

Avaliação de Políticas Públicas usando o R

Introdução à Avaliação de Impacto

Vítor Wilher

Mestre em Economia | Cientista de Dados



Plano de Voo

Pacotes

Introdução

Definição do problema

Causalidade

Validação interna X Validação externa

Heterogeneidade de impacto

Estratégias de identificação

Pacotes

Os pacotes que usaremos nessa seção estão listados abaixo. Sempre que possível, colocaremos os pacotes utilizados logo no início de cada seção.¹

```
library(tidyverse)
library(truncnorm)
library(broom)
library(kableExtra)
```

¹É possível que algum dos pacotes listados tenha sido retirado sem prévio aviso do CRAN. Nesses casos, por favor, entre em contato com o nosso suporte.

Introdução

Política pública refere-se às ações e normas formuladas pelos governos com o objetivo de solucionar problemas que impactam a sociedade. Assim, os propósitos podem ser os mais diversos, como, por exemplo, aumentar a renda, melhorar a aprendizagem nas escolas ou aumentar a arrecadação de impostos.

Introdução

Para avaliar se as políticas atingem os objetivos propostos, é importante que haja uma análise constante e rigorosa, com base em métricas que realmente respondam por um critério adequado ao objetivo.

Introdução

Esse tipo de avaliação faz parte de uma agenda global das últimas décadas de **políticas públicas baseadas em evidências**, que visa retirar o foco dos insumos e colocar nos resultados [Gertler et al., 2016].

Introdução

A maior motivação para se fazer políticas baseadas em evidência é simples, mas muito forte, e é a primeira lição da maioria dos manuais de introdução à economia: **os recursos são escassos**. Portanto, é necessário que se gaste efetivamente.

Introdução

Exemplos de Perguntas:

- Mais recursos para escolas melhora o desempenho dos alunos?
- Mais policiais nas ruas implica em redução da criminalidade?
- Mais fiscalização sobre recursos públicos implica em redução da corrupção?
- Filho(a)s de pais que recebem o BF têm melhora no desempenho escolar?
- BF reduz mortalidade infantil?
- Saneamento básico reduz mortalidade infantil?

Introdução

A utilização da avaliação de impacto de uma política pública tem dois tipos de usos principais: interno e externo.

No uso interno, a depender dos resultados encontrados, ela tem como objetivo subsidiar decisões acerca da operação do programa.

Já o uso externo tem como principal objetivo subsidiar a aplicação ou alteração de programas semelhantes em outros lugares ou períodos por outros *policymakers*.

Definição do problema

A questão típica da avaliação de políticas públicas é centrada na definição do **tratamento**, T ou seja, a intervenção da política em uma pessoa, grupo de pessoas ou local. Assim:

$$T_i = \begin{cases} 1 & \text{Se o indivíduo recebeu o tratamento} \\ 0 & \text{Se não recebeu} \end{cases}$$

O interesse da avaliação tipicamente é **analisar qual foi o efeito do tratamento em uma determinada variável de interesse**, Y . O problema é que nós nunca conseguimos observar um mesmo indivíduo com e sem tratamento na mesma situação.

Definição do problema

Por exemplo, suponha que alguém queira implementar uma política em determinado local.

O impacto dessa política pode ser definido como a diferença entre duas realidades: o que aconteceu posteriormente à ação governamental e o que teria acontecido no lugar sem essa política.

Definição do problema

A maior dificuldade dessa formulação é justamente mostrar o que aconteceu na segunda realidade, pois ela não ocorreu de fato, ou seja, **é hipotética**. Assim, para definir o impacto, é preciso buscar meios de se aproximar desse contrafactual [Peixoto et al., 2012].

Causalidade

Uma das questões centrais desse tipo de avaliação é isolar o **efeito causal** da política sobre a variável de interesse.

Causalidade

Caso fosse possível garantir que nada mais poderia afetar o comportamento do indicador, não seria nem necessária uma comparação com a situação hipotética de não-intervenção, dado que toda variação da variável de interesse seria causada pelo tratamento.

Causalidade

Na prática, o uso da modelagem estatística visa a aproximar essa situação e extrair o efeito, descontando os demais fatores.

Nesse sentido, é preciso que a escolha da metodologia seja adequada para defender que o efeito da política foi de fato isolado.

A simples análise estatística pode levar a conclusões erradas sobre os fenômenos.

Correlação espúria

Nesse sentido, é preciso que a escolha da metodologia seja adequada para defender que o efeito da política foi de fato isolado.

