

# Metodologia para tomada de decisões

- **Exemplo** Se nosso objetivo é estimar o valor desconhecido da proporção  $p$  de indivíduos que possuem um determinado atributo, tomamos como base o número de portadores  $X = x$  desse atributo em uma amostra casual simples de tamanho  $n$ . O objetivo é verificar se  $X = x$  dá suporte à rejeição de uma certa suposição para  $p$ .

# Metodologia para tomada de decisões

- **Exemplo 1:** Achamos que uma moeda não é honesta. Observamos 65 caras em 100 lançamentos, Podemos afirmar que os dados suportam nossa opinião?
- **Exemplo 2:** A diretoria de uma distribuidora de TV a cabo acredita que com a reengenharia de seu sistema de cabeamento, a proporção de assinantes satisfeitos com o serviço é agora maior do que os 82% existentes anteriormente. Se uma pesquisa com 300 assinantes revela que 251 deles estão satisfeitos com a distribuidora, a diretoria pode dizer que esses dados dão suporte à sua opinião?

# Metodologia

- **Exemplo 3:** Se em uma amostra casual simples de 2100 intervenções cirúrgicas feitas em 1997, apenas 181 são cirurgias cárdio-vasculares, é possível dizer que os dados suportam a conjectura (do centro nacional de estatísticas médicas) de que menos de 10% das cirurgias realizadas em 1997 foram cárdio-vasculares?

# Metodologia

- Se nosso objetivo é calcular ou estimar o valor desconhecido da proporção  $p$  de indivíduos em uma população que possui um determinado atributo, tomamos como variável de estudo a variável aleatória  $X$  que representa o número total de portadores do atributo, em uma amostra casual simples de tamanho  $n$  desses indivíduos.
- O objetivo é determinar se o valor  $x$  de  $X$ , observado na amostra, dá ou não suporte à uma hipótese referida ao valor de  $p$ .

# Metodologia

- Portanto, quando estamos interessados em estudar uma proporção  $p$ , baseamos nossa inferência em:
  - $X \sim \text{Bin}(n, p)$
  - $E(X) = np$
  - $\text{Var}(X) = np(1 - p)$
- E para valores adequados de  $n$  e  $p$  (ou  $n$  muito grande):
  - $\frac{X - np}{\sqrt{np(1-p)}} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0, 1)$

# Metodologia

- Relativo ao exemplo 1:
  - O que aconteceria em 100 lançamentos se a moeda fosse honesta ( $p = \frac{1}{2}$ )?
    - O número esperado de caras nos 100 lançamentos será 50
  - Comparado com essa expectativa, o que foi observado?
    - Observamos um desvio de  $|65 - 50| = 15$  unidades em relação ao número esperado de caras

# Metodologia

- Relativo ao exemplo 1:
  - Se a moeda for honesta, um desvio como o observado será pouco ou muito provável?
    - Se a moeda fosse honesta, teríamos que:
$$p\text{-valor} = P(|X - 50| \geq 15) \simeq P(|Z| \geq 3) \simeq 0.0027$$
    - Podemos usar  $p$  – valor para medir a magnitude da evidência contida nos dados contra a hipótese de honestidade da moeda.

# Metodologia

- Relativo ao exemplo 1:
  - Como decidir se a evidência  $p$  – valor é forte para rejeitar a suposição de honestidade da moeda?
    - Escolhemos um valor  $\alpha$  (nível de significância)
    - Se  $p$  – valor  $\leq \alpha$ , reconhecemos na amostra evidência suficiente para rejeitar a honestidade da moeda (rejeitar  $H_0$ ).
  - $H_0 : p = \frac{1}{2}$  vs  $H_1 : p \neq \frac{1}{2}$ 
    - para  $\alpha \geq 0.0027$ , rejeitamos  $H_0$  com nível de significância  $\alpha$



# Hipóteses Alternativas Unilaterais e Bilaterias

- **Exemplo 1:**  $H_0 : p = \frac{1}{2}$  vs  $H_1 : p \neq \frac{1}{2}$
- **Exemplo 2:**  $H_0 : p = 0.82$  vs  $H_1 : p > 0.82$
- **Exemplo 3:**  $H_0 : p = 0.1$  vs  $H_1 : p < 0.1$

