Estatística para Cursos de Engenharia e Informática

Pedro Alberto Barbetta / Marcelo Menezes Reis / Antonio Cezar Bornia

São Paulo: Atlas, 2004

Cap. 8 – Testes de hipóteses

APOIO:
Fundação de Apoio à Pesquisa Científica e Tecnológica do Estado de Santa Catarina (FAPESC)
Departamento de Informática e Estatística – UFSC (INE/CTC/UFSC)

Teste de hipóteses

População

Conjectura (hipótese) sobre o comportamento de variáveis

Decisão sobre a admissibilidade da hipótese

Amostra

Resultados reais obtidos

Hipóteses

a) Substituindo o processador *A* pelo processador *B*, altera-se o tempo de resposta de um computador.

b) Aumentando a dosagem de cimento, aumenta-se a resistência do concreto.

c) Uma certa campanha publicitária produz efeito positivo nas vendas.

d) A implementação de um programa de melhoria da qualidade em uma empresa prestadora de serviços melhora a satisfação de seus clientes.

Hipóteses em termos de parâmetros

a) A *média* dos tempos de resposta do equipamento com o processador *A é diferente* da *média* dos tempos de resposta com o processador *B*.

b) A *média* dos valores de resistência do concreto com a dosagem d_2 de cimento *é maior* do que a *média* dos valores de resistência com a dosagem d_1.

c) A *média* das vendas depois da campanha publicitária *é maior* do que a *média* das vendas antes da campanha publicitária.

d) A *proporção* de reclamações após a realização do programa de melhoria da qualidade *é menor* do que antes da realização do programa.

Hipóteses nulas

a) $H_0: \mu_A = \mu_B$ e $H_1: \mu_A \neq \mu_B$

onde:

μ_A é o tempo médio de resposta com o processador *A*; e

μ_B é o tempo médio de resposta com o processador *B*.

b) $H_0: \mu_2 = \mu_1$ e $H_1: \mu_2 > \mu_1$

onde:

μ_2 é a resistência média do concreto com a dosagem d_2 de cimento; e

μ_1 é a resistência média do concreto com a dosagem d_1 de cimento.

Hipóteses nulas

c) H_0: $\mu_2 = \mu_1$ e H_1: $\mu_2 > \mu_1$

onde: μ_1 é o valor médio das vendas antes da campanha publicitária; e

μ_2 é o valor médio das vendas depois da campanha publicitária.

d) H_0: $p_2 = p_1$ e H_1: $p_2 < p_1$

onde: p_1 é a proporção de reclamações antes do programa de melhoria da qualidade; e

p_2 é a proporção de reclamações depois do programa de melhoria da qualidade.

Conceitos básicos

Exemplo:

- Suspeita-se que uma moeda não seja perfeitamente equilibrada (probab. de cara ≠ probab. de coroa ≠ 0,5)

p = probab. de cara

H_0: $p = 0{,}5$

H_1: $p \neq 0{,}5$

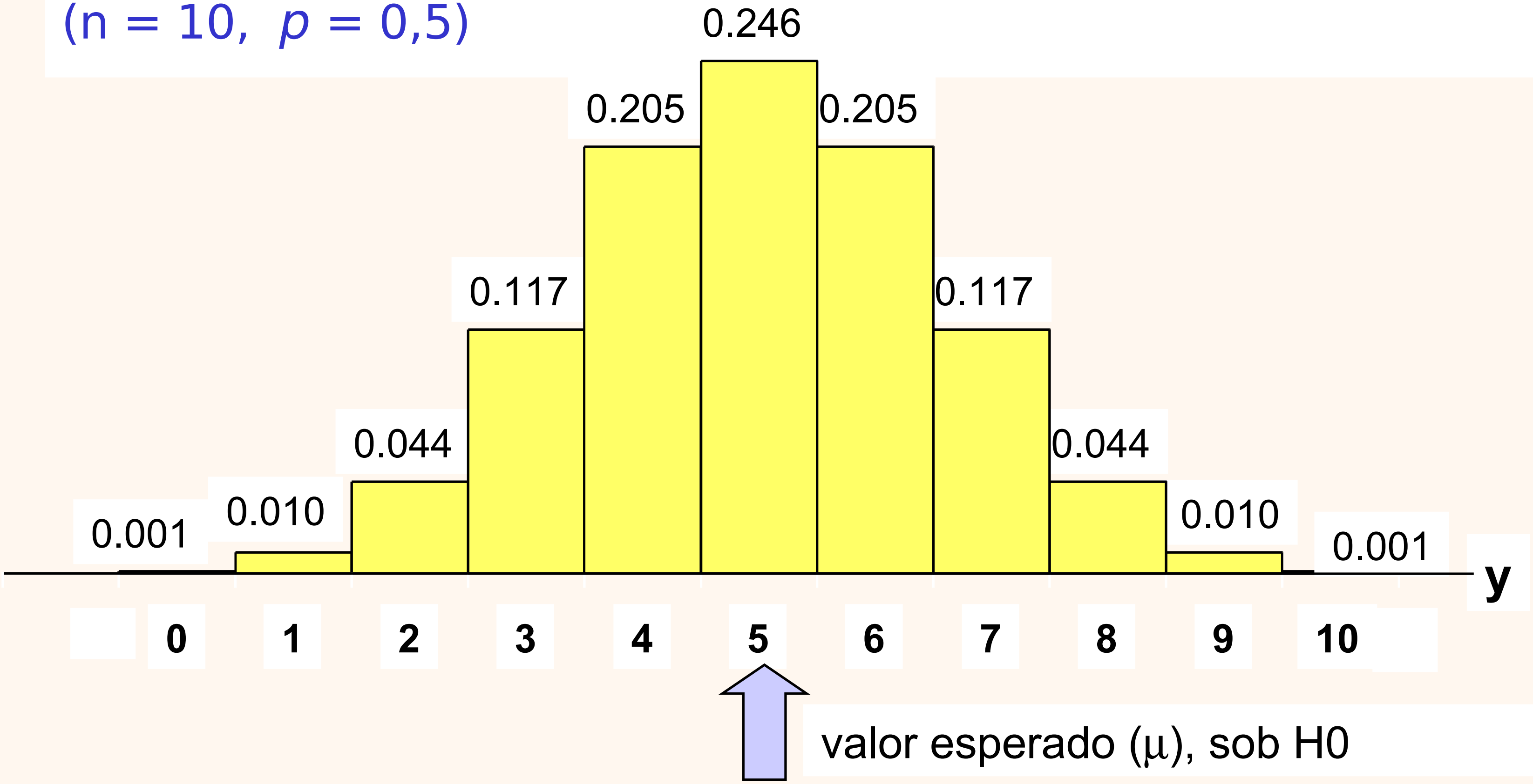
Resultado da amostra

- **Situação 1**: Valor obtido: **y = 10** caras.

  Qual seria a conclusão?

