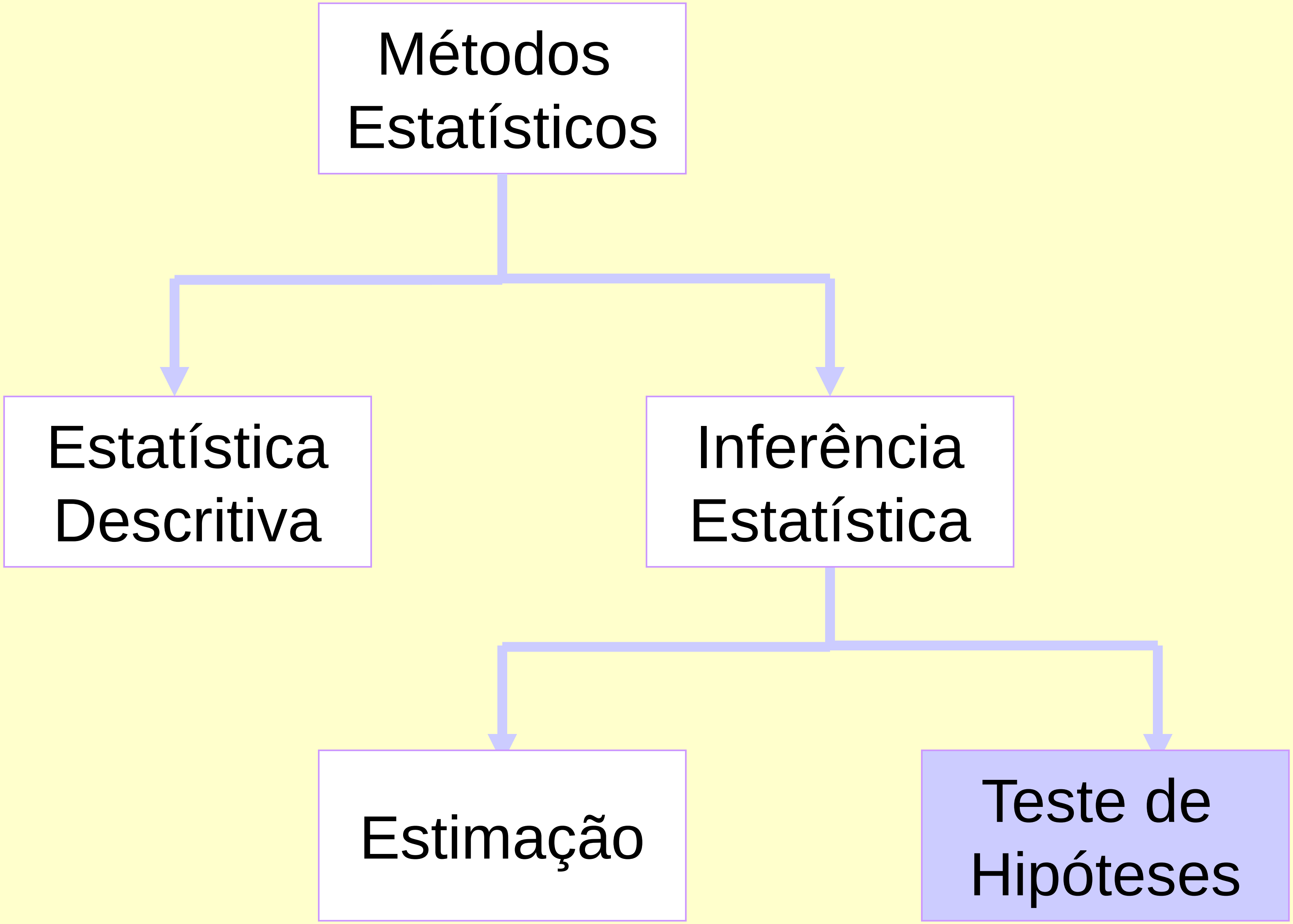
NOÇÕES DE TESTE DE HIPÓTESES (I)

Teste de hipóteses para a proporção populacional

Métodos Estatísticos



TESTE DE HIPÓTESES

O que é uma hipótese?

- É uma conjectura sobre um parâmetro populacional.

  Por exemplo, a proporção p é um parâmetro populacional.

- A hipótese deve ser estabelecida antes da análise.

