

Testes de hipótese: Motivação

Prof. Paulo Justiniano Ribeiro Junior

Departamento de Estatística
Universidade Federal do Paraná



Falar sobre **população** a partir da observação da **amostra**.

Na **inferência estatística** os dois principais objetivos são:

1. **Estimar** um parâmetro populacional.
 - ▶ Estimativa pontual.
 - ▶ Estimativa intervalar.
2. **Testar** uma hipótese ou afirmativa sobre um parâmetro populacional.



Figura 1. Analogia ao processo de estimação.
Extraído de bestbowreviews.com.

Testes de hipótese

Hipótese

É uma afirmativa sobre uma **propriedade** da população.

Teste de hipótese

- ▶ É um procedimento para se testar uma **afirmativa** sobre uma propriedade da população.
- ▶ Permite tomar **decisões** sobre a população com base em informações de dados amostrais.

Um exemplo: proporção sexual em peixes

- ▶ Deseja-se estudar a proporção de peixes machos e fêmeas de uma mesma espécie em uma lagoa.
- ▶ Sem nenhuma informação prévia, supõe-se que a proporção sexual é de 50% ($p = 0.5$).
- ▶ Se, em uma amostra de 100 peixes:
 - ▶ 54 forem fêmeas.
 - ▶ 65 forem fêmeas.
 - ▶ 92 forem fêmeas.
- ▶ Qual a evidência necessária para concluir que a proporção de fêmeas é maior que a de machos nessa população?

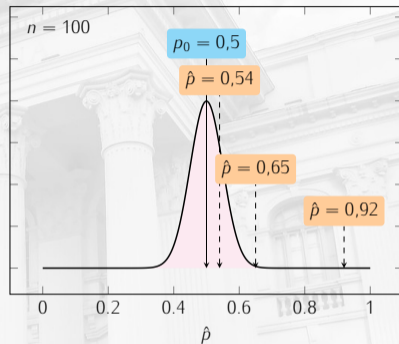


Figura 2. Proporções amostrais supondo $p = 0.5$ na população.

Um exemplo: cardápio vegano

- ▶ Um restaurante deseja caracterizar o perfil de seus clientes.
- ▶ Questionário para uma *amostra* de clientes.
- ▶ **Q1:** Há interesse por opções veganas? ($\hat{p} = 0.12$ em uma amostra).



Figura 3. Foto de Pexels.

Intervalo de confiança

- ▶ **Pergunta:** Qual a proporção que prefere pratos veganos?
- ▶ 0.12 ± 0.035 ou $(0.085, 0.155)$

Teste de hipótese

- ▶ **Pergunta:** A proporção de clientes que prefere pratos veganos supera 10%?
- ▶ $\hat{p} = 0.12$ é **significativamente maior** do que $p = 0.10$?

Um exemplo: caracterização dos clientes

Q2: Qual será a *idade média* dos clientes?
 ($\bar{y} = 32$ em uma amostra)



Figura 4. Foto de Adrienn no Pexels.

Intervalo de confiança

- ▶ **Pergunta:** Qual a idade média dos clientes?
- ▶ 32 ± 2.5 ou $(29.5, 34.5)$

Teste de hipótese

- ▶ **Pergunta:** A idade média dos clientes é igual a 35 anos?
- ▶ $\bar{y} = 32$ é **significativamente diferente** de $\mu = 35$?

Um exemplo: caracterização dos clientes por sexo

Q3: Qual será a *idade média* dos clientes **por sexo**?

($\bar{y}_h = 34$ para homens e $\bar{y}_m = 31$ para mulheres em uma amostra)

Intervalo de confiança

- ▶ **Pergunta:** Qual a idade média dos clientes homens e mulheres?
- ▶ 34 ± 2.3 ou (31.7, 36.3) para homens
- ▶ 31 ± 2.8 ou (28.2, 33.8) para mulheres
- ▶ ou ainda a diferença de idade
- ▶ 3 ± 2.5 ou (0.5, 5.5)

Teste de hipótese

- ▶ **Pergunta:** Existe diferença (**significativa**) entre a idade média dos clientes homens e mulheres?
- ▶ $\bar{y}_h - \bar{y}_m = 34 - 31 = 3$ é **significativamente diferente** de $\mu_h - \mu_m = 0$?

Fundamento lógico do teste de hipótese

Testamos uma afirmativa na tentativa de distinguir entre resultados que:

- ▶ Podem **facilmente** ocorrer por *acaso* na amostra.
- ▶ São **altamente improváveis** de ocorrer por *acaso* na amostra.

A ocorrência de **resultados altamente improváveis** pode ser explicada de uma das duas formas:

- ▶ Ou um evento raro realmente **ocorreu**.
- ▶ Ou a **suposição** subjacente não é verdadeira.

Regra do evento raro

Se, sob uma dada **suposição**, a probabilidade de um evento observado particular é **extremamente pequena**, concluímos que a suposição **provavelmente** não é verdadeira.