



DCP098

Fundamentos para Avaliação Quantitativa de Políticas Públicas

Heteroscedasticidade.

Aula 19
06 de junho de 2022

Ana Paula Karruz

Para que o estimador MQO seja **BLUE**, é preciso atender às chamadas “premissas clássicas”

O que eventualmente ficou de fora não é correlacionado com as variáveis incluídas – se for, teremos o problema do viés de variável omitida

- O modelo de regressão é **linear nos parâmetros** (i.e., coeficientes são adicionados e com expoente = 1)

- ε , o termo de erro aleatório, tem **média populacional = 0**
- Não há correlação entre cada uma das variáveis explicativas e o termo de erro – esta é a premissa de **exogeneidade: correlação $(X_j, \varepsilon) = 0$**

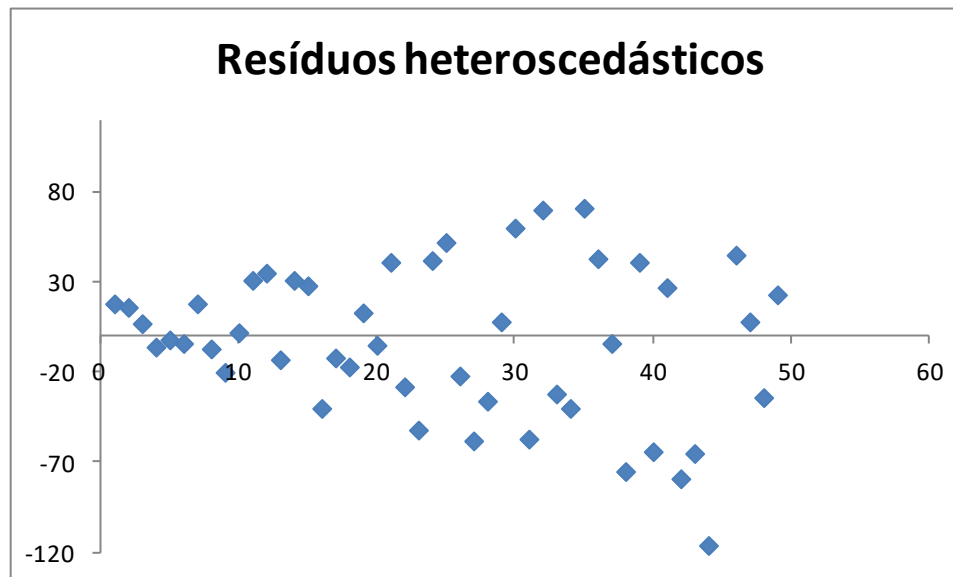
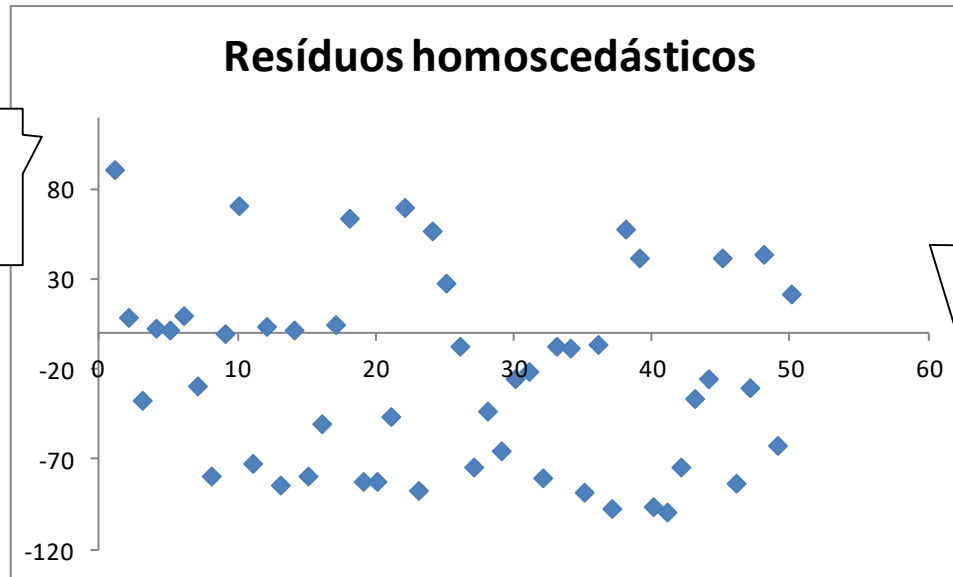
Premissa da média condicional zero é resultado destas duas premissas

