Skill Lab

Curso de Data Science

Aula 06 - Estatística pt.4

O que você irá aprender nesta aula?

- Teste de hipóteses: P-valor
- Tipos de erros

Teste de hipóteses

Exemplo 1:

Uma empresa quer melhorar o desempenho do seu website.
As páginas demoram, em média, 3.125 segundos para carregar, com desvio padrão de 0.700 segundos. Eles contratam, então, um desenvolvedor para diminuir este tempo, e um cientista de dados para analisar os resultados.

A diretoria espera um nível de confiança de 99%. Uma amostragem foi realizada com 40 acessos à página, com uma média amostral de 2.875 segundos.
Podemos afirmar (com 99 % de certeza) que os tempos do teste estão estatisticamente menores do que anteriormente?

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

Teste de hipóteses

Exemplo 1: Método do P-valor

α = 0.01
Teste estatístico: **Z = -2.259**
P-valor: P(-2.259) = 0.0119

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

α < P
0.01 < 0.0119
Nível de significância **α** < P-valor

O teste estatístico está fora da área de rejeição

Teste de hipóteses

Exemplo 1: Método do P-valor

α = 0.01
Teste estatístico: **Z = -2.259**
P-valor: P(-2.259) = 0.0119

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

Afirmações:

- Portanto, NÃO conseguimos rejeitar a hipótese de que os tempos são maiores (é o que gostaríamos de fazer...)
- Não podemos dizer que as novas páginas são, estatisticamente, mais rápidas.

Teste de hipóteses

Exemplo 1: Comparação dos métodos

Tradicional

- Compara valor crítico com z-score do teste estatístico

Teste estatístico: **Z = -2.259**
Valor crítico: z = -2.326

P-valor

- Compara nível de significância alpha com P-valor

α = 0.01
Teste estatístico **Z = -2.259**
P-valor: P(-2.259) = 0.0119

Teste de hipóteses

Exemplo 2:

Uma fábrica de monitores de computador quer garantir que seu produto tenha um peso determinado. A equipe que projetou o monitor disse que ele deve pesar, em média, 3.125 kg, com desvio padrão de 0.700 kg (700 gramas).

A diretoria espera um nível de confiança de 99% na sua linha de produção. Eles contratam um cientista de dados para realizar um experimento. Uma amostragem foi realizada com 40 monitores escolhidos aleatoriamente na linha de produção. O peso das amostras foi medido, e obteve-se uma média de 2.875 kg.

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

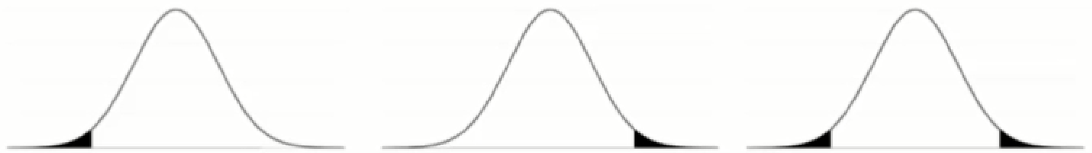
Podemos afirmar (com 99 % de certeza) que o peso médio dos monitores fabricados é estatisticamente igual a 3.125 kg?

Teste de hipóteses

Exemplo 2:

Hipótese nula (H0)	µ = 3.125
Hipótese alternativa (H1)	µ != 3.125

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$



Teste de hipóteses

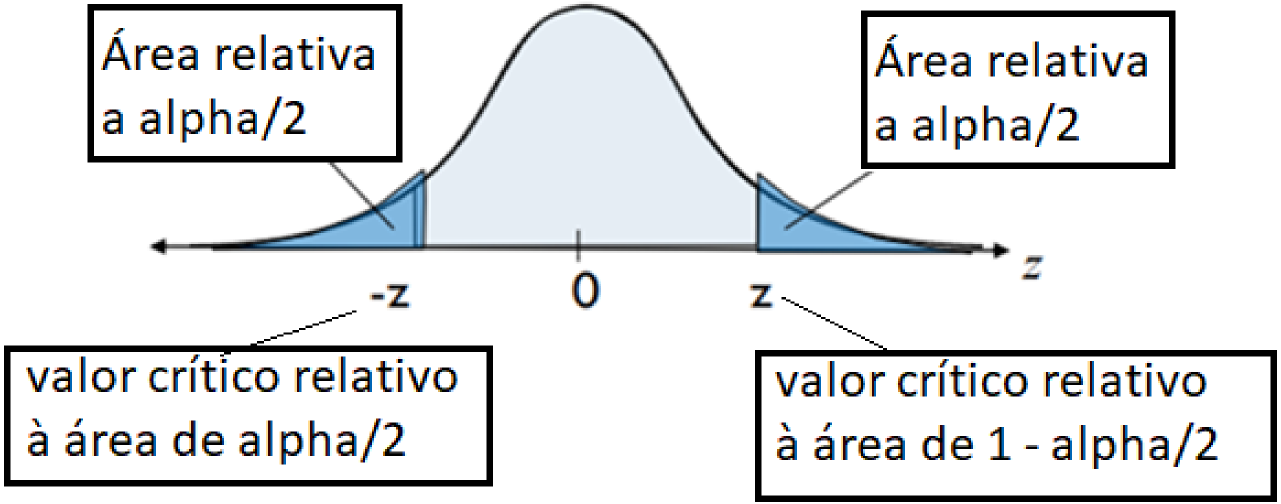
Exemplo 2:

Hipótese nula (H0)	μ = 3.125
Hipótese alternativa (H1)	μ != 3.125

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

Teste de hipóteses

Exemplo 2:



$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

Teste de hipóteses

Exemplo 2: Método tradicional

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

Teste estatístico:

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} =$$

Valor crítico:

Teste de hipóteses

Exemplo 2: Método tradicional

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

Teste estatístico:

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{2.875 - 3.125}{0.7/\sqrt{40}} = -2.259$$

Valores críticos:

```
z1 = stats.norm.ppf(alpha/2)
z2 = stats.norm.ppf(1 - alpha/2)
```

```
print(z1)
print(z2)
```

```
-2.5758293035489
2.5758293035489004
```

Teste de hipóteses

Exemplo 2: Método tradicional

Teste estatístico: **Z = -2.259**
Valores críticos: z1 = -2.576, z2 = 2.576

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

z1 **<** **Z** **<** z2

-2.576 **<** **-2.259** **<** 2.576

Valor crítico 1 < Teste estatístico **<** Valor crítico 2

O teste estatístico está fora da área de rejeição

Teste de hipóteses

Exemplo 2: Método tradicional

Teste estatístico: **Z = -2.259**
Valores críticos: z1 = -2.576, z2 = 2.576

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

Afirmações:

- O teste estatístico está fora da área de rejeição
- Logo, não conseguimos rejeitar H0

Teste de hipóteses

Exemplo 2: Método tradicional

Teste estatístico: **Z = -2.259**
Valores críticos: z1 = -2.576, z2 = 2.576

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

Afirmações:

- Portanto, aceitamos a afirmação H0 (mas não a provamos)
- Estatisticamente, e com nível de confiança de 99 %, pode-se dizer que a média de peso dos monitores fabricados é igual a 3.125 kg

Teste de hipóteses

Exemplo 2: Método do P-valor

$\alpha/2 = 0.005$, $1-\alpha/2 = 0.995$
Teste estatístico: **$Z = -2.259$**
P-valor: P(-2.259) = 0.0119

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

$$\alpha/2 < \text{P-valor} < 1-\alpha/2$$
$$0.005 < 0.0119 < 0.995$$

O teste estatístico está fora da área de rejeição

Teste de hipóteses

Exemplo 2: Método do P-valor

$\alpha/2 = 0.005$, $1-\alpha/2 = 0.995$
Teste estatístico: **$Z = -2.259$**
P-valor: P1(-2.259) = 0.0119

$$\mu = 3.125$$
$$\sigma = 0.700$$
$$\alpha = 1 - 0.99 = 0.01$$
$$n = 40$$
$$\bar{x} = 2.875$$

Afirmações:

- Portanto, aceitamos a afirmação H0 (mas não a provamos)
- Estatisticamente, e com nível de confiança de 99 %, pode-se dizer que a média de peso dos monitores fabricados é igual a 3.125 kg

Teste de hipóteses

Considerações finais: Conclusões do teste de hipóteses

	Afirmação	
Decisão	Afirmação é H_0	Afirmação é H_1
Rejeitar H_0	Há evidência suficiente para rejeitar a afirmação.	Há evidência suficiente para apoiar a afirmação.
Não rejeitar H_0	Não há evidência suficiente para rejeitar a afirmação.	Não há evidência suficiente para apoiar a afirmação.

Teste de hipóteses

Desafio 1: crie um programa Python que automatiza o teste de hipóteses para o caso do exemplo 2, em que se espera um valor fixo para a média. Sugere-se utilizar a resolução “tradicional”.
O programa deve receber os dados de média populacional, desvio padrão populacional, nível de significância alpha, número de elementos da amostra e média da amostra.
Ao final, deve imprimir “Rejeito H0” ou “Não rejeito H0”, dependendo do resultado do teste de hipóteses.

Teste de hipóteses

Teste do desafio 1: a mesma empresa faz o projeto de uma tv 4K, voltado agora ao público de entretenimento.
No projeto, a tv deveria ter 50 polegadas, com desvio padrão de 2 polegadas. Realiza-se uma amostragem com 25 tvs na linha de produção, e se encontra o valor de média amostral de 51.3 polegadas. Com um nível de significância de 5 %, podemos afirmar que as tvs possuem o tamanho médio de 50 polegadas?

Resposta: devemos rejeitar H0 (estatisticamente, as tvs não parecem ter, em média, 50 polegadas, a um nível de significância de 5%)

Teste de hipóteses

Considerações finais: tipos de erros

	A verdade de H_0	
Decisão	H_0 é verdadeira	H_0 é falsa
Não rejeitar H_0	Decisão correta	**Erro tipo II**
Rejeitar H_0	**Erro tipo I**	Decisão correta

Teste de hipóteses

Desafio 2: deve-se realizar uma pesquisa com um número indefinido de candidatos, em duas etapas. Eles são escolhidos aleatoriamente durante o período de um dia, entre participantes de uma feira de tecnologia. Em um primeiro momento, os candidatos são entrevistados. Deve-se selecionar, aleatoriamente, 10.5% dos candidatos para participar da segunda etapa da pesquisa. No entanto, cada candidato deve receber a informação se vai ou não participar da segunda etapa NO MOMENTO em que forem entrevistados. Como fazer isso?

Então, nesta aula vimos:

- Teste de hipóteses: P-valor
- Tipos de erros

Muito obrigado!

Skill Lab