A simples análise estatística pode levar a conclusões erradas sobre os fenômenos.

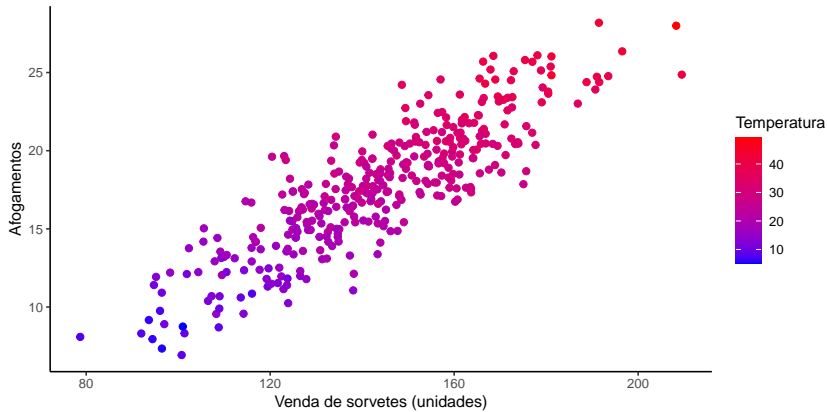
Correlação espúria

Um exemplo clássico e bastante ilustrativo é a correlação positiva entre vendas de sorvete e afogamentos no mar. Uma análise descuidada pode levar à tese de que tomar sorvete aumenta sua chance de se afogar, o que logicamente não é o caso.

Correlação espúria

```
## Vamos criar uma equação para venda de sorveter e para afogamentos
set.seed(2000)
Temperatura <- rnorm(365, mean = 25, sd = 8)
venda_sorvetes <- 50+Temperatura*2.5 + rnorm(365, mean = 30, sd = 10)
afogamentos <- Temperatura*0.5 + rnorm(365, mean = 5, sd = 1)
df <- data.frame(Temperatura, venda_sorvetes, afogamentos)
```

Correlação espúria



```
print(cor(df$venda_sorvetes, df$afogamentos))
```

```
[1] 0.8712027
```

Validação interna X Validação externa

Nem todos os desenhos de avaliação de impacto são iguais e há custos em optar por um tipo de análise ou outra. Dois componentes importantes para avaliarmos os resultados da nossa estimação do tratamento são sua validade interna e sua validade externa.

- **Validade interna** - Significa que podemos estabelecer uma relação causal entre a política e a variável de interesse dentro do contexto do estudo;
- **Validação externa** - Ocorre quando podemos generalizar os resultados do nosso estudo para outros contextos.

Validação interna X Validação externa

No caso de políticas públicas, é muito difícil fazer uma avaliação que tenha validação externa completa, pois alterações de local, grupos tratados, períodos e até pequenas alterações na política podem alterar completamente o resultado final.

Validação interna X Validação externa

Entretanto, a depender do caso, se bem executada, a análise pode ter um bom grau de **validação externa**. Assim, é importante estar atento a sensibilidade que os resultados de uma política têm em relação ao contexto em que ela foi aplicada, principalmente para um gestor que a use como referência para um outro programa.

Heterogeneidade de impacto

Normalmente, uma avaliação de impacto vai se preocupar com **a média do impacto no grupo tratado**.

Entretanto, por definição, os objetos das políticas públicas são as indivíduos pertencentes a um ou mais grupos.

Heterogeneidade de impacto

Como essas pessoas não são iguais entre si, o efeito que a ação governamental tem sobre elas vai ser naturalmente diferente.

Por exemplo, um programa de transferência de renda hipotético que destina 200 reais para pessoas que recebem até 1000 reais vai atuar em grupo de pessoas bem heterogêneas, por isso, provavelmente, o resultado também será heterogêneo.

Por isso, independente do indicador que pretende-se avaliar, seja impacto na educação, mortalidade infantil ou até renda após o programa, é esperado que o efeito não seja homogêneo para todas as pessoas.

Estratégias de identificação

Como dissemos, obviamente não é possível ter acesso à realidade alternativa em que não houve a política pública que queremos estimar o impacto.

A solução mais usual nas ciências naturais, como a Física e a Química, é realizar experimentos.

No caso de grande parte das questões dessas disciplinas é possível contruir um meio de controlar o efeito de fatores externos e testar uma hipótese de maneira isolada.

