

Quarta Lista de Exercícios

Mestrado Profissional em Economia
Professor: Hedibert Freitas Lopes
Insper Instituto de Educação e Pesquisa

Econometria Avançada
Monitor: Guilherme Piantino
November/2024

Questão 1: Regressão: endogeneidade e variável instrumental

O arquivo `mroz.txt` apresenta o salário (log) como variável dependente de interesse. O objetivo é estimar o efeito da educação no salário, mas existe o potencial de correlação entre educação e outras variáveis que também influenciam o salário. Esse é o ambiente onde a simples estimação de uma regressão não funcionará pois uma das hipóteses é a de exogeneidade da variável explicativa (educação). Nesse caso, vimos que, quando presente, uma (ou várias) variável instrumental pode mitigar esse viés e ajudar na obtenção do efeito puro da educação no salário, como por exemplo a educação do pai e a educação da mãe, ambos em anos. Aqui, temos as seguintes variáveis:

Outcome: Salário (log) (`lwage`)

Treatment: Educação em anos (`educ`)

Variável exógena: Experiência no mercado de trabalho em anos (`exper`)

Variáveis instrumentais: Educação do pai e da mãe em anos (`fatheduc + motheduc`)

Mostre que a regressão tradicional, onde `lwage` é função linear de `educ` e `exper` com erros Gaussianos, superestima o efeito da educação no salário. Faz diferença no resultado final (estimação do efeito) a presença da variável exógena `exper` na equação de “outcome” e na equação de “treatment”?

```
data = read.table("mroz.txt",header=TRUE)
attach(data)
n = nrow(data)
par(mfrow=c(1,2))
plot(educ,lwage,xlab="Educação (em anos)",ylab="Log salário")
plot(educ,motheduc,xlab="Educação da mãe (em anos)",ylab="Educação (em anos)")
```

Questão 2: Método de regressão descontínua

O arquivo `acidente-idade.txt` contém a taxa de mortalidade (por 100.000) em acidentes de automóvel para coorte de idades em volta dos 21 anos, idade em que bebida alcoólica passou a ser legalmente consumida por jovens adultos. Utilize o método de regressão descontínua para obter estimativa do impacto da mudança de idade tem no nível de acidentes de automóvel. Compare tais estimativas quando as funções que relacionam acidentes (y) à idade (x) são i) lineares, ii) quadráticas e iii) cúbicas. Comente seus achados.

```
data = read.table("acidente-idade.txt",header=TRUE)
n = nrow(data)
attach(data)
plot(idade,acidente,xlab="Idade (em anos)",ylab="Taxa de mortalidade (por 100.000)",pch=16)
abline(v=21,lty=2)
```