	F
1		**Teste de hipótese para uma proporção populacional**				
2						
3		n =	**200**	digite o tamanho da amostra		
4		$\hat{p}$=	**0,85**	digite a proporção de sucessos na amostra [0 ; 1]		
5		α=	**0,02**	digite o nível de significância do teste		
6		p_0 =	**0,9**	digite o valor da proporção hipótetica		
7		sinal em HA	**<**	digite o sinal na hipótese HA (<, > ou < >)		
8		H0 :	p >= 0,9	hipótese nula		
9		HA :	p < 0,9	hipótese alternativa		
10		tipo do teste	unicaudal esquerdo	tipo do teste (unicaudal ou bicaudal)		
11		Z =	-2,35702	estatística de teste		
12		valor-p	0,00921	valor-p do teste		
13						
14						
15		**Observações importantes:**		**Utilize <> quando a o sinal de HA for de ≠**		
16				**Bicaudal é sinônimo de bilateral**		
17						
18						
19						
20						
21						

Teste_Z_Média | Teste_t_Média | **Teste_Z_Proporção**

Exercício 3

Utilizando a função binom.test do R para a resolução desse exercício:

```
> binom.test(170, 200, p = 0.9, alternative = c("less"), conf.level = 0.98)

        Exact binomial test

data:  170 and 200
number of successes = 170, number of trials = 200, p-value = 0.01633
alternative hypothesis: true probability of success is less than 0.9
98 percent confidence interval:
 0.0000000 0.8983476
sample estimates:
probability of success
                  0.85
```

Exercício 3

Utilizando a função prop.test COM correção de continuidade de Yates:

```
> prop.test(170, 200, p = 0.9, alternative = c("less"), conf.level = 0.98,
correct=TRUE)

        1-sample proportions test with continuity correction

data:  170 out of 200, null probability 0.9
X-squared = 5.0139, df = 1, p-value = 0.01257
alternative hypothesis: true p is less than 0.9
98 percent confidence interval:
 0.0000000 0.8966986
sample estimates:
   p
0.85
```

Exercício 3

Utilizando a função prop.test SEM correção de continuidade de Yates:

```
> prop.test(170, 200, p = 0.9, alternative = c("less"), conf.level = 0.98,
correct=FALSE)

        1-sample proportions test without continuity correction

data:  170 out of 200, null probability 0.9
X-squared = 5.5556, df = 1, p-value = 0.009211
alternative hypothesis: true p is less than 0.9
98 percent confidence interval:
 0.0000000 0.8945941
sample estimates:
   p
0.85
```

Exercício 3

- ✓ Utilizando a distribuição Normal:

Valor p = 0,00921

- ✓ Utilizando a distribuição Binomial:

Valor p = 0,01633

- ✓ Utilizando a distribuição Qui-quadrado *com a correção* de Yates:

Valor p = 0,01257

- ✓ Utilizando a distribuição Qui-quadrado *sem a correção* de Yates:

Valor p = 0,009211

Testes de hipóteses no Python

Testes para uma e duas médias, variâncias e normalidade:
https://tmfilho.github.io/pyestbook/math/03_scip.html?highlight=test#testes-de-hipotese

Teste para uma proporção:
https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.binomtest.html#scipy.stats.binomtest

Ponderações importantes

1. A escolha do nível de significância deve ser realizada antes de iniciar o processo de teste de hipóteses;
2. Saber diferenciar e identificar as hipóteses estatísticas é fundamental;
3. Conhecer o conceito do valor p é essencial para a conclusão do teste;
4. Existe um mundo de testes de hipóteses diferentes! Saiba a indicação de uso de cada um deles para que você possa fazer escolhas certas!

PUC Minas
Virtual