Estratégias de identificação

Nas ultimas décadas, a popularidade dos estudos que utilizam experimentos para avaliar políticas públicas cresceu muito, pois são considerados o padrão ouro da ciência. Entretanto, a aplicação desse tipo de abordagem nem sempre é possível nessa área, pois experimentos às vezes não são exequíveis, ou são muito caros.

Atribuição aleatória

O ponto mais importante para a robustez dos métodos experimentais é a atribuição aleatória entre grupos de **controle** e **tratamento**.

Isto é, no primeiro momento os indivíduos não estão em nenhum grupo e são, posteriormente, colocados no grupos de forma aleatória e então, os que estiverem no grupo de tratamento sofrem a intervenção.

Assim, não há diferença entre os grupos no primeiro momento e, portanto, caso o experimento seja bem desenhado, as diferenças de trajetórias na variável de interesse são causadas exclusivamente pelo tratamento.

Atribuição aleatória - Exemplo

Para demonstrar o poder da atribuição aleatória, iremos utilizar o exemplo dado anteriormente do programa de transferência de renda de 200 reais.

Suponha que, para avaliar o impacto da política na renda futura das pessoas, pudéssemos, antes do início da política, sortear parte do grupo focal para tratamento e parte para controle.

Ou seja, não estaríamos de fato implementando a política em sua totalidade, mas sim, a testando.

Atribuição aleatória - Exemplo

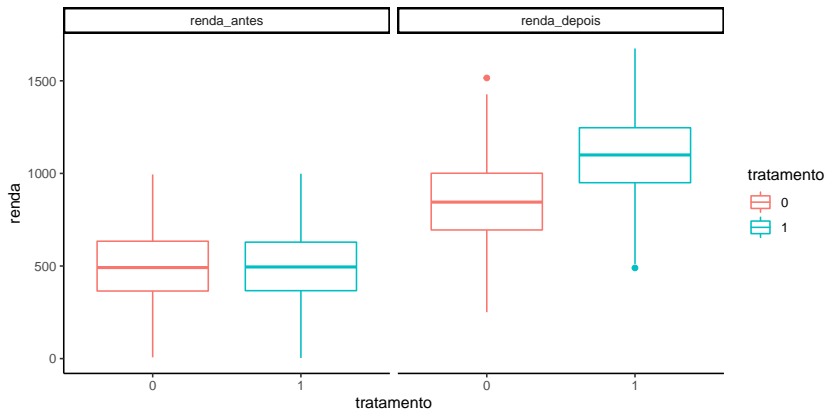
Para fins didáticos, vamos supor que a renda no momento anterior a política tenha uma distribuição normal. Também assumiremos que o impacto do tratamento é homogêneo e igual a 250 reais após a aplicação do programa, ou seja, afeta todos os indivíduos da mesma forma.

Atribuição aleatória - Exemplo

```
id <- seq(1:5000) #id de cada indivíduo

# distribuição normal truncada entre 0 e 1000
renda_antes <- rtruncnorm(n=5000, a=0, b=1000, mean=500, sd=200)
tratamento <- as.factor(rbinom(n=5000, size=1, prob=0.5))
df <- data.frame(id, renda_antes, tratamento)
df$renda_depois <- renda_antes*1.1 + as.numeric(tratamento)*250 +
  rnorm(5000, mean = 50, sd = 50)
```

Atribuição aleatória - Exemplo



Atribuição aleatória - exemplo

Assim, nesse exemplo hipotético, poderíamos estimar o impacto do programa com algum procedimento estatístico simples, como uma regressão linear. Assim, podemos ver que o impacto encontrado, medido pelo parâmetro da regressão, é praticamente igual ao efeito definido.

```
#diferença entre antes e depois  
df$diff <- df$renda_depois - df$renda_antes  
  
lm(diff ~ tratamento, data = df) %>%  
  tidy() %>%  
  kable()
```

term	estimate	std.error	statistic	p.value
(Intercept)	349.7236	1.058283	330.4632	0
tratamento1	251.0386	1.504786	166.8268	0

Viés de seleção

No caso de políticas públicas, raramente temos atribuições aleatórias, o que gera um problema de **viés de seleção**.

Primeiramente pois a própria definição de grupo de controle pode não ser óbvia. No caso do programa de transferência de renda, a definição do grupo de controle seria especialmente difícil, pois se todas as pessoas com renda condizente com o programa estivessem dentro dele, não haveria indivíduos com renda baixa no local para fazer a comparação.

Viés de seleção

Além disso, mesmo que uma política ainda pudesse dispor de um grupo de comparação, pouco provável que o tratamento seja 100% aleatório.

