DCP098

Fundamentos para Avaliação Quantitativa de Políticas Públicas

Teste de hipótese

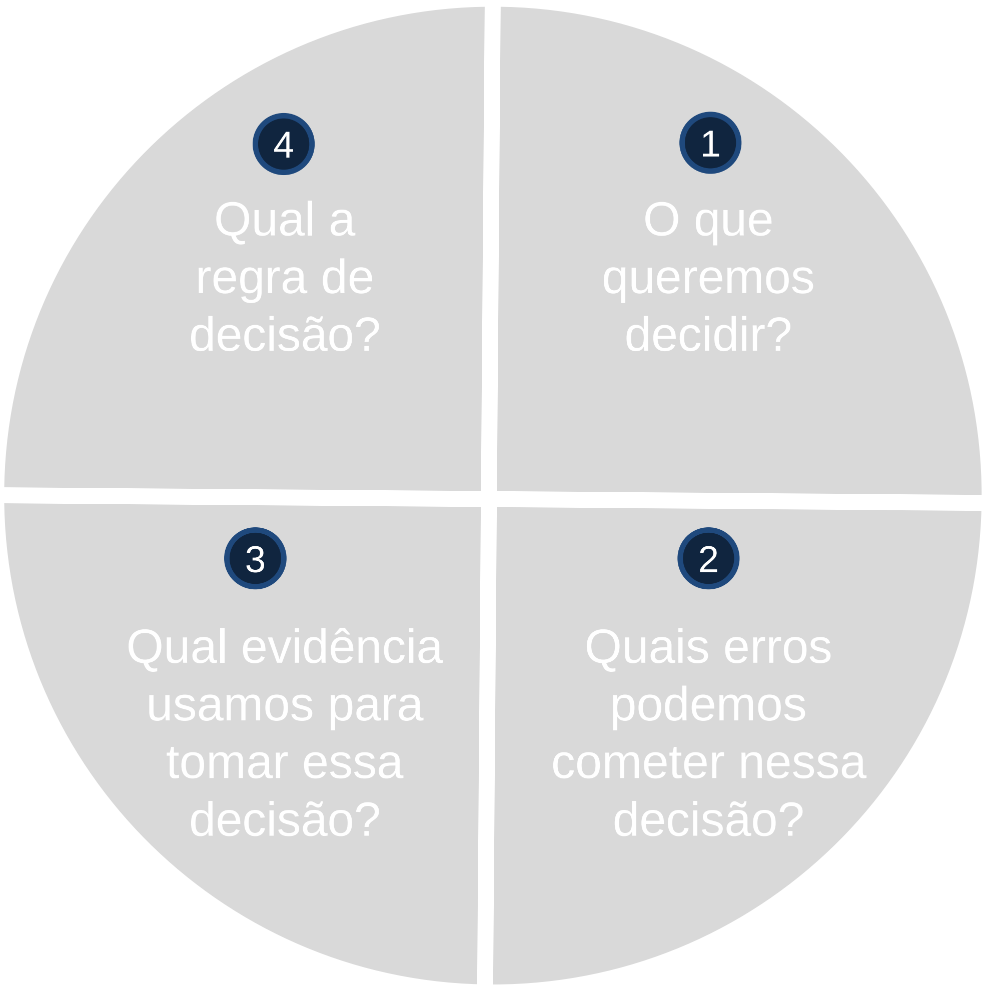
Aula 13

16 de maio de 2022

Ana Paula Karruz

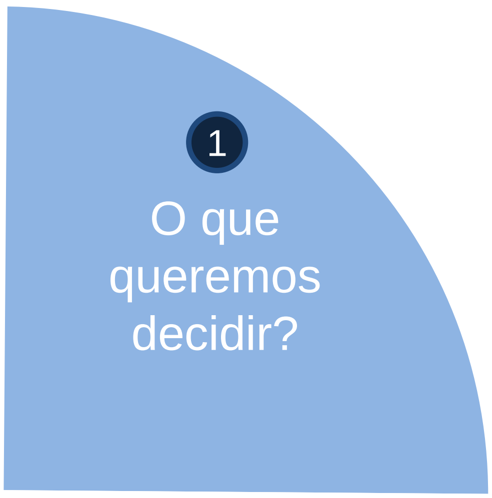
Conclusões sobre efeito de X em Y requerem tomada de decisão sobre o significado da evidência apurada

