

ACH3657

Métodos Quantitativos para Avaliação de Políticas Públicas

Aula prática 08
Regressão Múltipla

Alexandre Ribeiro Leichsenring
alexandre.leichsenring@usp.br



Desempenho em Matemática e o Programa de Merenda Escolar

Seja $mate10$ a percentagem de alunos do primeiro ano do ensino médio aprovados em um exame de matemática. Suponha que desejamos estimar o efeito do programa de merenda escolar financiado pelo governo sobre o desempenho dos alunos. Esperamos que o programa de merenda tenha um efeito *ceteris paribus* positivo sobre o desempenho: todos os outros fatores permanecendo iguais, se um estudante pobre torna-se beneficiário de um programa de merenda e passa a ter refeições regularmente, seu desempenho deveria melhorar. Seja $prgalm$ a percentagem de estudantes participantes do programa de merenda escolar em cada escola. Portanto, o modelo de regressão simples é

$$mate10 = \beta_0 + \beta_1 prgalm + u, \quad (1)$$

em que u contém características da escola e do estudante que afetam o desempenho escolar total.

O modelo ajustado é:

$$\widehat{mate10} = 32,14 - 0,319 prgalm$$

- A equação prevê que se a participação dos estudantes no programa de merenda escolar aumenta em dez pontos percentuais, a percentagem de estudantes que passa no exame de matemática cai cerca de 3,2 pontos percentuais.
- Realmente devemos acreditar que a participação maior no programa de merenda escolar causa, de fato, um desempenho pior?
- Muito provavelmente não.
- Uma explicação melhor é que o termo erro u na Equação (1) esteja correlacionado com [prgalm](#).
- De fato, u contém fatores como a taxa de pobreza das crianças que frequentam a escola, que afeta o desempenho dos estudantes e está altamente correlacionada com a qualificação no programa de merenda.
- Variáveis como qualidade e recursos da escola também estão contidas em u e, provavelmente, estão correlacionados com [prgalm](#).
- É importante lembrar que a estimativa -0,319 é somente para essa amostra particular, mas seu sinal e magnitude nos fazem suspeitar de que u e x sejam correlacionadas, de modo que a regressão linear é viesada.

Exercício 1

i) Estime o modelo:

$$mate10 = \beta_0 + \beta_1 \log(gasto) + \beta_2 prgalm + u$$

e relate os resultados da maneira habitual, inclusive o tamanho da amostra e o R-quadrado. Os sinais dos coeficientes de inclinação são os que você esperava?

- ii) O que você conclui do intercepto que você estimou na parte (i)? Particularmente, faz sentido definir as duas variáveis explicativas com valor zero? [*Sugestão*: Lembre-se que $\log(1) = 0$].
- iii) Agora execute a regressão simples de $mate10$ sobre $\log(gasto)$, e compare o coeficiente de inclinação com a estimativa obtida na parte (i). O efeito do gasto estimado agora é maior ou menor do que na parte (i)?
- iv) Encontre a correlação entre $\log(gasto)$ e $prgalm$. O seu sinal faz sentido para você?
- v) Use a parte (iv) para explicar suas constatações na parte (iii).