ACH3657

# Métodos Quantitativos para Avaliação de Políticas Públicas

## Aula Teórica 3

## Estimação dos parâmetros

Alexandre Ribeiro Leichsenring
alexandre.leichsenring@usp.br

# Estimativas de mínimos quadrados ordinários

- Trataremos de como estimar os parâmetros $\beta_0$ e $\beta_1$ da equação de regressão:

$$y = \beta_0 + \beta_1 x + u$$

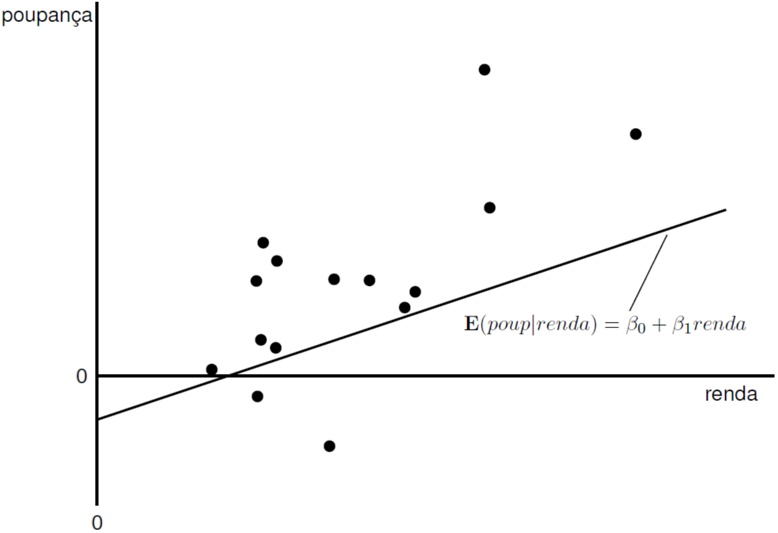
- Seja $\{(x_i, y_i) : i = 1, \ldots, n\}$ uma amostra aleatória de tamanho $n$ da população. Então podemos escrever:

$$y = \beta_0 + \beta_1 x_i + u_i,$$

  onde $u_i$ é o termo erro para a observação $i$, contendo todos os fatores, além de x que afetam $y$.

- Como um exemplo, $x$, poderia ser a renda anual e $y$ a poupança anual para a família $i$ durante um determinado ano. Se coletarmos dados de 15 famílias, então $n = 15$.

Gráfico de dispersão de poupança e renda de 15 famílias e a regressão populacional

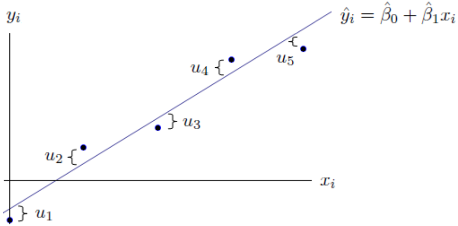


- Como usar esses dados, a fim de obter estimativas do intercepto e da inclinação na regressão populacional da poupança sobre a renda?
- Há muitas maneiras.
- Uma maneira é a minimização dos quadrados dos resíduos (termos de erro) estimados.

➤ Embora Wooldridge use outra abordagem, ele chega aos mesmos estimadores, e inclusive dá a eles o nome de *Estimadores de mínimos quadrados ordinários* (MQO).

## Estimação dos parâmetros por mínimos quadrados

- Queremos que nossas estimativas para os parâmetros $\beta_0$ e $\beta_1$ minimizem a distância entre os pontos previstos ($\hat{y}_i$) e os observados ($y_i$).




- Devemos escolher $\beta_0$ e $\beta_1$ que minimizem os resíduos.
- Uma boa opção é minimizar:

$$\sum_{i=1}^{n} u_i^2 \ ,$$

conhecida por soma de quadrado dos resíduos (SQRes).

$$
\begin{aligned}
\text{SQRes} &= \sum_{i=1}^{n} u_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \\
&= \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2
\end{aligned}
$$

Assim, minimizar **SQRes** equivale minimizar

$$
\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2.
$$

Para qualquer amostra, os valores de $\hat{\beta}_0$ e $\hat{\beta}_1$ que minimizam **SQRes** são dados por:

$$
\begin{aligned}
\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x} \\
\hat{\beta}_1 &= \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}.
\end{aligned}
$$

- A equação para $\hat{\beta}_1$ corresponde à covariância amostral entre $x$ e $y$, dividida pela variância amostral de $x$.
- Uma implicação imediata é que se $x$ e $y$ são positivamente correlacionados na amostra, então $\beta_1$ é positivo; se $x$ e $y$ são negativamente correlacionados, então $\beta_1$ é negativo.

