

# Laboratório de Avaliação de Projetos

Aula 6 - Como avaliar?

Resultados potenciais

---

Felipe Nunes, Ph.D. (felipenunes@ufmg.br)

May 30, 2022

UFMG - 8/2020

# **Arcabouço Metodológico Formal de Resultados Potenciais**

---

## Resultados Potenciais

- Estamos interessados em avaliar o impacto de um programa sobre uma variável de resultado.
- Denotando por  $Y$  essa variável de resultado, podemos pensar que seu valor para indivíduo  $i$  será  $Y_i(1)$  caso ele esteja na situação de tratado, mas poderia assumir, pelo menos potencialmente, o valor  $Y_i(0)$  caso ele não seja tratado.
- Se fosse possível observar o indivíduo  $i$  nas duas situações, a diferença

$$\beta_i = Y_i(1) - Y_i(0)$$

forneceria o impacto do programa para esse indivíduo.

## Resultados Potenciais

- Cada indivíduo  $i$  possui um par de resultados potenciais  $[Y_i(1), Y_i(0)]$  e características observáveis pelo avaliador, que denotaremos por meio do vetor  $X_i$ .
- Exemplos típicos de características observadas que são incluídas em avaliações são o gênero, a idade e a escolaridade da pessoa.
- Uma maneira amplamente utilizada para especificar o modelo de geração de valores dos resultados potenciais é o seguinte:

$$Y_i(1) = X_i\alpha + \beta_i + \epsilon_i$$

$$Y_i(0) = X_i\alpha + \epsilon_i$$

## Resultados Potenciais

- Vamos assumir  $T_i$  uma variável discreta que assume valor 1 se o indivíduo participa do programa e valor 0 caso ele não participe.
- Então podemos expressar a equação de resultados potenciais, como:

$$\begin{aligned} Y_i &= T_i Y_i(1) + (1 - T_i) Y_i(0) \\ &= Y_i(0) + T_i(Y_i(1) - Y_i(0)) \end{aligned}$$

## Resultados Potenciais

- Aplicando as equações para  $Y_i(1)$  e  $Y_i(0)$ , na equação dos resultados potenciais, temos:

$$\begin{aligned} Y_i &= T_i Y_i(1) + (1 - T_i) Y_i(0) \\ &= Y_i(0) + T_i (Y_i(1) - Y_i(0)) \\ &= X_i \alpha + \epsilon_i + T_i (X_i \alpha + \beta_i + \epsilon_i - X_i \alpha - \epsilon_i) \\ &= X_i \alpha + \beta_i T_i + \epsilon_i \end{aligned}$$

- Ou seja, uma equação de regressão linear cujo  $\beta$ , valor que capta o efeito causal de  $T_i$ , varia por indivíduo!

## Resultados Potenciais

- O maior problema para estimar corretamente o impacto do programa por meio dessa equação é que a variável  $T_i$  é potencialmente correlacionada com o componente não observável  $\epsilon_j$ .
- Na realidade, é provável que isso ocorra, já que a participação no programa normalmente não se dá de forma aleatória.
- Ao contrário, a participação geralmente decorre de uma decisão voluntária do indivíduo (ou da família), a qual pode ter sido tomada com base em informações que não são observadas pelo avaliador e, portanto, estão embutidas no componente  $\epsilon_j$  da equação.
- **Exemplo:** indivíduos que participam de programas de treinamento são considerados mais motivados.

## Resultados Potenciais

- Esse problema de estimação incorreta de  $\beta$  devido à correlação entre a participação no programa ( $T$ ) e a parte não observável ( $\epsilon$ ) recebe o nome de **problema de autoseleção**.
- A magnitude dessa incorreção é chamada de **viés de autoseleção**.
- Sua denominação decorre do fato de que são os próprios tratados que se selecionam para participar do programa.
- Se essa autoseleção estiver associada com características não observáveis dos indivíduos, esse grupo se torna um grupo “especial” da população (por exemplo, mais motivados), o que pode afetar diversas variáveis de resultado, tornando difícil o isolamento do efeito causal do programa.



## Resultados Potenciais

- Boa parte do esforço de pesquisa realizado até os dias de hoje na área de avaliação tem sido dedicado a encontrar soluções para contornar o **problema da autosseleção**.
- A fim de apresentá-lo de forma mais geral, considere as seguintes médias populacionais da variável de resultado para os grupos de tratados e não tratados:

$$D_{11} = E[Y_i(1) | T_i = 1]$$

$$D_{10} = E[Y_i(0) | T_i = 1]$$

$$D_{01} = E[Y_i(1) | T_i = 0]$$

$$D_{00} = E[Y_i(0) | T_i = 0]$$

- Note que somente  $D_{11}$  e  $D_{00}$  são médias que podem ser calculadas diretamente dos dados,  $D_{10}$  e  $D_{01}$  representam contrafactuais.

## Resultados Potenciais

- Em geral, na área de avaliação, estamos interessados em computar o que se costuma denominar de efeito médio do programa sobre os tratados (TOT):

$$TOT = D = E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 1] = D_{11} - D_{10}$$

a diferença entre a média da variável de resultado para os tratados na situação de tratados e a média para esses mesmos indivíduos na situação contrafactual de eles não terem passado pela intervenção.

- O TOT é um parâmetro importante, uma vez que informa o impacto médio do programa para aquela parcela da população que foi efetivamente tratada.

## Resultados Potenciais

- É comum o uso de um grupo de indivíduos não tratados como o grupo que representa o contrafactual desejado. Quando isso ocorre, geralmente calcula-se a diferença:

$$ITT = D = E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 0] = D_{11} - D_{00}$$

que é tomada como o impacto médio do programa.

- É preciso cuidado para usar essa expressão, já que o grupo de não tratados não necessariamente funciona como bom contrafactual.

**Perguntas? Dúvidas?**  
**Questionamentos?**

---

## **Próxima aula: Métodos de Seleção Aleatória**

---