Exemplo da moeda

Distribuição
binomial
(n = 10, p = 0,5)

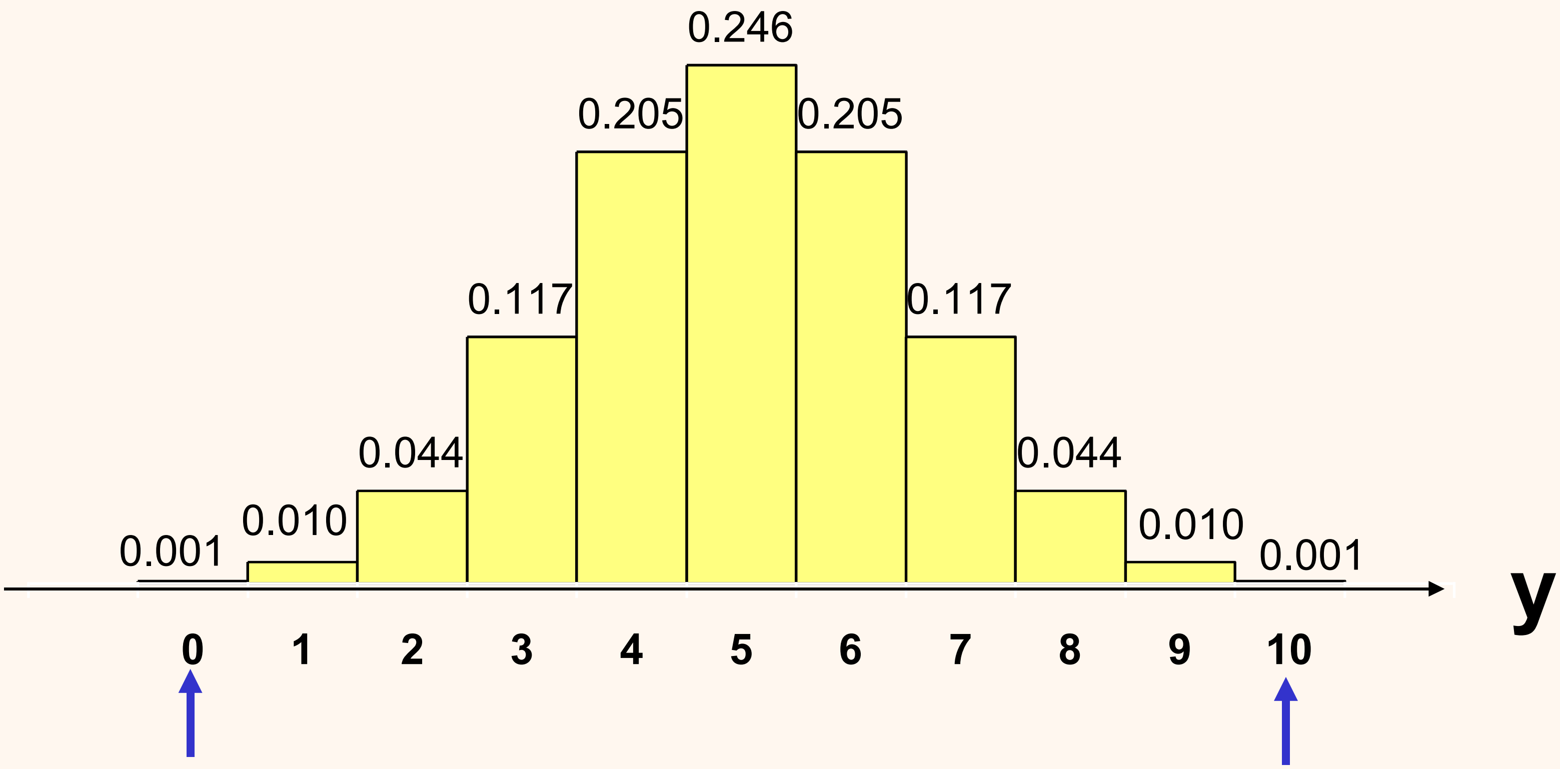


BARBETTA, REIS e BORNIA – Estatística para Cursos de Engenharia e Informática. Atlas, 2004

Probabilidade de Significância ou valor p

- Probabilidade da estatística do teste acusar um resultado tão (ou mais) distante do esperado quanto o resultado ocorrido na amostra observada, supondo H_0 como a hipótese verdadeira.

Situação 1



Valor **p** = 0,002 ou 0,2%

Conclusão...

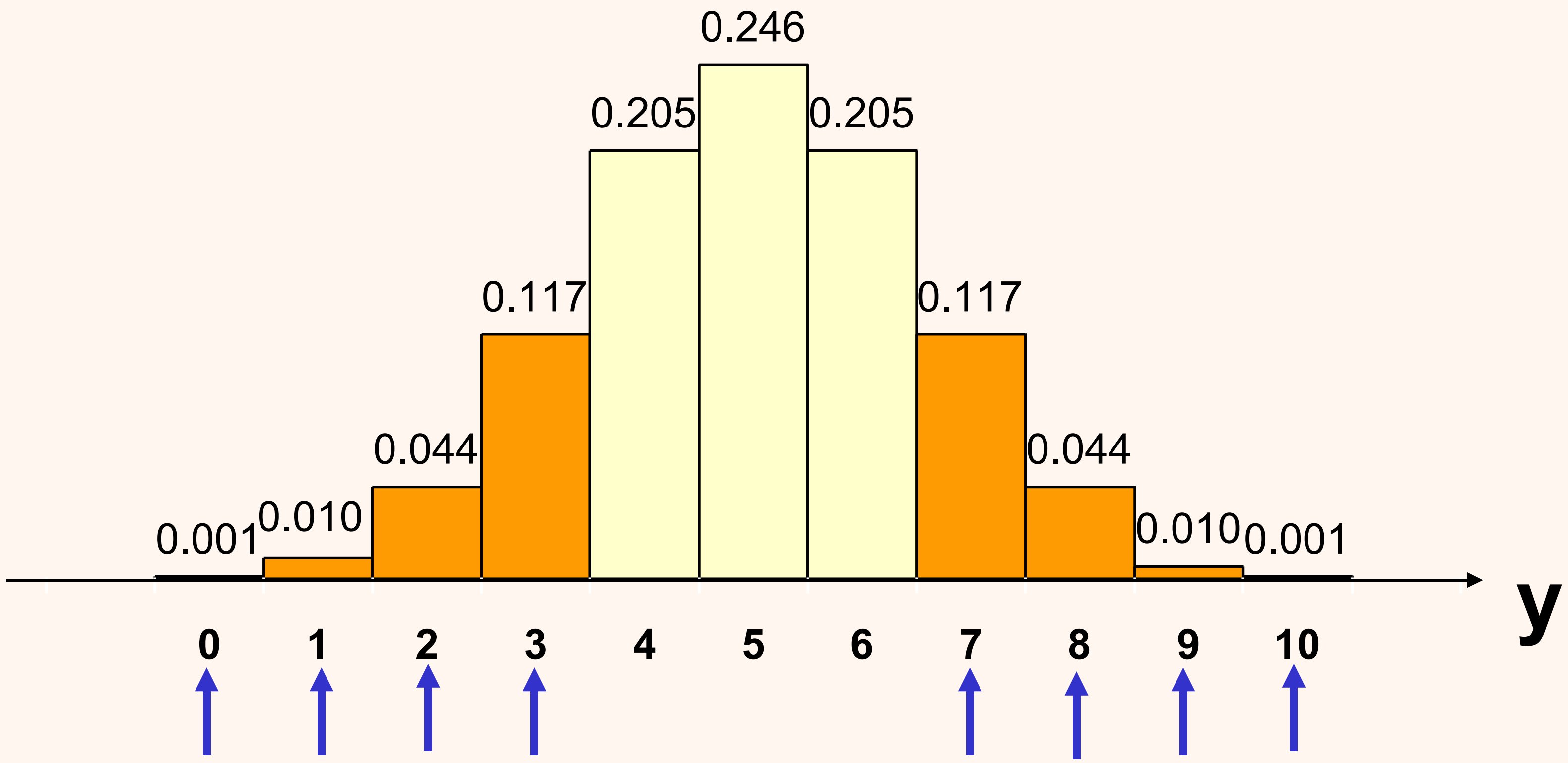
- Valor **p** = 0,2% (probabilidade de uma moeda honesta acusar um valor tão distante quanto ao que se observou na amostra). Probabilidade muito pequena!!!
- Qual é a conclusão?
- O teste rejeita H_0, ou seja, prova-se estatisticamente que a moeda é viciada.