- A única hipótese necessária para se calcular as estimativas para uma amostra particular é:

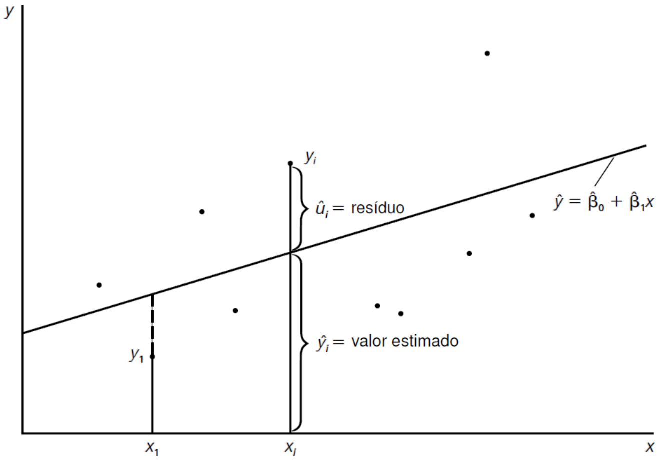
$$\sum_{i=1}^{n}(x_i - \bar{x})^2 > 0$$

que é verdadeira sempre que os $x_i$, na amostra não são todos iguais a um mesmo valor

- Se a hipótese não se sustentar, então fomos infelizes em obter nossa amostra da população, ou não especificamos um problema interessante ($x$ não varia na população!)
- Por exemplo, se $y = salarioh$ e $x = educ$, então a hipótese não se mantém apenas se todos na amostra têm a mesma quantidade de anos de educação formal.
- Se apenas uma pessoa tem uma quantidade diferente de anos de educação formal, então a hipótese se sustenta, e as estimativas de MQO podem ser calculadas.

- Para cada valor $x_i$ na amostra, há um valor estimado $\hat{y}_i$ para $y$
- O *resíduo* para a observação $i$ é a diferença entre o valor verdadeiro de $y$, e seu valor estimado:

$$\hat{u}_i = y_i - \hat{y}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)$$



- $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1$ é chamada função de regressão amostral (FRA)
- É importante lembrar que a função de regressão populacional (FRP) é algo fixo, porém desconhecido, na população
- Como a FRA é obtida para uma dada amostra de dados, uma amostra nova gerará uma inclinação e um intercepto diferentes
- Em muitos casos, $\hat{\beta}_1$ é de interesse fundamental
  - Ela nos diz o quanto varia $y$ quando $x$ aumenta em uma unidade.

$$\Delta\hat{y} = \hat{\beta}_1 \Delta x$$

  - Dada qualquer variação em $x$ (seja positiva ou negativa), podemos calcular a variação prevista em $y$.

## Exemplo: Salários de diretores executivos e retornos de ações

- Para a população de diretores executivos, seja $y$ o salário anual em milhares de dólares
- Seja $x$ o retorno médio da ação sobre o patrimônio ($rma$), dos três anos anteriores, da empresa do diretor executivo

  (Por exemplo, se $rma = 10$, então o retorno médio da ação sobre o patrimônio é de 10%)
- Para estudar a relação entre essa medida do desempenho das empresas e a remuneração dos seus diretores executivos, postulamos o modelo simples:

$$salario = \beta_0 + \beta_1 rma + u$$

- O parâmetro de inclinação $\beta_1$, mede a variação no salário anual em milhares de dólares quando o retorno da ação aumenta em um ponto percentual.
- Como um $rma$ mais elevado é melhor para a empresa, esperamos que $\beta_1 > 0$.
- Na amostra, o salário médio anual é \$ 1.281.120; (mínimo e máximo são \$ 223.000 e \$ 14.822.000, respectivamente.
- O retorno médio das ações para os anos 1988,1989 e 1990 é de 17,18%, sendo que os valores menor e maior são de 0,5 e 56,3%, respectivamente.