Estimação	Teste de Hipóteses
Qual é a probabilidade de “cara” no lançamento de uma moeda?	A moeda é honesta ou é desequilibrada?
Qual é a proporção de votos que o candidato A terá na próxima eleição?	O candidato A vencerá a eleição?
Qual é a proporção de motoristas habilitados de SP que tiveram suas carteiras apreendidas após a vigência da nova lei de trânsito?	A proporção dos moto-ristas habilitados de SP que tiveram suas carteiras apreendidas após a nova lei é maior que 2% ou não?

Introdução

Em **estimação,** o objetivo é "estimar" o valor desconhecido de um parâmetro, por exemplo, a proporção $\boldsymbol{p}$ de "indivíduos" em uma população com determinada característica.

A estimativa é baseada no número $\boldsymbol{x}$ de "indivíduos" com a característica numa amostra aleatória de tamanho $\boldsymbol{n}$.

Entretanto, se o objetivo for saber se o valor observado $\boldsymbol{x}$ nessa amostra dá ou não suporte a uma conjectura sobre o valor de $\boldsymbol{p}$, trata-se de um **teste de hipóteses.**

Exemplo 1: Queremos avaliar se uma moeda é honesta.

Ou seja, queremos testar a

hipótese nula H: a moeda é honesta

contra a

hipótese alternativa A: a moeda não é honesta.

Em linguagem estatística, essas hipóteses podem ser reescritas como:

$$H: \boldsymbol{p} = 0{,}5$$

$$A: \boldsymbol{p} \neq 0{,}5$$

com $\boldsymbol{p}$ a probabilidade de "cara" da moeda.

Hipóteses

De maneira geral, uma **hipótese estatística** é uma afirmação ou conjectura sobre um parâmetro da distribuição de probabilidades de uma variável aleatória.

No caso especial de teste de hipóteses sobre a proporção populacional p, temos:

Hipótese nula: afirmação sobre p geralmente relacionada a um valor de referência, ou a uma especificação padrão ou histórica.

Hipótese alternativa: afirmação sobre p que suspeitamos ser verdadeira.

No nosso exemplo, o parâmetro é a probabilidade p de sair "cara".

Se consideramos 12 lançamentos independentes da moeda e denotamos por X o número de caras obtidas nesses lançamentos, então

$$X \sim \textbf{binomial}\ (12;\ p).$$

Note que o número de lançamentos está fixado (n=12), portanto fazer conjecturas sobre p é similar a fazer conjecturas sobre o número esperado de sucessos (esperança de X).

Se observarmos 5 caras em 12 lançamentos independentes da moeda, o que podemos concluir?

E se observarmos 4 caras? Ou 10 caras?

Podemos considerar uma **regra de decisão**, por exemplo,

"Se nos 12 lançamentos da moeda, observarmos 0,1, 2, 3, 9, 10, 11 ou 12 caras, então rejeitamos a hipótese nula H de que a moeda é honesta.

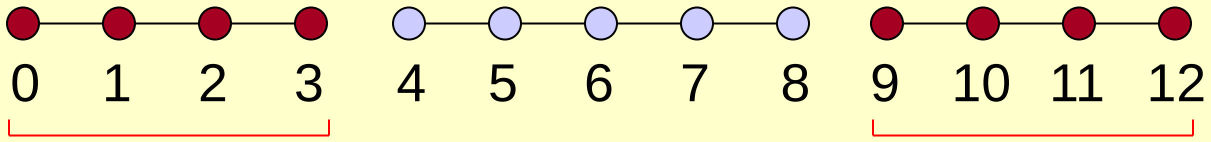
Caso contrário, não rejeitamos a hipótese H."

Testar uma hipótese estatística é estabelecer uma regra que nos permita, com base na informação de uma amostra, decidir pela rejeição ou não de H.

No exemplo, segundo a regra de decisão, o conjunto de valores de X que levam à rejeição da hipótese nula H é {0, 1, 2, 3, 9, 10, 11, 12}.