Resultado da amostra

Situação 2: Valor obtido: **y = 7** caras.

Qual seria a conclusão?

Situação 2



Valor **p** = 0,344 ou 34,4%

Conclusão...

- Valor **p** = 34,4% (probabilidade de uma moeda honesta acusar um valor tão distante quanto ao que se observou na amostra). Não é muito pequeno!!!
- Qual é a conclusão?
- O teste aceita H_0, ou seja, não se pode afirmar que a moeda é viciada.

Nível de Significância (α)

- Representa a probabilidade tolerável de se rejeitar H_0 quando esta for verdadeira.

- Os valores mais comuns para o nível de significância são **5%,** 10% e 1%.

Regra de decisão

- **Valor p ≤ α**
- **Valor p > α**

Rejeita H_0 (prova-se estatisticamente H_1)

Aceita H_0 (os dados não mostram evidência para afirmar H_1)

Exercício

- Para testar se existe diferença entre dois sistemas computacionais (A e B), observou-se o desempenho com **12** cargas de trabalho. Em **3** casos o sistema A apresentou melhor desempenho do que o B. Nos demais, o sistema B foi melhor. Qual a conclusão ao nível de significância de 5%?

Exercício - Resp.

- Hipóteses:

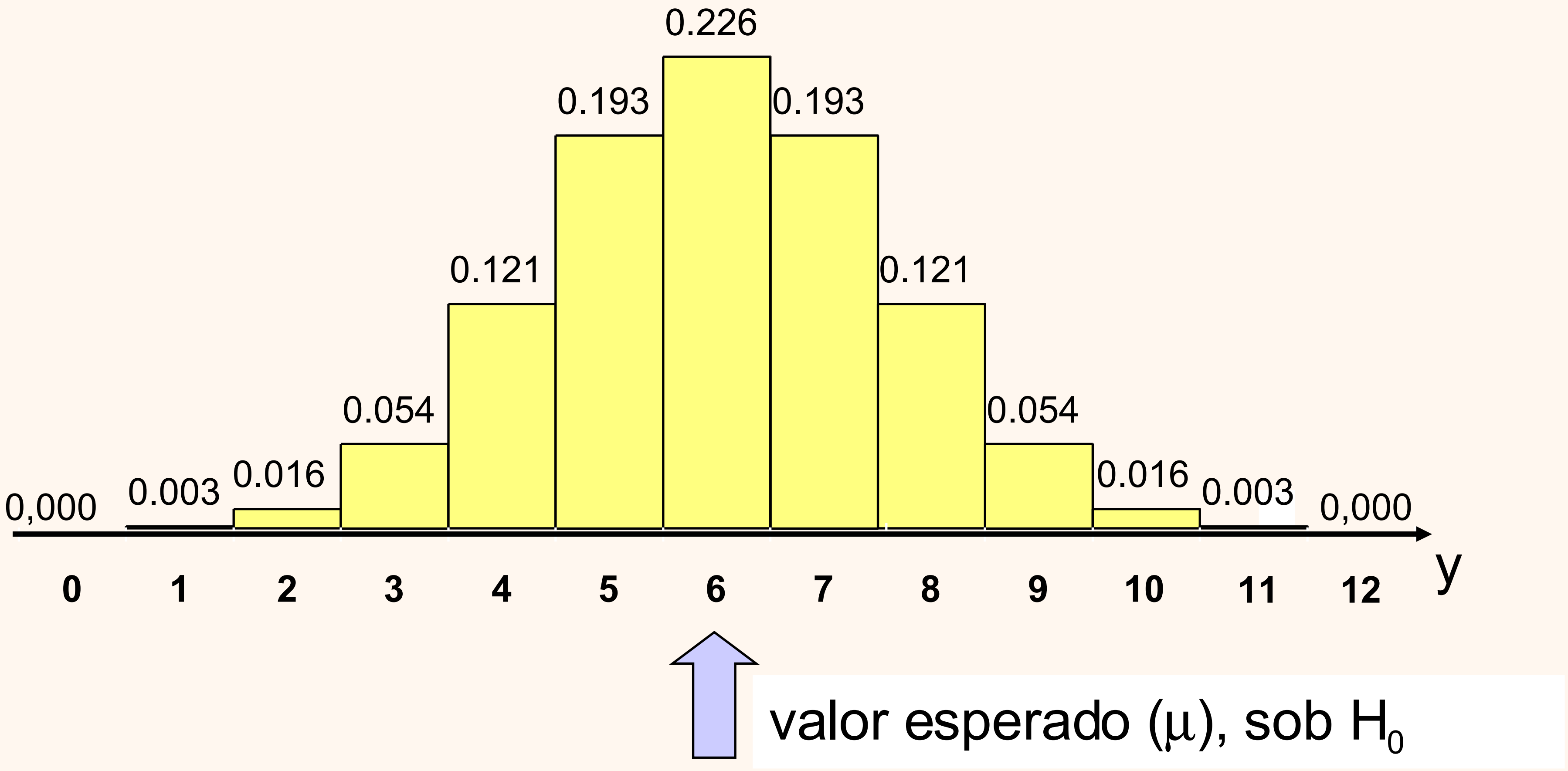
H_o: p = 0,5

H_1: p ≠ 0,5

p = probabilidade do sistema A apresentar melhor desempenho do que o sistema B.

