

Avaliação de Políticas Públicas B

Aula 10- Como avaliar?

Métodos de Pareamento - Propensity Scores

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

June 13, 2022

UFMG - 2/2022

Avaliação de Impacto com Dados Observacionais

Avaliação de Impacto com Dados Observacionais

- **Objetivo:** avaliar o efeito de uma política ou de um programa em um contexto em que é impossível realizar randomização.
- **Métodos disponíveis:**
 1. Pareamento (matching, weighting e propensity score)
 2. Variável instrumental
 3. Regressão descontínua
 4. Diferença em diferença
 5. Modelos de painel

Avaliação de Impacto com Dados Observacionais

- Por que a comparação de média em dados observacionais não permite a inferência de efeitos causais?

Grupos	Canada	U.K.	U.S.
Não fumantes	20.2	11.3	13.5
Fumantes de cigarro	20.5	14.1	13.5
Fumantes de chaturro	35.5	20.7	17.4

- Se aplicarmos o modelos de resultados potenciais para Canada:

$$\alpha = \bar{Y}_{Cigarro} - \bar{Y}_{NF} = 20.5 - 20.2 = 0.3$$

$$\alpha = \bar{Y}_{Charuto} - \bar{Y}_{NF} = 35.5 - 20.2 = 15.3$$

Avaliação de Impacto com Dados Observacionais

- Qual pressuposto que nós discutimos está sendo violado?
 1. É possível assumir que não há outra variável afetando o resultado?
 2. Qual seria a melhor forma de averiguar se há ou não outra variável interferindo na relação entre W (fumar) e Y (tx de mortalidade)?
 3. Quais tipos de viés nós podemos imaginar nesta situação?
 4. Esta comparação isola todos os outros possíveis efeitos observáveis?
- A análise de dados observacionais fere o pressuposto da aleatoriedade de variáveis omitidas:

$$\{Y_1, Y_0\} \perp W$$

Avaliação de Impacto com Dados Observacionais

X_3	Mortalidade Fumantes	Mortalidade Não-fumantes	Número Fumantes	Total pacientes
Velhos	28	24	3	10
Jovens	22	16	7	10
Total			10	20

- Se desconsiderarmos a terceira variável IDADE, qual seria o efeito do cigarro sobre a mortalidade?

$$[28(3/10) + 22(7/10)] - [24(7/10) + 16(3/10)] = 2.2$$

- Mas se considerarmos idade, qual o efeito?

$$(28 - 24) \cdot \frac{10}{20} + (22 - 16) \cdot \frac{10}{20} = 5$$

Avaliação de Impacto com Dados Observacionais

- Por que análise de regressão usando dados observacionais não permite a inferência de efeitos causais?
 1. Modelo básico de regressão:

$$Y_i = \alpha + \beta_1 X_i + \epsilon$$

2. Como estimar os parâmetros α e β_1 ?

$$\hat{\beta}_1 = \frac{n \sum (XY) - \sum X \sum Y}{n \sum X^2 - (\sum X)^2}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}_1 \bar{X}$$

3. Como interpretar os parâmetros $\hat{\alpha}$ e $\hat{\beta}_1$?

$\hat{\alpha}$ = Média de Y quando X igual a zero

$\hat{\beta}_1$ = correlação entre X e Y

Avaliação de Impacto com Dados Observacionais

- Diante dessas dificuldades, nós precisamos nos ancorar em outro pressuposto: **ignorabilidade condicional**

$$\{Y_1, Y_0\} \perp W | \{X_2, X_3, X_4, \dots\}$$

- Que tal a seguinte solução então?

$$Y_i = \alpha + \beta_1 W + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \epsilon$$

- Se o dado é observacional, não será uma modelagem de regressão que será suficiente para solucionar o problema da randomização.
- Ou encontramos uma variável que transforma nossa variável tratamento em aleatória ou construímos um banco de dados aleatório e válido!

Métodos de Pareamento

- Os métodos de pareamento utilizam, essencialmente, técnicas estatísticas para construir um grupo de comparação artificial ao identificar, para cada observação possível de tratamento, uma observação de não tratamento (ou um conjunto de observações de não tratamento) que possua a maior quantidade possível de características semelhantes.

Métodos de Pareamento

- Considere um caso em que você esteja tentando avaliar o impacto de um programa e exista um conjunto de dados que contenha tanto famílias inscritas no programa quanto as não inscritas.
- O programa que você está tentando avaliar não possui regras claras de alocação (tais como uma seleção aleatória ou um índice de elegibilidade), que expliquem por que algumas famílias se inscreveram no programa e outras, não.
- Em tal contexto, os métodos de pareamento permitirão que você identifique o conjunto de famílias não inscritas que mais se assemelham às famílias de tratamento, com base nas características disponíveis em sua base de dados.
- Essas famílias não inscritas “pareadas” serão o grupo de comparação que você usará para estimar o cenário contrafactual.