Denominamos esse conjunto **região crítica (RC)** ou **região de rejeição de H.**

RC = {0, 1, 2, 3, 9, 10, 11, 12} : **região de rejeição**

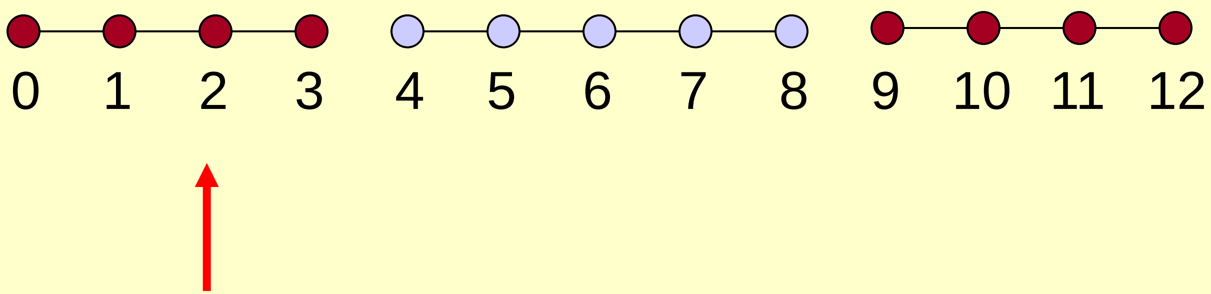


RC^c = {4, 5, 6, 7, 8} : **região de não rejeição de H**

Regra de decisão (teste)

Seja x o <u>valor observado</u> da variável aleatória X.

No exemplo, suponha que observamos 2 caras, isto é, $x = 2$.



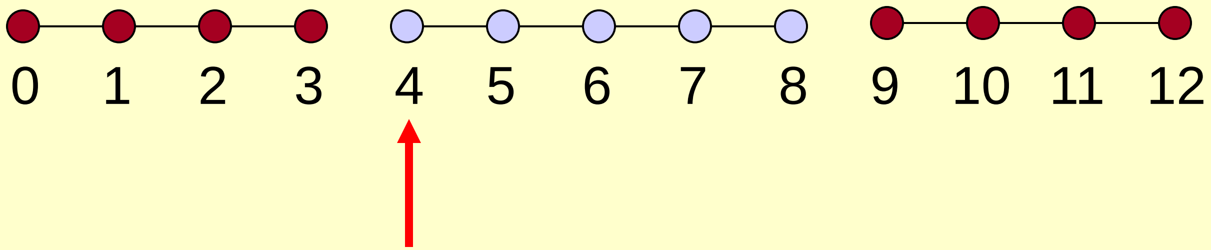
Valor observado na amostra

$x \in$ RC $\Rightarrow$ rejeitamos H.

Será que a nossa conclusão está correta?

Regra de decisão (teste)

Agora suponha que observamos 4 caras, isto é, x = 4.



Valor observado na amostra

Como $x \notin$ RC $\Rightarrow$ **não rejeitamos** H (não temos evidência suficiente de que a moeda seja desequilibrada).

Será que a nossa conclusão está correta?

Regra de decisão (teste)

$x \in$ **RC** $\Rightarrow$ **rejeitamos** H

$x \notin$ **RC** $\Rightarrow$ **não rejeitamos** H

Ao decidir pela rejeição ou não da hipótese nula H, podemos cometer *dois tipos de erro*.

Erros

Erro tipo I: Rejeitar H quando H é verdadeira.

(Afirmar que a moeda não é honesta quando na verdade ela é).

Erro tipo II: Não rejeitar H quando H é falsa.

(Afirmar que a moeda é honesta quando na verdade ela é desequilibrada).

Exemplo: Uma pessoa está sendo julgada.

Como pela lei uma pessoa é inocente até que se prove o contrário, as hipóteses são:

H: A pessoa é inocente.

A: A pessoa é culpada.

- Erro I: A pessoa é condenada apesar de ser inocente.
- Erro II: A pessoa é absolvida apesar de ser culpada.
- Naturalmente, a Justiça procura reduzir a possibilidade de ocorrer o Erro I, pois entende-se que é mais grave condenar inocentes do que absolver criminosos.

Probabilidades de erros

P(erro I) = P(rejeitar H **|** H é verdadeira**) =** α

α **: nível de significância do teste**

P(erro II) = P(não rejeitar H **|** H é falsa**) =** β

1 - β **: poder do teste**

Observações:

- α e β têm uma relação inversa.
- Em geral, só podemos controlar um dos erros (fixando sua probabilidade de ocorrência).