Teste de hipótese é rotina para essa tomada de decisão



Conclusões sobre efeito de X em Y requerem tomada de decisão sobre o significado da evidência apurada

Teste de hipótese é rotina para essa tomada de decisão



Como nos convencemos de que uma associação é real

(e não apenas fruto da aleatoriedade)

- **Testes de hipótese permitem avaliar se os dados observados são compatíveis com uma proposição** (e.g., donuts não afetam peso)

- Resultado de teste de hipótese **não é determinístico** (ou seja, resultado não é do tipo "donuts engordam" ou "donuts não engordam")

- Ao contrário, resultado **expressa um grau de confiança** em relação à aderência entre a afirmação contida na hipótese e a evidência coletada (e.g., "com base nesta amostra, estamos 95% confiantes de que a afirmação 'donuts não afetam peso' é falsa")

- No contexto de uma regressão, testes de hipótese informam, por exemplo, a **probabilidade de se ter obtido o $\beta_1$hat que obtivemos (dada sua imprecisão)**, em certo cenário

  – Cenário: previsto na hipótese testada (hipótese nula): H_0: $\beta_1 = 0$

  – Se a **probabilidade** de observar o **$\beta_1$hat (com respectivo grau de imprecisão)** que observamos for bem **baixa** dado o cenário de H_0, então **desconfiamos do cenário declarado em H_0**

Hipóteses testadas são **sobre parâmetros populacionais (βs)**, não sobre estimativas calculadas a partir da amostra (βhats)

A relação entre X e Y é "estatisticamente significante"?

Em outras palavras, a evidência suporta o cenário em que $\beta_1 = 0$?

As ferramentas estatísticas ***não*** *nos permitem* ***provar [...] uma hipótese nula****. Em vez disso,* ***"rejeitamos"*** *ou* ***"deixamos de rejeitar"*** *as hipóteses nulas.*

Bailey (2016: 138)

Quando ***deixamos de rejeitar*** *uma hipótese nula, estamos dizendo que o* ***$\beta_1$hat*** *que observamos* ***não*** *seria particularmente* ***improvável*** *se a hipótese nula fosse verdadeira. Por exemplo, normalmente [não] rejeitamos a hipótese nula quando observamos um* ***pequeno $\beta_1$hat****. Esse resultado não seria nada surpreendente se $\beta_1 = 0$. Também podemos deixar de rejeitar hipóteses nulas quando a* ***incerteza é alta****. Ou seja, um* ***grande $\beta_1$hat pode não ser muito surpreendente*** *mesmo quando $\beta_1 = 0$* ***se a variância de $\beta_1$hat for grande em relação ao valor de $\beta_1$hat****.*

Bailey (2016: 139)

Se, no entanto, observarmos um grande $\beta_1$hat com um pequeno erro padrão, ***rejeitaremos*** *a hipótese nula e nos referiremos ao* ***coeficiente como estatisticamente significativo****.*