Por exemplo, digamos que nosso objetivo seja avaliar o impacto da construção de uma linha do metrô na renda dos moradores próximos a ela.

Não é factível colocar quem mora longe das estações como grupo de controle, pois, muito provavelmente, são grupos muito distintos. Isso acontece porque a decisão de onde colocar a linha do metrô não é aleatória, já que leva em consideração características do local, como renda média, densidade, demografia, entre outras.

Métodos quasi-experimentais

Para tratar esse tipo de problema, existem alguns métodos estatísticos que podem ser aplicados.

Esse métodos visam aproximar a análise a um experimento, são os chamados métodos **quasi-experimentais** ou experimentos naturais.

Métodos quasi-experimentais

A diferença entre essas duas definições é que nos Quase-experimentos o critério de atribuição de tratamento é definido pelo investigador, enquanto nos experimentos naturais a exposição ao tratamento é determinada pela natureza ou por fatores externos.

Entretanto esses dois conceitos são semelhantes, dado que sempre há um papel do pesquisador em definir a pergunta a ser feita.

Métodos quasi-experimentais

Existem diversos métodos e procedimentos que se encaixam nesse tipo de abordagem, os quais serão trabalhados ao longo do curso. Cada método tem um propósito e aplicação diferente e mais adequado. Boa parte do desafio da avaliação de políticas públicas é construir uma abordagem que utilize das ferramentas mais adequadas.

Pareamento

As técnicas de pareamento são utilizadas quando o tratamento não ocorre de maneira aleatória e os grupos de tratamento e controle acabam apresentando características muito diferentes.

Esses métodos têm como objetivo balancear os grupos com base em covariáveis, ou seja, outros dados além da variável de interesse.

Na prática, essa técnica seleciona os indivíduos do grupo de controle que sejam os mais semelhantes possíveis com relação aos indivíduos que receberam o tratamento. Normalmente, esse procedimento faz com que algumas observações do grupo de controle sejam eliminadas da amostra.

Pareamento

Por isso, a estratégia mais comum é parear utilizando **escores de propensão**.

Resumidamente, é estimada uma regressão logística, em que a variável dependente é o tratamento.

Assim, com base nos parâmetros estimados, é calculada a probabilidade do indivíduo pertencer ao grupo de tratamento. Assim, cada indivíduo do grupo de tratamento é pareado com um, ou mais, do grupo de controle que tenham essa probabilidade, ou “escore”, semelhante.

Pareamento - Exemplo

Um exemplo de trabalho que utiliza **matching** para avaliar uma política pública é Andrade et al. [2012].

Os autores investigaram os impactos do **Programa Bolsa Família** na imunização de crianças de 0 a 6 anos utilizando dados de 2005.

Nele foram comparadas famílias que estavam no programa com famílias elegíveis, mas não presentes no programa.

O ponto principal da utilização deste método é que esses grupos apresentam características diferentes, portanto é necessária uma readequação do grupo de controle. Os resultados encontrados mostraram que o programa não afetou a imunização das crianças.

Diferenças-em-diferenças

É um método de dados em painel muito popular na literatura econômica que calcula o efeito de um tratamento comparando a mudança média ao longo do tempo na variável de interesse para os tratados, em comparação com a mudança média para o grupo de controle.

Diferenças-em-diferenças

O principal requisito para esse método poder ser aplicado é o de que as tendências de ambos os grupos na variável de interesse sejam paralelas antes do tratamento. Ou seja, a hipótese é de que as diferenças não observáveis entre os grupos se mantenha constante ao longo do tempo. Assim, o efeito do tratamento é o responsável por alterações de trajetória.

Diferenças-em-diferenças - Exemplo

Muitas vezes, a hipótese de tendência comum não é atendida, o que faz com que muitos estudos utilizem pareamento para atingir isso.

Mattos and Mazetto [2019] utilizam exatamente essa abordagem para estimar o efeitos de curto prazo do programa Mais Médicos, lançado em 2013.

Assim, eles usam dados de municípios de 2010 até 2015 e mostram que o programa teve ao menos dois efeitos positivos: o aumento em consultas, agendamentos e visitas a domicílios, e o impacto negativo na hospitalização. Entretanto, não encontraram efeito significativo na mortalidade.

Controle sintético

Esse método é de certa forma semelhante a um modelo de **diferenças-em-diferenças**, mas utilizado normalmente quando há apenas uma unidade tratada.