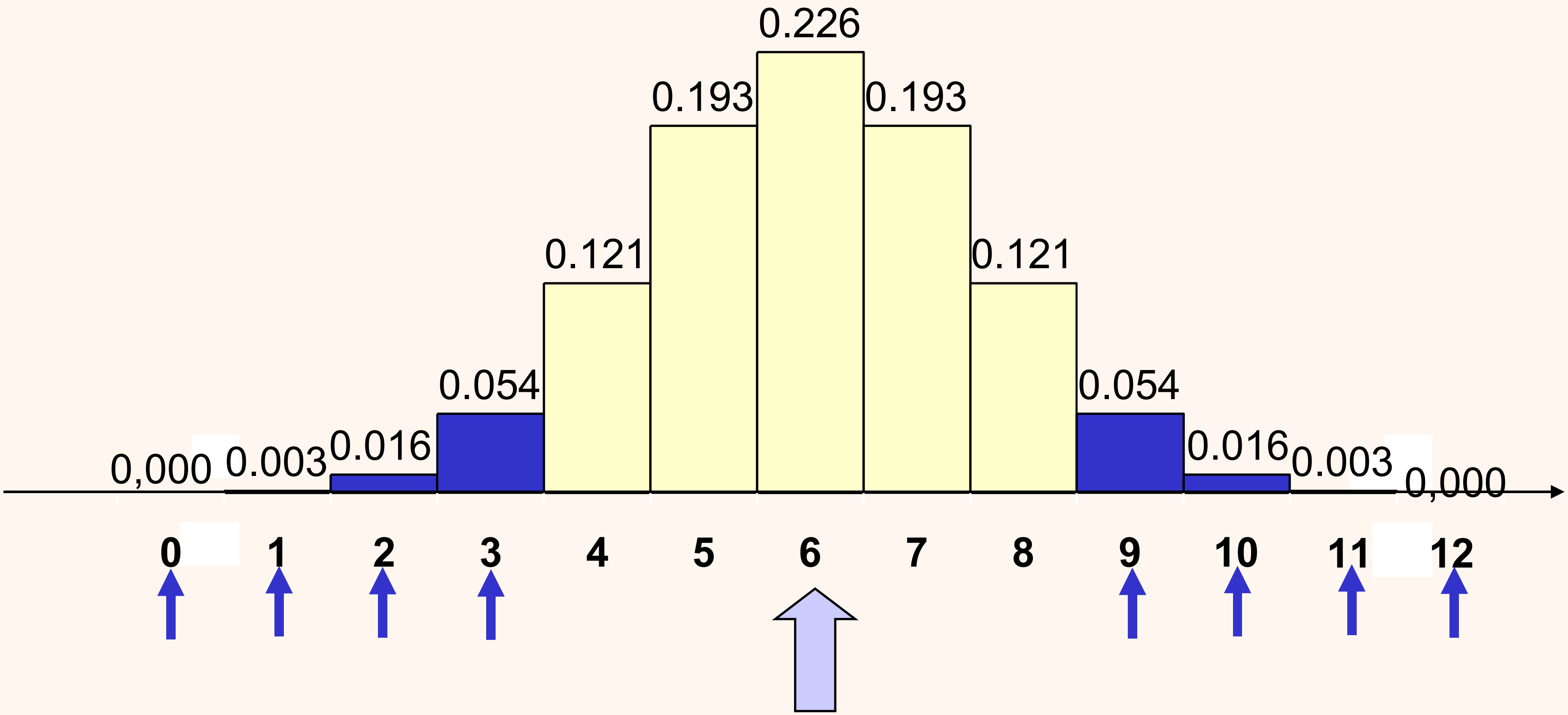
Exercício - Resp.

- Distribuição binomial (n = 12, p = 0,5).



Exercício - Resp.

- Valor **p** = P{(X < 3) ou (X > 9)}



Valor **p** = 0,146 ou 14,6%

Exercício - Resp.

- Valor **p** = 14,6%   **> 5%**   (α = 5%)
- O teste **aceita** H_0, ao nível de significância de 5%.
- **Não** se pode afirmar (ao nível de significância de 5%) que existe diferença entre os dois tipos de sistemas, em termos de desempenho.

Tipos de erro num teste estatístico

Realidade (desconhecida)	Decisão do teste	
	aceita H_0	rejeita H_0
H_0 verdadeira	decisão correta (probab = $1 - \alpha$)	erro tipo I (probab = α)
H_0 falsa	erro tipo II (probab = β)	decisão correta (probab = $1 - \beta$)

Tipos de erro num teste estatístico

Realidade (desconhecida)	Decisão do teste	
	aceita H_0	rejeita H_0
H_0 verdadeira	decisão correta (probab = $1 - \alpha$)	erro tipo I (probab = α)
H_0 falsa	erro tipo II (probab = β)	decisão correta (probab = $1 - \beta$)

$$P(\text{erro tipo I}) = P(\text{rejeitar } H_0 \mid H_0 \text{ é verdadeira}) = \alpha$$

$$P(\text{erro tipo II}) = P(\text{aceitar } H_0 \mid H_0 \text{ é falsa}) = \beta$$

Abordagem clássica

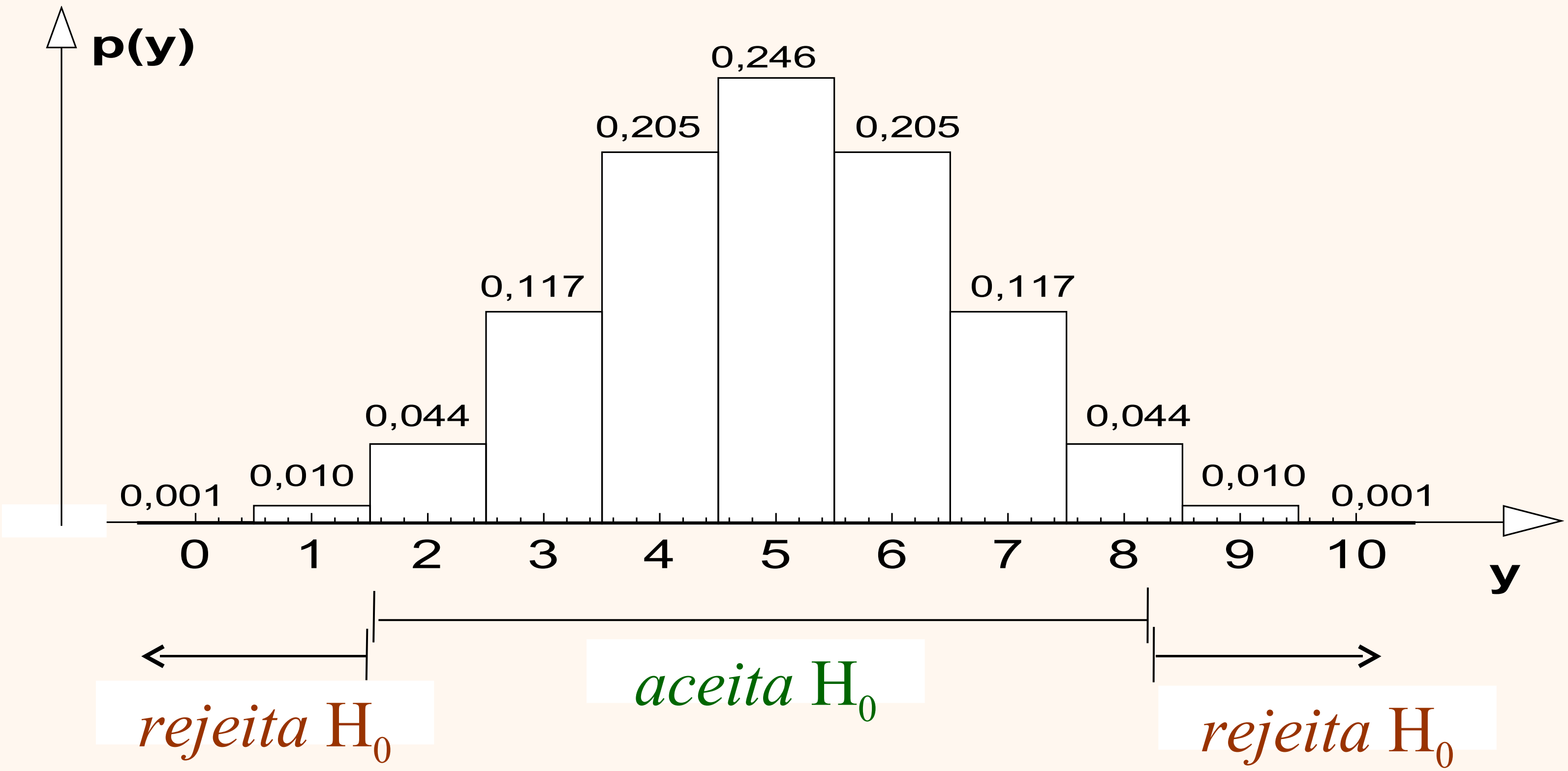
- Constrói a regra de decisão antes de observar a amostra