No exemplo da moeda:

H: $p = 0{,}5$
A: $p \neq 0{,}5$

X **~ binomial (**12; p**)**

RC = {0,1,2,3,9,10,11,12}

α = P(erro I) = P(rejeitar H | H verdadeira)

= $P(X \in \textbf{RC} \mid p = 0{,}5)$

= $P(X=0 \mid p=0{,}5) + \ldots + P(X=3 \mid p=0{,}5) + P(X=9 \mid p=0{,}5) +$

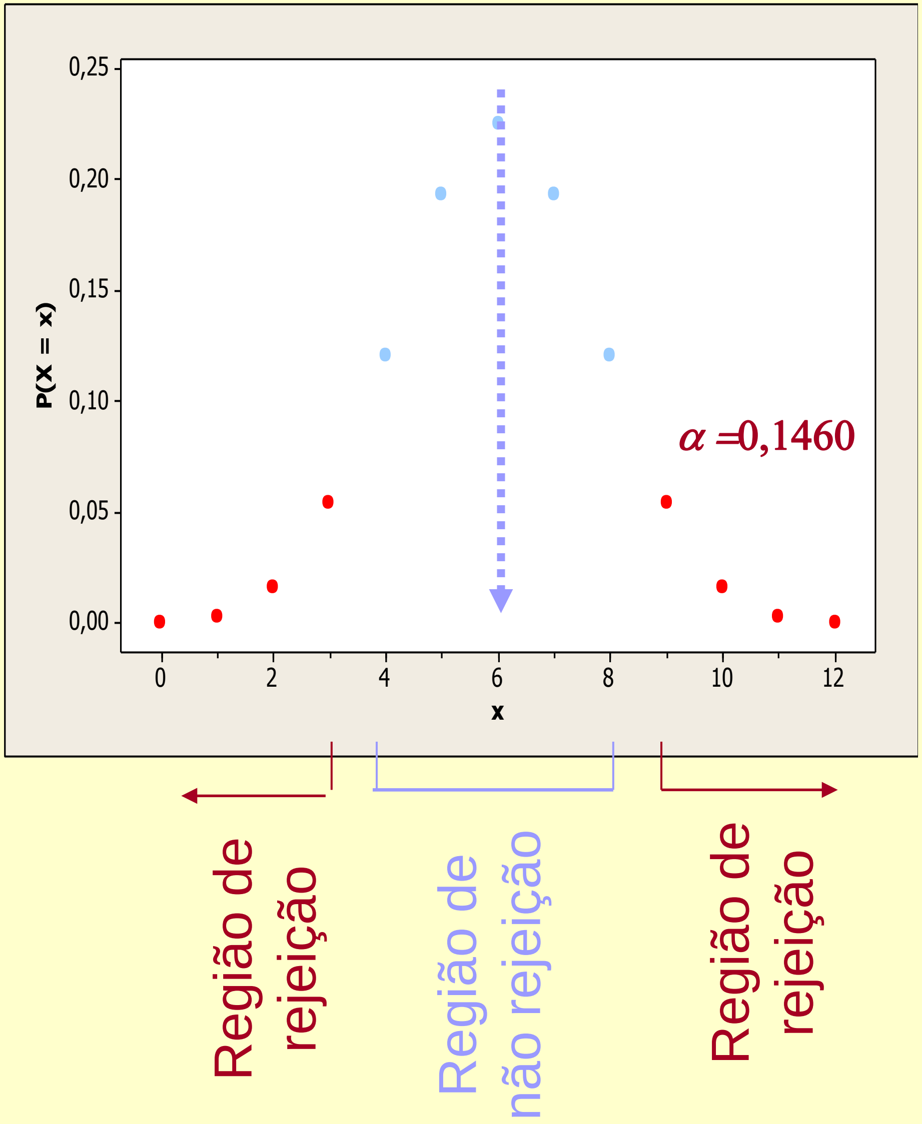
$\ldots + P(X=12 \mid p=0{,}5)$ $\Rightarrow$

= 0,000244 + 0,00293 + 0,016113 + 0,053711 + 0,053711 +

0,016113 + 0,00293 + 0,000244

= **0,1460**

Valor de $E(X)$ sob H



Decisão	Verdadeiro valor de p	
	p **= 0,5** (H é verd.)	$p \neq$ **0,5** (A é verd.)
Não rejeitar ***H***	Decisão correta **1 -** α = 0,8540	**Erro II** β
Rejeitar ***H***	**Erro I** α = 0,1460	Decisão correta **1 -** β

Se alterarmos a regra de decisão para **RC** = {0, 1, 2, 10, 11, 12}, isto é, concluiremos que a moeda é desonesta se o número de caras for 0, 1, 2, 10, 11 ou 12, o que acontece com o nível de significância α do teste (probabilidade de erro tipo I)?