- Usando os dados do arquivo ceosal1.sav, a reta de regressão de MQO que relaciona salário a rma é:

$$\hat{salario} = 963,191 + 18,501 rma$$

- Usamos “salário chapéu” para indicar que é uma equação estimada

**Interpretação**

1. $rma = 0 \Rightarrow$ salário previsto é o intercepto ($\beta_0$)
2. Podemos escrever a variação prevista no salário como uma função da variação em $rma$:

$$\Delta \hat{salario} = 18,501(\Delta rma)$$

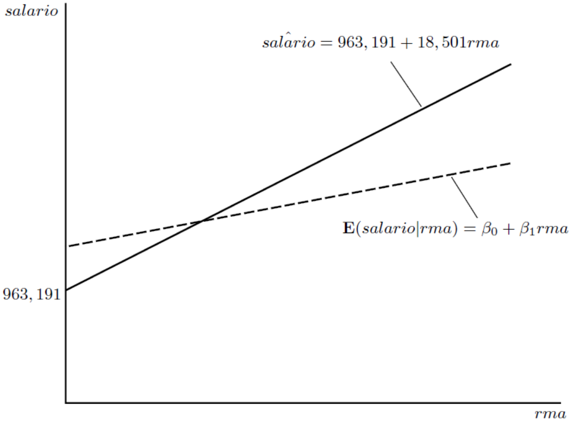
    - Se o retorno da ação aumenta um ponto percentual, $\Delta rma = 1$, então espera-se que salário variará cerca de 18.500
    - Como a equação é linear, esse valor é a variação estimada independentemente do salário inicial

3. Podemos usar a equação estimada para comparar salários previstos para valores diferentes de $rma$:

    - Suponha $rma = 30$. Portanto

$$\hat{salario} = 963,191 + 18,501(30) = 1.518,221$$

- Entretanto, isso não significa que um determinado diretor executivo, cuja empresa tenha um rma = 30, ganhe $ 1.518.221
- Muitos outros fatores afetam o salário
- Essa é somente a nossa previsão a partir da reta de regressão de MQO




- Não podemos dizer o quão próxima a FRA está da FRP
- Outra amostra de dados levará a uma reta de regressão diferente...

## Exemplo: Salários e educação

Para a população de pessoas na força de trabalho em 1976, seja $y = salarioh$, em que $salarioh$ é mensurado em \$ por hora. Seja

- $x = educ$ (e.g., $educ = 12$ corresponde ao ensino médio completo)
- O salário horário médio na amostra é \$ 5,90
- Usando os dados do arquivo wage1.sav, com n = 526 indivíduos, obtemos a seguinte reta de regressão de MQO:

$$\widehat{salarioh} = -0,90 + 0,54\,educ$$

### Interpretação

- $\hat{\beta}_0 = -0,90$ (!)
  - Apenas 18 pessoas na amostra têm $educ \leq 8$. A regressão não deve fazer boas previsões para níveis de educação formal "baixos"
- Para $educ = 8$, $\widehat{salarioh} = -0,90 + 0,54(8) = 3,4$
- $\Delta educ = 1 \Rightarrow \Delta salarioh = 0,54$
- $\Delta educ = 4 \Rightarrow \Delta salarioh = 4(0,54) = 2,16$
- Devido à natureza linear da equação, outro ano de educação formal aumenta o salário na mesma quantidade...

## Exemplo: Resultados eleitorais e gastos de campanha

O arquivo vote1.sav contém dados sobre resultados eleitorais e gastos de campanha de 173 disputas entre dois partidos nos EUA. Há dois candidatos em cada disputa: A e B. Seja:

- $votoA$: % de votos recebida pelo Candidato A
- $partA$: % dos gastos totais de campanha que cabem ao Candidato A

Muitos outros fatores além de $partA$ afetam o resultado eleitoral (incluindo a qualidade dos candidatos e os valores absolutos dos gastos de A e B). No entanto, podemos estimar um modelo de regressão simples para descobrir se gastar mais do que o concorrente implica uma percentagem maior de votos.

A equação estimada usando as 173 observações é:

$$votoA = 26,81 + 0,464 partA$$

- Se a parte dos gastos do Candidato A aumenta em um ponto percentual, o Candidato A recebe quase meio ponto percentual (0,464) a mais da votação total
- Não fica claro se isso revela ou não um efeito causal, mas isso é crível
- Se $partA = 50$, prevê-se que $votoA$ será cerca de 50, ou metade da votação