É feita a construção de uma combinação linear ponderada da variável de interesse das unidades do grupo de controle antes do tratamento ocorrer, o que faz com que a trajetória dessa combinação se torne semelhante à trajetória da unidade tratada.

Assim, se as duas trajetórias divergirem após à implementação de uma política, isso pode se dever ao tratamento.

Controle sintético - Exemplo

Por sua característica, esse é um método interessante para aplicação de avaliação de políticas nacionais. Um exemplo disso é o paper de Chamon et al. [2017], que avaliou um programa de intervenção cambial por meio de swaps do Banco Central do Brasil em 2013. Utilizando outros países emergentes como controle, eles estimaram que a política levou a uma apreciação do Real em um excedente de 10%.

Variáveis instrumentais

É um método que busca uma variável externa (instrumento) que é correlacionada com as variáveis explicativas, mas não com a variável de interesse.

Assim, essa técnica permite estimar uma regressão quando as variáveis explicativas são correlacionadas com o termo de erro, ou seja, quando há endogeneidade.

Basicamente, este método permite o instrumento dividir as variáveis explicativas em duas partes. Uma delas é correlacionada com o erro e a outra não. Assim, tornando possível separar a parte que não tem correlação para a estimação da regressão de forma não viesada.

Variáveis instrumentais - Exemplo

Um exemplo bastante comum do uso de variáveis instrumentais é a estimação do retorno da educação nos salários.

Isso acontece pois existe um fator não observável, que é a habilidade pessoal.

Pessoas com maior habilidade tem maior probabilidade de fazer ensino superior. Ao mesmo tempo, a habilidade também afeta os salários de forma independente à educação formal. Assim, um bom instrumento precisa ser correlacionado com a probabilidade de ir à universidade, mas não pode ser correlacionado com os salários. Card [1993], propõem estimar o efeito utilizando a distância em relação às universidades. Outros autores, como Blackburn and Neumark [1993] utilizam a educação dos pais como instrumento.

Regressão descontínua

O objetivo dessa abordagem é comparar observações que estejam perto de um limite que estabelece um limite entre o grupo de tratamento e o de controle. A lógica desse método é a de que indivíduos próximos a essa limite seriam extremamente semelhantes, sendo a designação para um ou para outro grupo completamente aleatória. A aplicação é possivelmente mais restritiva do que os outros métodos citados e muito dependente dos dados disponíveis e do critério de aplicação da política pública.

Regressão descontínua - Exemplo

Chamon et al. [2019] utilizaram uma descontinuidade nas regras eleitorais municipais de 2000 para investigar se a competição política teria impacto nas escolhas das políticas fiscais. A descontinuidade em questão era a de que, para municípios menores de 200 mil habitantes, não havia segundo turno. Os resultados mostram que a competição faz com que se aumentem os gastos com infraestrutura em relação aos gastos correntes. Além disso, o impacto encontrado foi maior para os casos em que os candidatos poderiam concorrer à reeleição.

- Mônica Viegas Andrade, Flávia Chein, Laeticia Rodrigues de Souza, and Jaume Puig-Junoy. Income transfer policies and the impacts on the immunization of children: the bolsa familia program. *Cadernos de saude publica*, 28:1347–1358, 2012.
- McKinley L Blackburn and David Neumark. Omitted-ability bias and the increase in the return to schooling. *Journal of labor economics*, 11(3):521–544, 1993.
- David Card. Using geographic variation in college proximity to estimate the return to schooling. *NBER working paper*, (w4483), 1993.
- Marcos Chamon, Márcio Garcia, and Laura Souza. Fx interventions in brazil: a synthetic control approach. *Journal of International Economics*, 108:157–168, 2017.
- Marcos Chamon, Sergio Firpo, João MP de Mello, and Renan Pieri. Electoral rules, political competition and fiscal expenditures: regression discontinuity evidence from brazilian municipalities. *The Journal of Development Studies*, 55(1):19–38, 2019.

Paul J Gertler, Sebastian Martinez, Patrick Premand, Laura B Rawlings, and Christel MJ Vermeersch. *Impact evaluation in practice*. The World Bank, 2016.

Enlinson Mattos and Debora Mazetto. Assessing the impact of more doctors' program on healthcare indicators in brazil. *World Development*, 123:104617, 2019.

Betânia Peixoto, Cristine Campos de Xavier Pinto, Lycia Lima, Miguel Nathan FOGUEL, Ricardo Paes BARROS, and Naercio MENEZES FILHO. *Avaliação econômica de projetos sociais*. Fundação Itaú Social, 2012.