- As variáveis explicativas não são uma função linear das outras (i.e., **não há multicolinearidade**)
- O termo de erro tem **variância constante** (i.e., **é não há heteroscedasticidade**)
- Os erros não são correlacionados entre si (i.e., **não existe autocorrelação** serial ou espacial)
- O **erro** tem distribuição **normal**: $\varepsilon \sim N(0, \sigma^2)$

Premissa requerida para teste de hipótese e intervalo de confiança, não para estimação de MQO

Ilustração: Homoscedasticidade vs. Heteroscedasticidade

Eixo vertical
representa
os resíduos



Heteroscedasticidade

Pode ocorrer em regressões simples ou múltiplas

A heteroscedasticidade é geralmente causada por variáveis que assumem valores altos, as quais se associam com erros altos. Exemplos: Pessoas de alta renda podem ter maior variância em seu consumo; escolas grandes podem ter maior variância na nota dos estudantes em testes de desempenho

Implicação	Deteção	Mitigação
<ul style="list-style-type: none">▪ Tende a subestimar erro padrão, portanto:<ul style="list-style-type: none">– Tende a aumentar artificialmente a precisão das estimativas (e a estreitar os intervalos de confiança)– Pode induzir-nos a “encontrar” significância estatística onde ela não existe	<ul style="list-style-type: none">▪ Examine os resíduos plotados contra a variável suspeita de causar heteroscedasticidade – esta abordagem só funciona em casos óbvios/ dramáticos▪ Use um teste estatístico em que a H_0 = não existe heteroscedasticidade (e.g., teste Breusch-Pagan, Park, White)	<ul style="list-style-type: none">▪ Use erros padrão robustos (heteroscedasticity-consistent standard errors)▪ Transforme as variáveis que possam estar causando heteroscedasticidade (e.g., logaritmo, per capita, razão)

Heteroscedasticidade não causa viés, porém requer que cálculo do $EP(\hat{\beta})$ seja ajustado

Calculando erros padrão robustos à heteroscedasticidade (exemplo com dados womenlabor.csv)



- Estime a regressão (como de costume); por exemplo:
`spec1 <- lm(my_formula, data = womenlabor)`
- Calcule erros padrão robustos à heteroscedasticidade
`coeftest(spec1, vcov = vcovHC(spec1, "HC1"))`

`vcov` matriz de variância-covariância do estimador de MQO; em sua forma matricial, OLS calcula os erros padrão a partir dessa matriz, cuja diagonal principal contém $\text{var}(\hat{\beta}_j)$

$$\begin{bmatrix} \text{var}(\hat{\beta}_1) & \text{cov}(\hat{\beta}_1, \hat{\beta}_2) & \dots & \text{cov}(\hat{\beta}_1, \hat{\beta}_k) \\ \text{cov}(\hat{\beta}_2, \hat{\beta}_1) & \text{var}(\hat{\beta}_2) & \dots & \text{cov}(\hat{\beta}_2, \hat{\beta}_k) \\ \vdots & \vdots & \vdots & \vdots \\ \text{cov}(\hat{\beta}_k, \hat{\beta}_1) & \text{cov}(\hat{\beta}_k, \hat{\beta}_2) & \dots & \text{var}(\hat{\beta}_k) \end{bmatrix}$$

`vcovHC` Função que calcula a `vcov` para erros padrão robustos (heteroscedasticity-consistent standard errors)

`type` Tipo de ajuste aplicado; vide documentação da função `vcovHC` no pacote "sandwich"

- Considere os erros padrão robustos à heteroscedasticidade para os testes de significância de β_0 e $\hat{\beta}_j$



DCP098

Fundamentos para Avaliação Quantitativa de Políticas Públicas

Heteroscedasticidade.

Aula 19
06 de junho de 2022

Ana Paula Karruz