**Hipótese Bilateral:** Quando desejamos detectar desvios em qualquer direção desde o valor da hipótese nula.

# Hipóteses Alternativas Unilaterais e Bilaterias

- Relativo ao exemplo 2:
  - $X$  = número de clientes satisfeitos
  - Sob  $H_0$  verdadeira:
    - $X \sim Bin(300, 0.82)$
    - $E(X) = 300 \times 0.82$
    - $Var(X) = 300 \times 0.82 \times 0.18$
  - $p$  - valor =  $P(X - 300 \times 0.82 \geq 251 - 300 \times 0.82) \simeq P(Z \geq 0.7514) \simeq 0.2266$  ● onde  $Z \sim N(0, 1)$
  - Logo, para  $\alpha \geq 0.2266$ , rejeitamos  $H_0$

**Observação:** os valores usuais de  $\alpha$  são 0.01, 0.05 e 0.1.

# Teste de Proporções

■  $H_0 : p = p_0$  vs  $H_1 : p \neq p_0$

$H_0 : p = p_0$  vs  $H_1 : p > p_0$

$H_0 : p = p_0$  vs  $H_1 : p < p_0$

■ Procedimento

- escolha  $\alpha$  (nível de significância) - limite máximo de probabilidade de cometer erro tipo I
- selecione uma amostra casual simples da população e determine o número  $x$  de indivíduos portadores do atributo
- determine o  $p$  – valor ou força de evidência contida nos dados
- se  $p$  – valor  $\leq \alpha$ , rejeita-se  $H_0$

## Cálculo do $p$ – valor

- $H_1 : p \neq p_0$

- $X \sim \text{Bin}(n, p)$

$$p\text{-valor} = P(|X - np_0| \geq |x - np_0|)$$

$$p\text{-valor} = 1 - P(np_0 - |x - np_0| < X < np_0 + |x - np_0|)$$

- usando TCL ( $Z \sim N(0, 1)$ )

$$p\text{-valor} = P(|X - np_0| \geq |x - np_0|)$$

$$p\text{-valor} \simeq P\left(|Z| \geq \frac{|x - np_0|}{\sqrt{np_0(1-p_0)}}\right)$$

$$p\text{-valor} \simeq 2 \left[ 1 - \Phi\left(\frac{|x - np_0|}{\sqrt{np_0(1-p_0)}}\right) \right]$$

# Cálculo do $p$ – valor – valor

- $H_1 : p > p_0$

- $X \sim \text{Bin}(n, p)$

- $p$  – valor =  $P(X \geq x)$

- usando TCL ( $Z \sim N(0, 1)$ )

- $p$  – valor =  $P(X - np_0 \geq x - np_0)$

- $p$  – valor  $\simeq P\left(Z \geq \frac{x - np_0}{\sqrt{np_0(1-p_0)}}\right)$

- $p$  – valor  $\simeq 1 - \Phi\left(\frac{x - np_0}{\sqrt{np_0(1-p_0)}}\right)$

## Cálculo do $p$ – valor

- $H_1 : p < p_0$

- $X \sim \text{Bin}(n, p)$

- $p$  – valor =  $P(X \leq x)$

- usando TCL ( $Z \sim N(0, 1)$ )

- $p$  – valor =  $P(X - np_0 \leq x - np_0)$

- $p$  – valor  $\simeq P\left(Z \leq \frac{x - np_0}{\sqrt{np_0(1-p_0)}}\right)$

- $p$  – valor  $\simeq \Phi\left(\frac{x - np_0}{\sqrt{np_0(1-p_0)}}\right)$

## Cálculo do $p$ – valor

- **Observação:** As hipóteses para o exemplo 3 são:

$$H_0 : p = 0.1 \quad \text{vs} \quad H_1 : p < 0.1$$

- Nesse tipo de situações, é suficiente calcular  $p$  – valor considerando o valor ( $p = 0.1$ ), por isso:

$$H_0 : p \geq 0.1 \quad \text{vs} \quad H_1 : p < 0.1$$

é equivalente a testar

$$H_0 : p = 0.1 \quad \text{vs} \quad H_1 : p < 0.1$$

- Comentário análogo vale quando as hipóteses forem

$$H_0 : p = 0.1 \quad \text{vs} \quad H_1 : p > 0.1.$$