α = P(erro I) = P(rejeitar *H* | *H* verdadeira) $=$ P($X \in$ **RC** | p = 0,5)

= P(X=0 | p=0,5) + ... + P(X=2 | p=0,5) + P(X=10 | p=0,5)+

... + P(X=12 | p=0,5) $\Rightarrow$

= 0,000244 + 0,00293 + 0,016113 + 0,016113 + 0,00293 +

0,000244

= **0,0384**

Regiões críticas e níveis de significância α (Exemplo 1: Moeda)

RC	α
{0, 1, 2, 3, 9, 10, 11, 12}	0,1460
{0, 1, 2, 10, 11, 12}	0,0384
{0, 1, 11, 12}	0,0063

RC $\downarrow$ $\quad \alpha \downarrow \quad \beta \uparrow$

Até agora, o procedimento foi

escolher RC $\Rightarrow$ determinar α.

Alternativamente, podemos

fixar α $\Rightarrow$ determinar RC.

Os valores de nível de significância α usualmente adotados estão entre 1% e 10%.

Determinação da região crítica

Exemplo 2: Suponha que um medicamento existente no mercado produza o efeito desejado em 60% dos casos nos quais é aplicado.

Um laboratório produz um novo medicamento e afirma que ele é melhor do que o existente.

Objetivo: Verificar estatisticamente se é verdadeira a afirmação do laboratório.

Aplicou-se o medicamento em n = 10 pacientes.

Seja X o número de pacientes, dentre os 10, para os quais o novo medicamento produz o efeito desejado.

Temos que:

$$X \sim b\,(10;\, p),$$

onde p é a proporção de pacientes para os quais o novo medicamento é eficaz.

(1) Hipóteses estatísticas:

$$H\colon p = 0{,}6$$
$$A\colon p > 0{,}6$$

que correspondem a

H: O novo medicamento é similar ao existente.

A: O novo medicamento é melhor (mais efetivo).

(2) Fixemos o nível de significância em 5% (α = 0,05).

(3) A região crítica deve ter a forma:

$$\mathbf{RC} = \{ X \geq k \}$$

O valor de k deve ser tal que

P(**erro I**) = P($X \in$ **RC** | p = 0,6) = P($X \geq k$ | p = **0,6**) = α

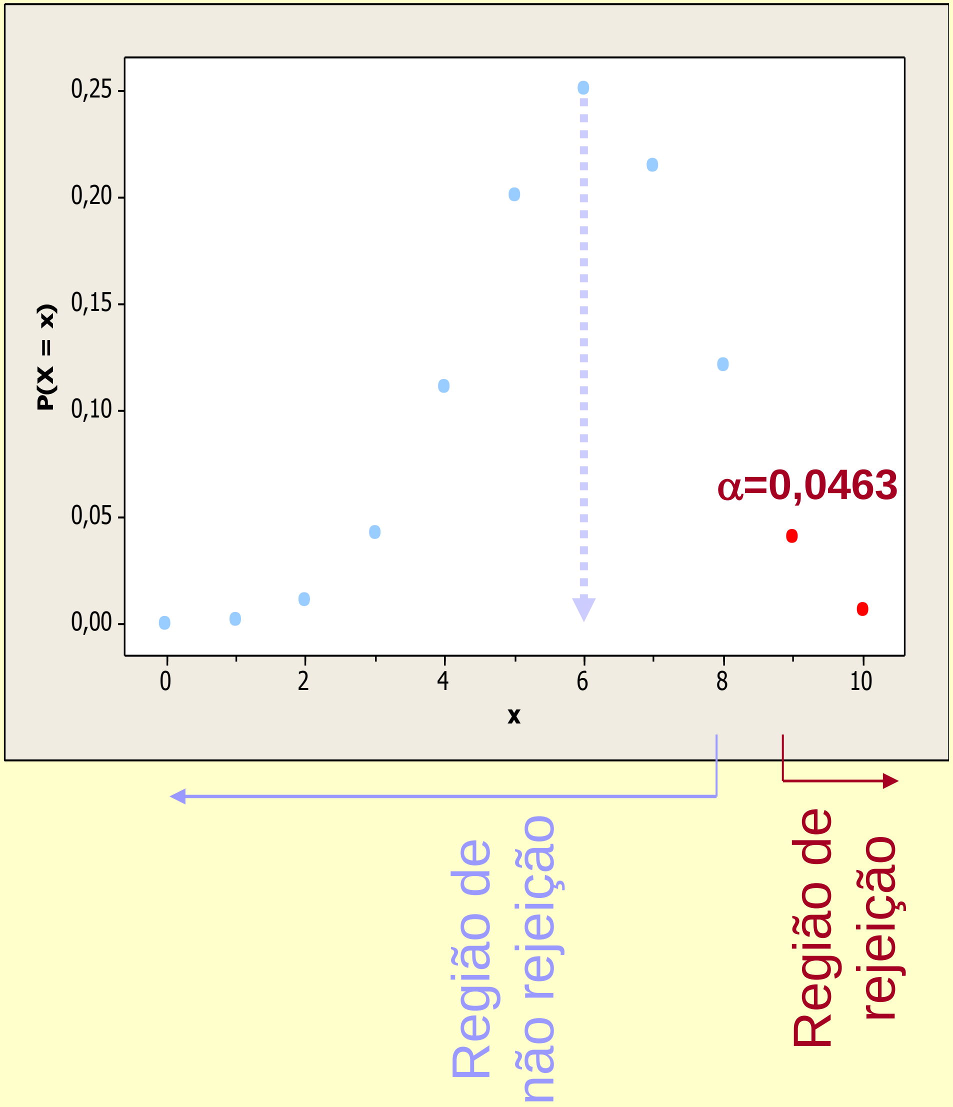
Pela tabela da **binomial** (10; 0,6), $\Rightarrow$