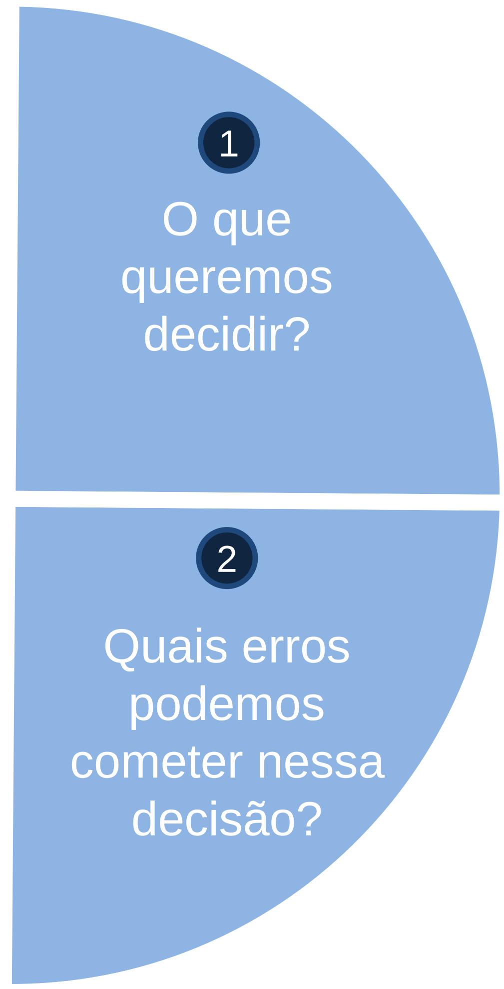
Bailey (2016: 138)

Quando rejeitamos uma hipótese nula, na verdade estamos dizendo que a probabilidade de se obter o $\beta_1$hat que estimamos é muito baixa se a hipótese nula fosse verdadeira.

Bailey (2016: 138)

Conclusões sobre efeito de X em Y requerem tomada de decisão sobre o significado da evidência apurada

Teste de hipótese é rotina para essa tomada de decisão



Nula (H_0) vs. Alternativa (H_1 ou H_a)

- **H_0: contém o cenário sendo testado**, e.g.:
  - H_0: $\beta_1 = 0$

- **H_1: contém todos os cenários concebíveis alternativos a H_0**, e.g.:
  - Hipótese alternativa bilateral (aka bicaudal):
    H_1: $\beta_1 \neq 0$
  - Hipótese alternativa direcional (aka unilateral ou unicaudal):
    H_1: $\beta_1 > 0$;
    outro exemplo: H_1: $\beta_1 < 0$

*Se **rejeitarmos a hipótese nula**, **aceitamos a hipótese alternativa**. **Não provamos** que a hipótese alternativa seja verdadeira. Em vez disso, a hipótese alternativa é a **ideia à qual nos apegamos** quando temos **evidência inconsistente** com a **hipótese nula**.*

Bailey (2016: 140)

O reconhecimento de que podemos incorrer em erro é central para o teste de hipótese

- Quando **rejeitamos H_0**, concluímos que o **cenário** nela contido (e.g., $\beta_1 = 0$) é **improvável dados o $\beta_1$hat** observado **e o se($\beta_1$hat)**. Nota: nossa conclusão não é que H_0 seja impossível

- Analogamente, quando **não rejeitamos H_0**, concluímos que o **cenário** nela contido (e.g., $\beta_1 = 0$) é **provável dados o $\beta_1$hat observado e o se($\beta_1$hat)**. Nota: nossa conclusão não é que H_0 seja verdadeira

- **Erros na conclusão de testes** de hipótese podem assumir 2 formas:

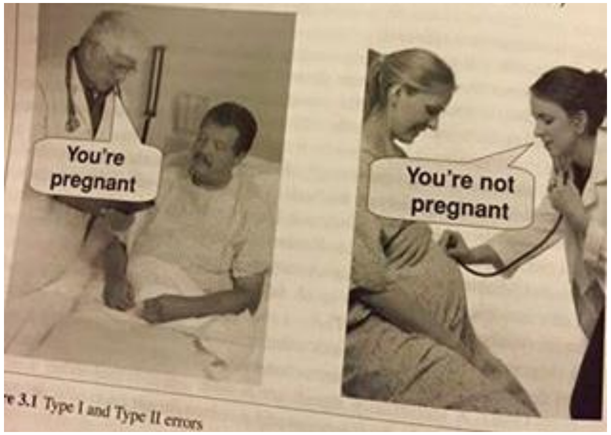
  – **Erro Tipo I:** Rejeitamos H_0 quando H_0 é verdadeira (e.g., saímos acreditando que X explica Y quando na verdade não explica); metáfora: julgar réu culpado quando ele não cometeu o crime