- Retomando o experimento de lançar 10 vezes a moeda, a regra de decisão para $\alpha = 0{,}05$ é construída com base na equação:

P (erro tipo I) = P (rejeitar H_0 | H_0 é verdadeira) = α = 0,05

Abordagem clássica

Regra de decisão em termos de *Y = número de caras em 10 lançamentos da moeda*, com $\alpha = 0{,}05$.

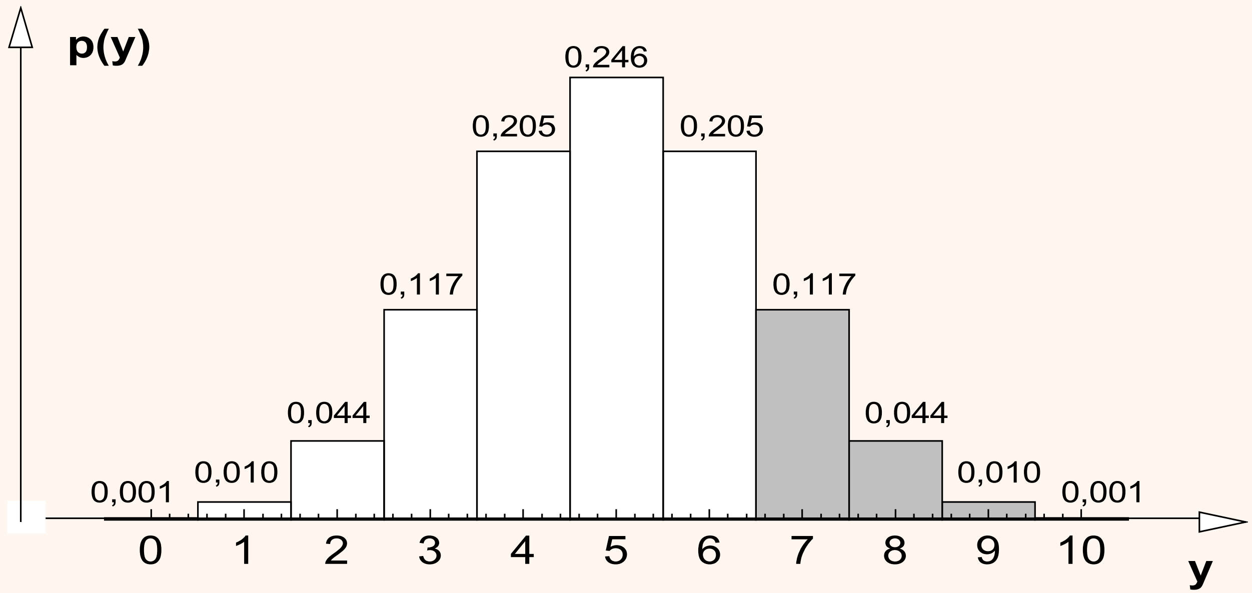


Testes unilaterais

- Quando a hipótese alternativa tem sinal > ou < (pelas características do problema em estudo).
- Ex.
  - H_0: $p = 0,5$ (a moeda é honesta) e
  - H_1: $p > 0,5$ (a moeda tende a dar mais caras do que coroas).

Testes unilaterais

- Cálculo do valor **p**, considerando n = 10:



Valor **p** = p(7) + p(8) + p(9) + p(10) = 0,172

Teste para proporção

- H_0: $p = p_0$ e H_1: $p \neq p_0$ (p_0 é um valor dado)

- No caso de teste unilateral, a hipótese alternativa seria H_1': $p > p_0$ (unilateral à direita) ou H_1'': $p < p_0$ (unilateral à esquerda).

- Suponha amostra suficientemente grande para aproximação da binomial à normal:
  $n.p_0 \geq 5$ e $n.(1 - p_0) \geq 5$

Teste para proporção

- Sejam:

$$\hat{p} = \frac{y}{n} = \frac{\text{número de elementos com } o \text{ atributo de interesse}}{n}$$

$$y' = y - 0{,}5 \quad \text{se } y > n.p_0; \quad \text{ou}$$

$$y' = y + 0{,}5 \quad \text{se } y < n.p_0 \quad \text{(correção de continuidade)}$$

- Cálculo da estatística do teste:

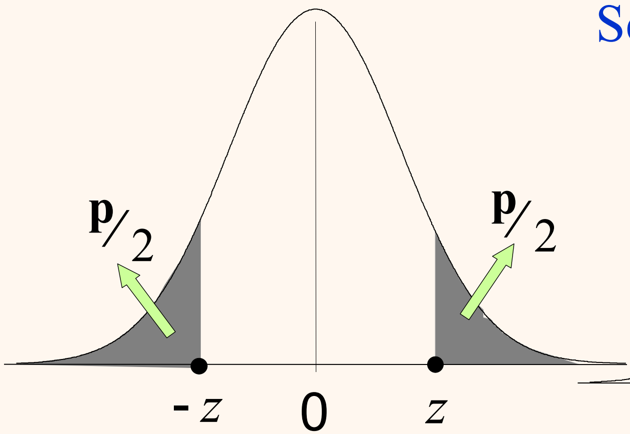
$$z = \frac{y' - n.p_0}{\sqrt{n.p_0.(1 - p_0)}}$$