para k = 9: P($X \geq$ 9) = 0,0463

para k = 8: P($X \geq$ 8) = 0,1672

Portanto, **RC = {$X \geq$ 9},** que garante um nível de significância menor que 5% (na realidade, α **= 4,63%**).

Valor de $E(X)$ sob H

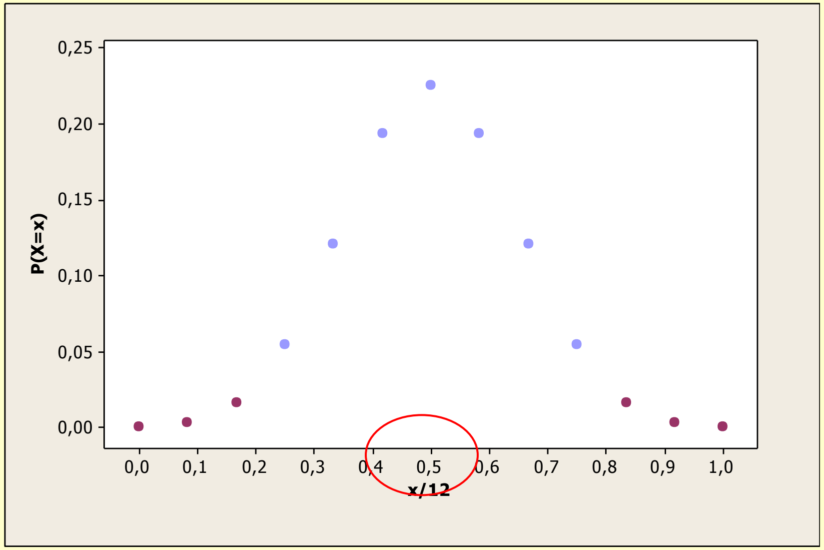


Hipóteses alternativas bilaterais e unilaterais

No Exemplo 1 (da moeda), as hipóteses são

$$H: p = \mathbf{0{,}5} \quad \text{e} \quad A: p \neq \mathbf{0{,}5}.$$

Dizemos que a hipótese alternativa é **bilateral** (queremos detectar desvios em torno de p = 0,5 em qualquer direção).