Métodos de Pareamento

Unidades tratadas			
Idade	Sexo	Meses desempregado	Diploma Secundário
9	1	3	0
35	1	12	1
41	0	17	1
23	1	6	0
55	0	21	1
27	0	4	1
24	1	8	1
46	0	3	0
33	0	12	1
40	1	2	0

Unidades tratadas			
Idade	Sexo	Meses desempregado	Diploma Secundário
24	1	8	1
38	0	2	0
58	1	7	1
21	0	2	1
34	1	20	0
41	0	17	1
46	0	9	0
41	0	11	1
19	1	3	0
27	0	4	0

Fonte: Autores, resultado de várias fontes.

Métodos de Pareamento

- Parece fácil, mas não é!
- Se a lista de características relevantes observadas for muito grande, ou se cada característica assume muitos valores, poderá ser difícil identificar um par correspondente para cada uma das unidades do grupo de tratamento.
- À medida que você aumenta o número de características ou dimensões com as quais quer parear as unidades que se inscreveram no programa, você pode incorrer na chamada “maldição da dimensionalidade”.

Propensity Score

Propensity Score

- Felizmente, a “maldição da dimensionalidade” pode ser facilmente resolvida através do “pareamento por escore de propensão” (Rosenbaum e Rubin, 1983).
- Nessa abordagem, não será mais necessário fazer o pareamento com o exato valor para todas as características de controle observadas.
- Ao invés disso, para cada unidade no grupo de tratamento e no conjunto de não inscritos, calcula-se a probabilidade de uma unidade vir a se inscrever no programa, com base nos valores observados das suas características.
- Esse escore é um único número, variando de 0 a 1, que resume todas as características observadas das unidades que influenciam a probabilidade de inscrição no programa.

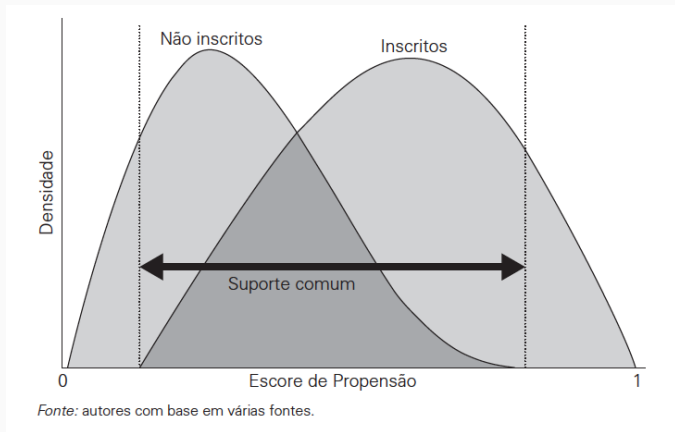
Propensity Score

- Uma vez que o escore de propensão tenha sido calculado para todas as unidades, poderá então ser feito o pareamento entre as unidades no grupo de tratamento e as unidades no conjunto de não inscritos que tenham escores de propensão o mais próximos possível.
- Essas “unidades próximas” serão o grupo de comparação e são usadas para produzir uma estimativa do contrafactual.
- A diferença nos resultados (Y) entre as unidades de tratamento ou inscritas e suas unidades de comparação correspondentes produz o impacto estimado do programa.

Propensity Score

- Para que o pareamento por escore de propensão gere estimativas válidas do impacto de um programa, todas as unidades de tratamento precisam ser pareadas satisfatoriamente com uma unidade de não inscritos.
- Pode acontecer que, para algumas unidades inscritas, nenhuma unidade no conjunto de não inscritos tenha escores de propensão semelhantes: **falta de suporte comum**
- Nesse caso, ausência de sobreposição entre os escores de propensão do grupo de tratamento ou grupo de inscritos e os do conjunto dos não inscritos não permite o uso do método.

Propensity Score



- Limitações do método de 'propensity score':
 1. Exige um conjunto extenso de dados e grandes amostras que, mesmo quando disponíveis, não impedem a falta de suporte comum entre o grupo de tratamento - ou de inscritos - e o conjunto de não participantes.
 2. O pareamento somente pode ser executado com base em características observadas. A suposição de que não houve viés de seleção oriundo de características não observadas é muito forte e, o que é mais problemático ainda, não pode ser testado.

**Perguntas? Dúvidas?
Questionamentos?**

Próxima aula: Propensity Score no R!