Teste para proporção – abordagem do valor **p**

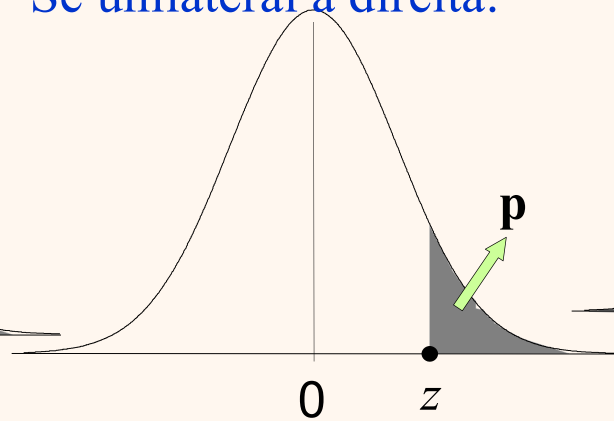
Amostra → Cálculo de z → Obtenção de **p** pela tabela da normal

$$z = \frac{y' - n.p_0}{\sqrt{n.p_0.(1 - p_0)}}$$

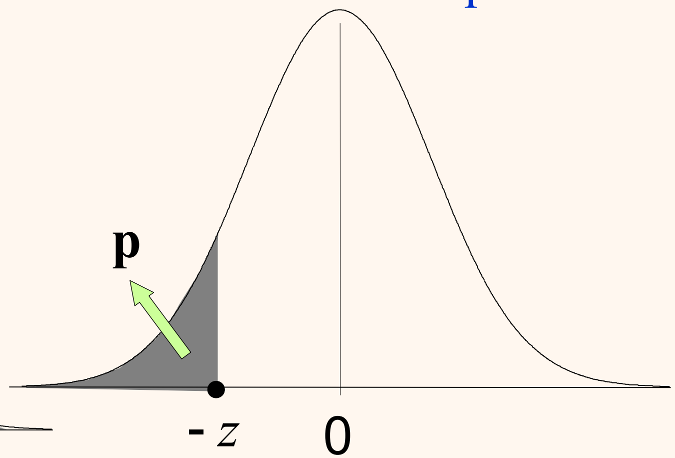
Se bilateral:



Se unilateral à direita:



Se unilateral à esquerda:



Teste para proporção – abordagem do valor **p**

Valor **p** $> \alpha$ $\Longrightarrow$ aceita H_0

Valor **p** $\leq \alpha$ $\Longrightarrow$ rejeita H_0

Teste para proporção – Exemplo 8.6:

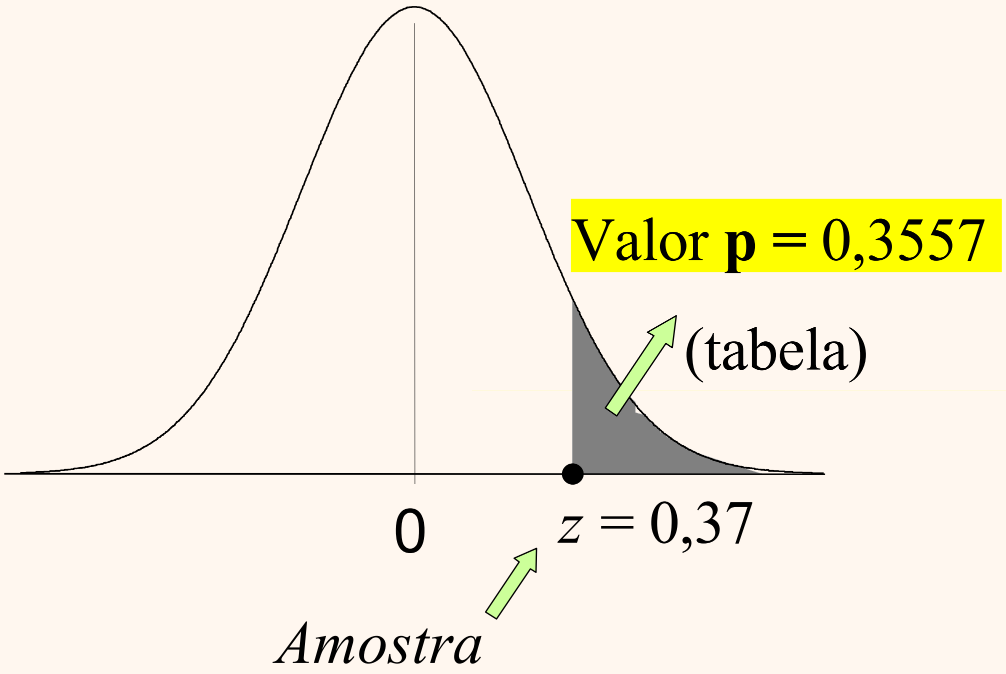
(ver enunciado no livro)

- H_0: p = 0,015 e H_1: p > 0,015. Usar α = 0,01.
- Amostra: y = 9 em n = 500.

$$\hat{p} = \frac{9}{500} = 0{,}018$$

$$z = \frac{y' - n.p_0}{\sqrt{n.p_0.(1-p_0)}} = \frac{8{,}5-(500).(0{,}015)}{\sqrt{(500).(0{,}015).(1-0{,}015)}} = \frac{1}{2{,}718} \approx 0{,}37$$

Teste para proporção – Exemplo 8.6:



- Aceita H_0 ao nível de significância de 1%.

Teste para proporção – abordagem clássica

Nível de significância α → Obtenção do valor crítico z_c pela tabela da normal → ...

Valores usuais de z_c, obtidos da distribuição normal padrão:

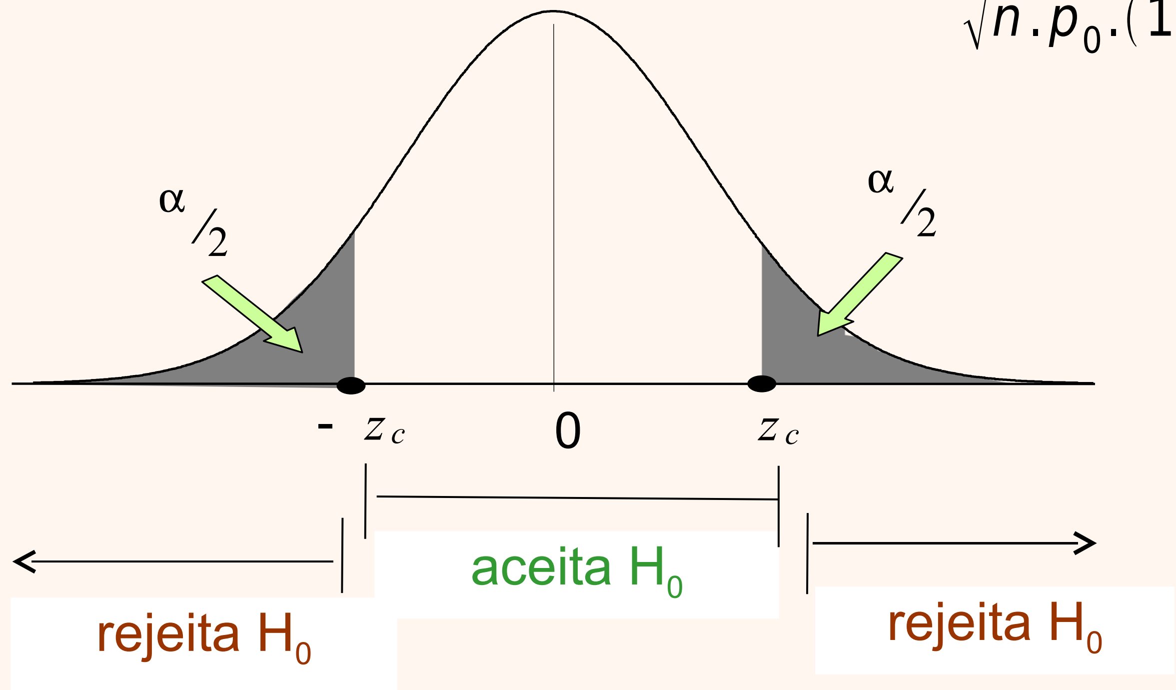
teste bilateral, α:	0,20	0,10	0,05	0,02	0,01	0.005
teste unilateral, α:	0,10	0,05	0,025	0,01	0,005	0,0025
valor crítico (z_c):	1,282	1,645	1,960	2,326	2,576	2,807

Teste para proporção – abordagem clássica

Nível de significância α → Obtenção do valor crítico z_c pela tabela da normal → Cálculo do valor z

$$z = \frac{y' - n.p_0}{\sqrt{n.p_0.(1-p_0)}}$$

Se bilateral:

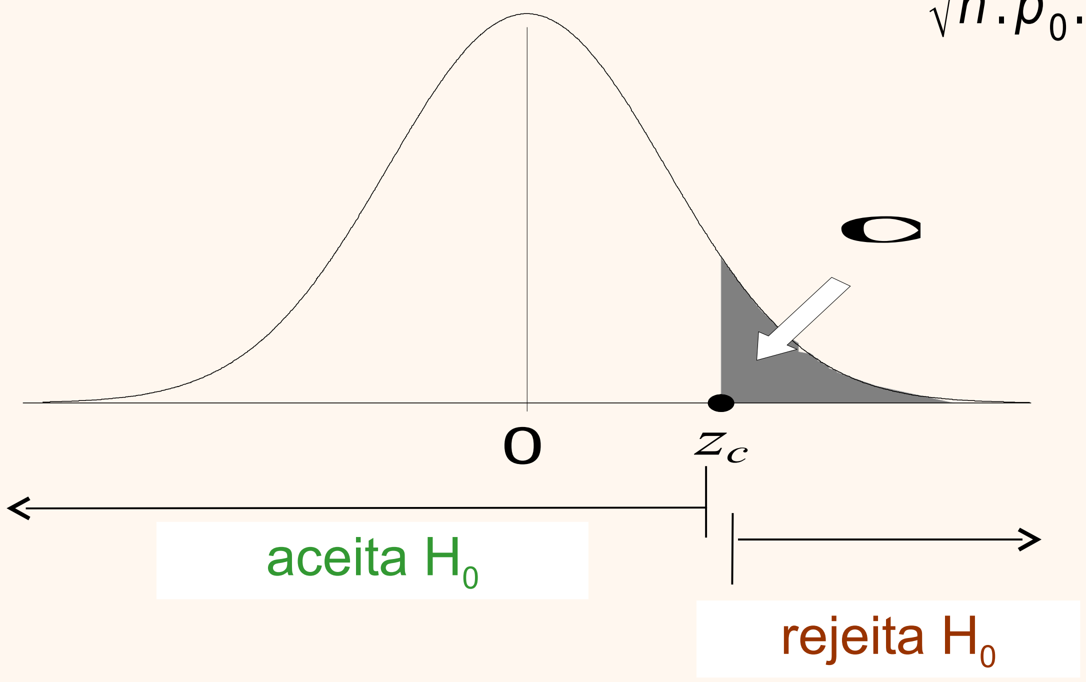


Teste para proporção – abordagem clássica

Nível de significância α → Obtenção do valor crítico z_c pela tabela da normal → Cálculo do valor z

$$z = \frac{y' - n.p_0}{\sqrt{n.p_0.(1-p_0)}}$$

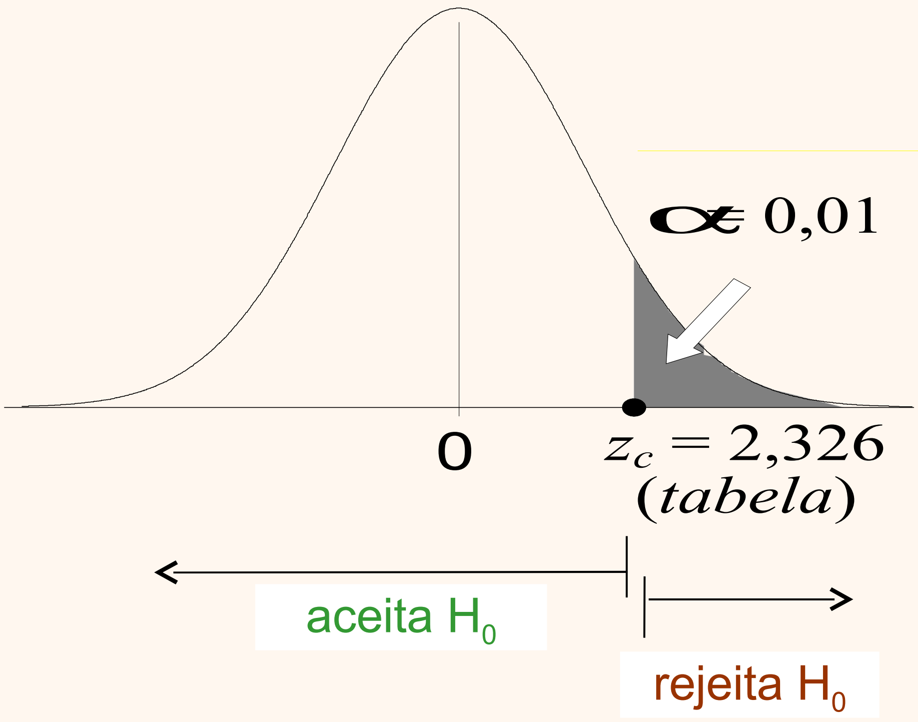
Se unilateral à direita:



Teste para proporção – Exemplo 8.6:

(ver enunciado no livro)

- H_0: p = 0,015   e H_1: p > 0,015. Usar α = 0,01.
- Regra de decisão:



Teste para proporção – Exemplo 8.6:

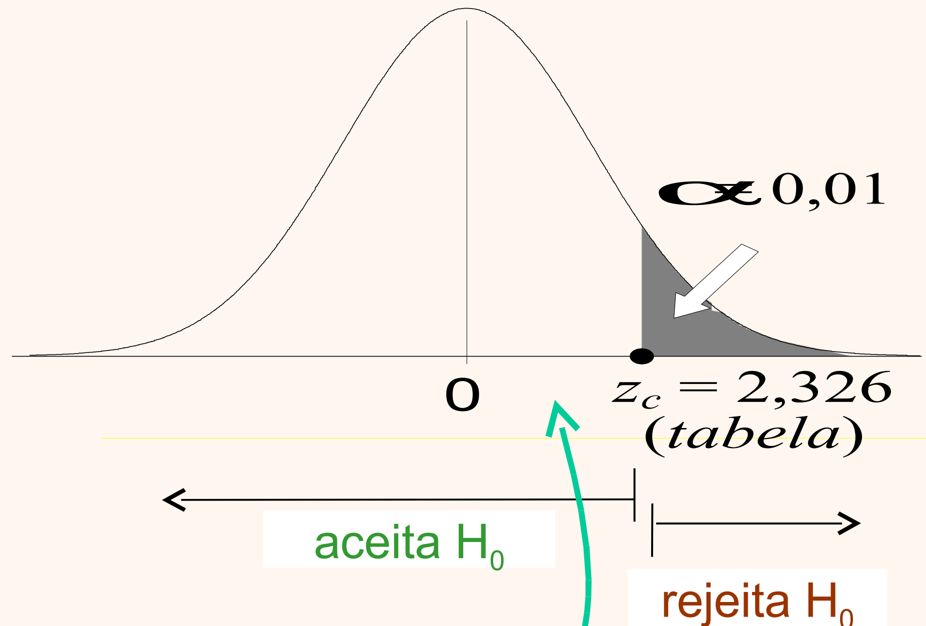
- Amostra: $y = 9$ em $n = 500$.

$$\hat{p} = \frac{9}{500} = 0,018$$

$$z = \frac{y' - n.p_0}{\sqrt{n.p_0.(1-p_0)}} = \frac{8,5 - (500).(0,015)}{\sqrt{(500).(0,015).(1-0,015)}} = \frac{1}{2,718} \approx 0,37$$

Teste para proporção – Exemplo 8.6:

- Conclusão:



Da amostra:

$$z = \frac{y' - n.p_0}{\sqrt{n.p_0.(1-p_0)}} \approx 0,37$$

➔ Aceita H_0.

(Ver comentários práticos no livro.)

Teste para média

- H_0: $\mu = \mu_0$ e H_1: $\mu \neq \mu_0$

- No caso de teste unilateral, a hipótese alternativa seria H1': $\mu > \mu_0$ (unilateral à direita) ou H1'': $\mu < \mu_0$ (unilateral à esquerda).

Teste para média – Caso de variância conhecida

- Cálculo da estatística do teste:

$$z = \frac{\left(\bar{x} - \mu_0\right) \cdot \sqrt{n}}{\sigma}$$

onde: μ_0 é o valor da média segundo H0;

n é tamanho da amostra;

σ é o desvio padrão populacional; e

$\bar{x}$ é a média da amostra.

O teste é feito com a distribuição normal, análogo ao da proporção.

Teste para média – Caso de variância desconhecida

- Cálculo da estatística do teste:

$$t = \frac{\left(\bar{x} - \mu_0\right) \cdot \sqrt{n}}{s}$$

onde: μ_0 é o valor da média segundo H0;

n é tamanho da amostra;

s é o desvio padrão da amostra; e

$\bar{x}$ é a média da amostra.

Uso da distribuição t com gl = n – 1
(supondo população com distribuição normal)

Teste para média – Caso de variância desconhecida
Exemplo 8.8 (ver enunciado no livro):

- H_0: $\mu = 7{,}4$ s
- H_1: $\mu < 7{,}4$ s

- Amostra:

  n = 10;

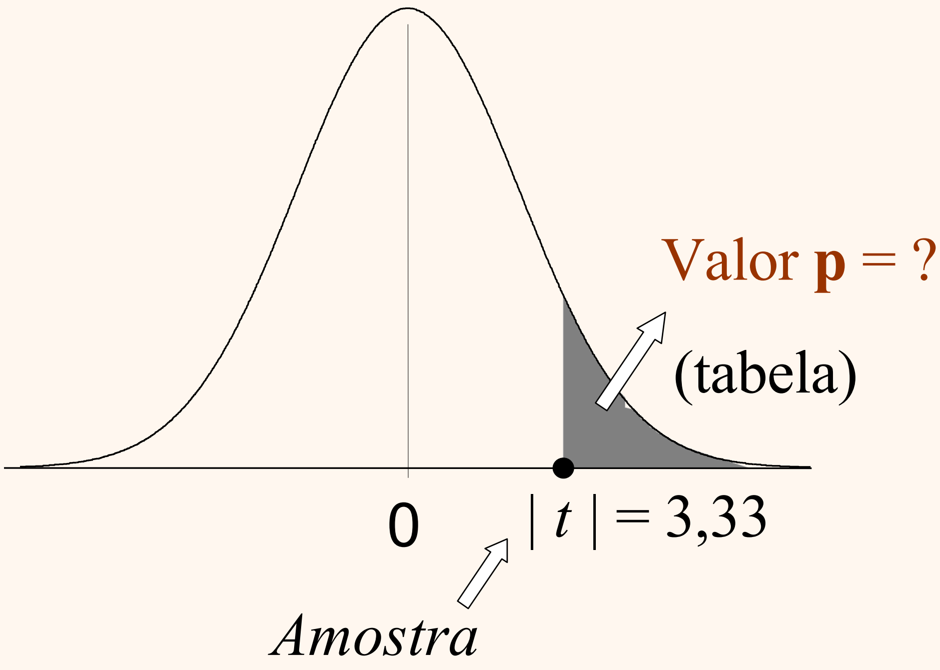
  média da amostra = 6,82;

  desvio padrão da amostra = 0,551

$$t = \frac{(\bar{x} - \mu_0)\cdot\sqrt{n}}{s} = \frac{(6{,}82 - 7{,}4)\cdot\sqrt{10}}{0{,}551} = -3{,}33$$

Teste para média – Caso de variância desconhecida
Exemplo 8.8 (ver enunciado no livro):

- Uso da tabela t para obter o valor **p**:



Teste para média – Caso de variância desconhecida
Exemplo 8.8. Abordagem do valor **p**:

- Uso da tabela t para obter o valor **p**:

dados observados

$|t| = 3{,}33$

	Área na cauda superior					
gl	0,25	...	0,01	0,005	0,0025	...
...						
9	0,703	...	2,821	3,250	3,690	...
...						

➔ 0,0025 < valor **p** < 0,005 ➔ valor **p** < 0,01

➔ Teste rejeita H_0. (Ver comentários práticos no livro.)