RC RC

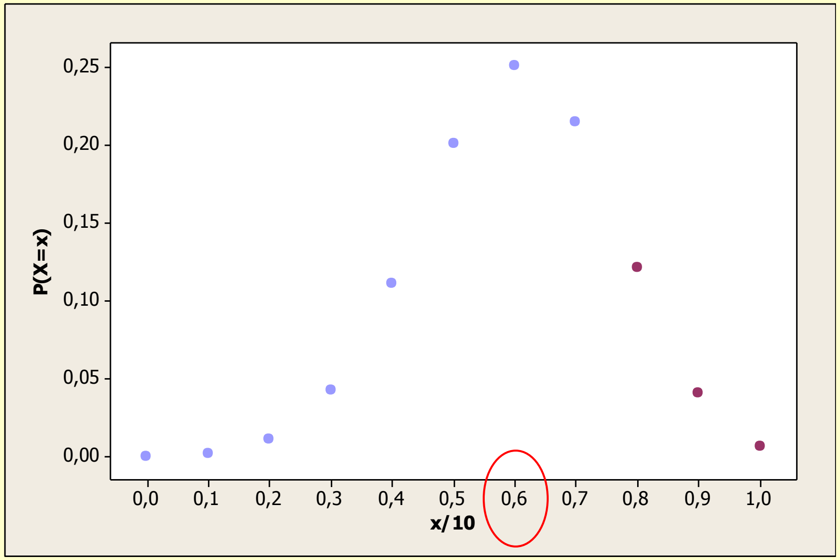
Hipóteses alternativas bilaterais e unilaterais

No Exemplo 2, as hipóteses são

$$H\text{: } p = \mathbf{0{,}6} \quad \text{e} \quad A\text{: } p > \mathbf{0{,}6}$$

isto é, desejamos detectar desvios em p em apenas uma direção (desvios à direita de 0,6).

Nesse caso, a hipótese alternativa é **unilateral**.



RC

Exemplo 3: A proporção de analfabetos em um município era de 15% na gestão anterior.

No início da sua gestão, o prefeito atual implantou um programa de alfabetização e após 2 anos afirma que reduziu a proporção de analfabetos.

Para verificar a afirmação do prefeito, n = 60 cidadãos foram entrevistados.

Seja X o número de analfabetos entre os 60 cidadãos entrevistados.

Então: $X \sim \mathbf{b}(60; p)$,

sendo p a proporção atual de analfabetos no município (após o programa de alfabetização).

(1) As **hipóteses de interesse** são:

H: A proporção de analfabetos no município não se alterou (a afirmação do prefeito está incorreta).

A: A proporção de analfabetos no município diminuiu (a afirmação do prefeito está correta).

Equivalentemente,

$$H: p = 0{,}15$$
$$A: p < 0{,}15$$

(2) Vamos fixar α **= 5%.**

(3) A **região crítica** deve ter a forma:

$$\mathbf{RC} = \{ X \leq k \}$$

O valor de k deve ser tal que P(erro I) = α, ou seja,

$$P(X \leq k \mid p = 0{,}15) = 0{,}05.$$

Pela tabela da **binomial**(60; 0,15), $\Rightarrow$

$$\boxed{\mathrm{RC} = \{X \leq 4\}.}$$

Na realidade, temos $\alpha = 0{,}0424$.

(4) Buscar a **evidência na amostra** para concluir.

Se observamos 6 analfabetos entre os 60 entrevistados, qual a conclusão?

(5) **Decisão e conclusão:**

$\mathbf{6 \notin RC} \Rightarrow$ Decidimos por não rejeitar $\boldsymbol{H}$ ao nível de significância 4,24%.

Concluímos que não temos evidência suficiente para afirmar que a proporção de analfabetos (após o programa de alfabetização) é inferior a 15%, isto é, não há evidência suficiente de que a afirmação do prefeito seja correta.

Resumo

(1) Estabelecer as **hipóteses**:

H: $p = p_0$ contra uma das alternativas

A: $p \neq p_0$, A: $p > p_0$ ou A: $p < p_0$.

(2) Escolher um **nível de significância α.**

(3) Determinar a **região crítica RC** da forma

$\{ X \leq k_1 \} \cup \{ X \geq k_2 \}$, $\{ X \geq k \}$ ou $\{ X \leq k \}$,

respectivamente às hipóteses alternativas.

(4) Selecionar uma **amostra aleatória** e determinar o número $\boldsymbol{x}$ de elementos na amostra com o atributo desejado.

(5) **Decidir**, usando a evidência $\boldsymbol{x}$, ao nível de significância α, e **concluir**.

$x \in$ RC $\Rightarrow$ rejeitamos H.

$x \notin$ RC $\Rightarrow$ não rejeitamos H.