

lista4_gabarito

Gabriel_Mardegan

2024-12-17

Lista 4: causalidade e resultados potenciais

Gabarito

- 1) O problema fundamental da inferência causal se refere à impossibilidade de observar, simultaneamente, os dois resultados potenciais - $Y(0)$ e $Y(1)$ - para a mesma unidade, sob diferentes condições de tratamento ($T=0$ e $T=1$). Ou seja, apenas conseguimos coletar dados sobre o que acontece de fato, o factual, e não sobre o que teria acontecido, o contrafactual. Isso acarreta em um problema de ausência de dados, o que gera a necessidade de se estimar o efeito causal.
- 2) A simples comparação entre um grupo tratamento de municípios (T) e um grupo controle de municípios (C) pode levar a uma análise enviesada, pois o fato de um município ter sido contemplado com a política pode estar correlacionado a outras características que também estão influenciando a variável resposta que se deseja aferir. Isso resulta em um problema de confusão, onde fatores não observados ou não controlados podem afetar tanto a probabilidade de receber o tratamento quanto os resultados, mas se não controlados, esses efeitos podem ser atribuídos ao tratamento, erroneamente. Como a inferência causal requer comparabilidade entre grupo de tratamento e grupo controle, é fundamental que se controle outras variáveis que podem estar associada à implementação da política.
- 3)
 - a. Viés de seleção se refere a diferenças sistemáticas entre os grupos de tratamento e de controle, consequência do processo não aleatório de seleção dos indivíduos. Essas diferenças sistemáticas entre os grupos impedem a aferição da causalidade.
 - b. Investigação sobre efeitos da rotatividade de diretores de escola sobre o desempenho dos alunos. O tratamento seria a baixa rotatividade dos diretores (por exemplo, um mesmo diretor se manter por cinco anos ou mais na mesma unidade), e o controle seriam escolas cuja diretoria se mantém, em média, menos de cinco anos sem alterações. A variável resposta seria o desempenho da escola no IDEB (índice de desenvolvimento da educação básica). Como não se trata de um estudo experimental, o grupo de tratamento e grupo de controle são formados a partir da realidade do conjunto de escolas.

Como o estudo é observacional, o viés de seleção ocorre porque escolas com características pré-existentes podem influenciar tanto a rotatividade quanto o desempenho no IDEB. Por exemplo, escolas com baixos índices de desempenho no IDEB podem ser menos atrativas para diretores, levando a uma maior rotatividade. Além disso, variáveis omitidas, como o nível socioeconômico dos alunos, a infraestrutura escolar ou o suporte

recebido pela escola da rede de ensino, podem afetar tanto a rotatividade dos diretores quanto o desempenho no IDEB, criando uma relação espúria entre tratamento e resultado.

4)

a.

$$E[Yi(0)|T = 1] - E[Yi(0)|T = 0]$$

- b. A equação do viés de seleção indica a diferença nos resultados esperados ($Yi(0)$) entre os grupos de tratamento ($T=1$) e controle ($T=0$) no cenário em que nenhum dos dois grupos recebe o tratamento. Essa diferença reflete a falta de comparabilidade entre os dois grupos antes da intervenção, ou seja, o viés gerado por características pré-existentes que influenciam a variável dependente. Em um caso de diferença igual a zero, os dois grupos seriam perfeitamente comparáveis.

No caso do exemplo da questão 3b, as mesmas escolas com baixa rotatividade, caso fossem manipuladas de modo a trocar a diretoria com frequência, deveriam ter os mesmos resultados que o grupo de escolas com alta rotatividade. Portanto, a equação do viés de seleção mede essa diferença não atribuída ao efeito do tratamento, mas sim às características intrínsecas dos grupos.

- 5) Um desenho experimental de pesquisa tende a eliminar o viés de seleção, pois utiliza a aleatorização para determinar quais unidades serão alocadas ao grupo de tratamento e ao grupo de controle. Esse método garante que as diferenças entre o grupo de tratamento e o grupo controle sejam apenas a presença da intervenção/tratamento, cujo efeito se pretende estimar.

6)

- a. Estimandos causais (também nomeados como quantidades de interesse ou parâmetros causais) são as quantidades causais que se deseja estimar. Eles são definidos matematicamente para descrever como a variável dependente (Y) mudaria em resposta à manipulação do tratamento (T).

- b. Para a avaliação do exemplo, o estimando causal mais apropriado é o ATT (Average Treatment on the Treated). Como apenas parte dos jovens interessados participou do programa, o ATT permite comparar os efeitos específicos sobre o perfil de jovem que teria interesse em participar do programa, possibilitando estimar com mais precisão seus efeitos. O uso do ATE, neste caso, tende a um cálculo impreciso do impacto do programa sobre os jovens, pois consideraria também o efeito sobre jovens que não participariam do programa.

- 7) Ao incluir uma variável omitida no modelo de regressão, busca-se atender à hipótese de média condicional do erro igual a zero ($E[e|X]=0$), de forma a não haver nenhuma correlação entre o erro e os regressores, eliminando, assim, o viés de seleção. Os efeitos dessa variável são controlados, e a regressão ajustada passa a refletir a relação causal com maior precisão. Assim, os parâmetros da regressão linear podem se alterar, pois estarão considerando apenas o efeito de determinada variável independente. Dessa maneira, o coeficiente angular anterior ao controle poderia estar subestimando ou superestimando o efeito da variável preditora, dependendo de quais variáveis foram omitidas no modelo anterior. Da mesma maneira, o novo cálculo dos parâmetros poderia alterar o intercepto do modelo.