  – **Erro Tipo II:** Não rejeitamos H_0 quando H_0 é falsa (e.g., saímos acreditando que X não explica Y quando na verdade explica); metáfora: julgar réu inocente quando ele cometeu o crime

Erro Tipo II tende a ocorrer quando a amostra é pequena e/ou quando o nível de confiança (100-nível de significância) pré-definido para o teste é alto

Ilustração: erros Tipo I e Tipo II no teste de gravidez

H_0: gravidez = 0



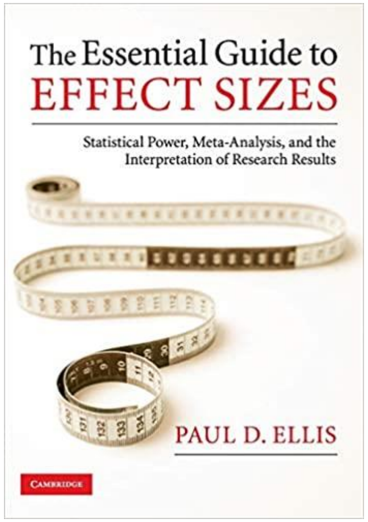


Tabela ou matriz de confusão (realmente, é assim que se chama!)

	Estado verdadeiro da natureza	
Decisão	**A hipótese nula é verdadeira**	**A hipótese nula é falsa**
Rejeitar a hipótese nula	**Erro tipo I** (rejeitar uma hipótese nula verdadeira); α	Decisão correta
Deixar de rejeitar a hipótese nula	Decisão correta	**Erro tipo II** (deixar de rejeitar uma hipótese nula falsa); β

Não é o mesmo β da equação de regressão

Atenção: Teste de hipótese não “prova” coisa alguma; incerteza não permitiria isso

(mesmo rejeitando H_0, incorremos em risco de erro Tipo I)

DECISÃO sobre hipótese nula	Implicação sobre hipótese alternativa
Rejeita	Aceita (mas não prova)
Não rejeita	**Não** rejeita

- Só mudamos nosso entendimento atual sobre proposição testada (i.e., H_0) se pudermos falseá-la; teste não nos permite concluir que H_0 seja verdadeira
  - Encontrar evidência de que X afeta Y na amostra pode nos autorizar a concluir que X afeta Y na população, com um certo grau de confiança
  - Não encontrar evidência de que X afete Y na amostra não nos autoriza a concluir que X não afeta Y na população (lembremos: a falta de evidência em favor de efeito pode advir da ausência de efeito na população ou da nossa incapacidade de detectá-lo)

Testes de hipótese focam no erro Tipo I, especificando a priori um nível aceitável (α) para esse erro

Quão improvável $\beta_j$hat (dada sua imprecisão) tem que ser para que rejeitemos H_0?

- O **nível de significância (α)** do teste, determinado a priori, especifica quão improvável $\beta_j$hat (com seu respectivo nível de imprecisão) tem que ser no cenário da H_0 para que rejeitemos essa hipótese

- Se $\alpha = 0{,}05$, então rejeitamos H_0 se observarmos um $\beta_j$hat de magnitude tão grande (em relação ao erro padrão) que esperaríamos um $\beta_j$hat desse tamanho ou maior em no máximo 5% das amostras, se a nula for verdadeira

  – **α = probabilidade de rejeitar H_0 quando H_0 é verdadeira**
  – **α = risco assumido de cometer erro Tipo I**
  – **α = 100% – nível de confiança**

Atenção: à medida que diminuímos α, aumentamos a probabilidade de cometer erro Tipo II

Níveis convencionais de significância (e confiança)

Nível de significância (α)	Nível de confiança
0,100 (10%)	90,0%
0,050 (5%)	95,0%
0,010 (1%)	99,0%
0,001 (0,1%)	99,9%

α = 0,10 (10%) é considerado nível de significância "marginal"

DCP098
Fundamentos para Avaliação Quantitativa de Políticas Públicas

Teste de hipótese

Aula 13
16 de maio de 2022

Ana Paula Karruz