

# [Lista 5] Estudo Dirigido

Giovanna Quintão

Abril 2020

Esta lista contempla a primeira parte do conteúdo do curso (Heterocedasticidade e Problemas adicionais de Especificações de Dados). As questões seguem a mesma ordem do livro do Wooldridge e servem como guia para prova, uma vez que abrangem os principais aspectos da matéria. Para um bom estudo, tente fazer da maneira mais completa possível.

As questões extras não precisam ser entregues, mas é útil para consolidar os conhecimentos. A lista valerá ponto e deverá ser entregue até dia **22/04 (Quarta-feira) às 23:59**

## 8 Heterocedasticidade

### 8.1 Consequências da Heterocedasticidade para o método MQO

1. Explique as consequências da heterocedasticidade para a estimação de MQO. O que acontece com a eficiência do estimador e a estatísticas de teste no caso de heterocedasticidade?

### 8.2 Inferência Robusta em Relação à Heterocedasticidade após a estimação MQO

2. Explique porque a seguinte fórmula da variância do estimador:  $var(\hat{\beta}) = \frac{\sigma^2}{SQT_x}$  não é mais válida sob heterocedasticidade. Dica: volte na derivação da fórmula na página 53 (Capítulo 2)
3. Qual é o problema do erro padrão robusto em amostras pequenas?
4. Qual é a importância de calcular a estatística de teste LM robusta em relação a heterocedasticidade?

### 8.3 O teste da existência de heterocedasticidade

5. Explique o sentido de testar, em um teste para heterocedasticidade, a hipótese nula

$$H_0 : \delta_1 = \delta_2 = \dots = \delta_k = 0$$

em que  $\delta_k$  são os parâmetros da seguinte função linear:

$$u^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \dots \delta_k x_k + v$$

6. Explique a diferença entre o teste de Breusch-Pagan e o teste de White.

### 8.4 Estimação de Mínimos Quadrados Ponderados

7. (Wooldridge Adaptado Q8.2) Considere um modelo linear para explicar o consumo mensal de cerveja, que satisfaz as 4 primeiras hipóteses de Gauss Markov e  $u_i$  tem distribuição normal:

$$cerveja = \beta_0 + \beta_1 renda + \beta_2 preco + \beta_3 educ + \beta_4 feminino + u$$

$$E(u \mid renda, preco, educ, feminino) = 0 \quad var(u \mid renda, preco, educ, feminino) = \sigma^2 renda^2$$

- Escreva a equação transformada que tenha um termo de erro homocedástico.
  - Esta equação **transformada** satisfaz as hipóteses do modelo linear clássico?
  - Se estimarmos a equação **transformada** por MQO obtemos estimadores eficientes?
8. Porque os estimadores de Mínimos Quadrados Generalizados (MQG) para correção de heterocedasticidade são também chamados de estimadores de Mínimos Quadrados Ponderados (MQP)?
9. Explique um caso em que os pesos para estimação de MQP surgem “naturalmente” de um modelo econométrico subjacente. *Dica: lembre-se dos casos que temos apenas dados em nível agregado, mas existe heterocedasticidade em nível individual.*
10. Explique o que é o método de Mínimos Quadrados Generalizados Factível.
11. Explique quais são as consequências de usar Mínimos Quadrados Ponderados (MQP) como uma função de variância mal especificada.

### 8.5 O modelo de probabilidade linear revisitado

12. Explique porque no caso de um Modelo de Probabilidade Linear com heterocedasticidade não podemos prosseguir com o método de Mínimos Quadrados Ponderados diretamente.

## 8.6 Questões Extras

13. (Anpec 2006 Q.9). Julgue os itens como verdadeiro ou falso.

O método dos Mínimos Quadrados Ordinários foi empregado para estimar o modelo de regressão abaixo, cujo objetivo é explicar as variações de renda entre 526 indivíduos de uma amostra aleatória:

$$\ln(\text{renda}) = 0,362 + 0,094\text{educ} + 0,014\text{exper} - 0,178\text{sexo} - 0,010\text{exper} * \text{sexo} + u$$

(0,128)            (0,008)(0,002)            (0,058)(0,002)

$$R^2 = 0,368 \quad n = 526$$

Em que sexo é uma variável dicotômica (valor 1, se for mulher, e 0, caso contrário), educ é o número de anos de escolaridade (0 educ 17), exper são anos de experiência profissional (0 exper 40) e u é a estimativa do erro. Os números entre parênteses são os erros padrões das estimativas, robustos à heterocedasticidade. Com base nos resultados acima, é correto afirmar:

- a) Ao nível de significância de 5%, o efeito de um ano a mais de experiência profissional para indivíduos do sexo masculino é estatisticamente maior do que o efeito para mulheres.
- b) Para um indivíduo com 10 anos de escolaridade, 1 ano adicional de estudo acarreta um aumento da renda de aproximadamente 9%.
- c) O efeito na renda de um aumento de 1 ano na experiência profissional para as mulheres é 1% menor do que para os homens.
- d) Pela inspeção dos resultados da estimação fica claro que os erros do modelo são heterocedásticos.
- e) Se a um nível de significância de 5%, o valor crítico do teste F para a regressão for 2,37, os coeficientes angulares serão conjuntamente diferentes de zero.

14. (Anpec 2007- Q8) Julgue as afirmativas:

- a) Heterocedasticidade ocorre quando o erro aleatório em um modelo de regressão é correlacionado com uma das variáveis explicativas.
- b) Quando o erro aleatório em um modelo de regressão é correlacionado com alguma variável explicativa, os estimadores de mínimos quadrados não são consistentes.
- c) Na presença de heterocedasticidade, estimadores de Mínimos Quadrados Ordinários são inefficientes.
- d) Os testes t e F usuais não são válidos na presença de heterocedasticidade.

- e) Na presença de heterocedasticidade, estimadores de Mínimos Quadrados Ordinários são não viesados, mas são inconsistentes.

## 9 Problemas adicionais de especificação de dados

### 9.1 Má especificação da forma funcional

1. Explique o que é a má-especificação da forma funcional. Explique também as consequências da má-especificação para a inferência.
2. Escreva a hipótese nula do teste RESET, e explique a ideia do teste.
3. Suponhamos que queremos testar o modelo

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u \quad \text{contra}$$

$$y = \beta_0 + \beta_1 \log(x_1) + \beta_2 \log(x_2) + u$$

Como poderemos testar?

### 9.2 Utilizando variáveis proxy para variáveis explicativas não observadas

4. Quais hipóteses precisamos para assegurar que a estimação com uma variável *proxy* gere estimativas consistentes dos estimadores?
5. Explique qual é o sentido de incluir a variável de crime defasado em uma regressão de crime como função de gastos com segurança pública.

### 9.4 Propriedades do Método MQO quando há erros de medida

6. Seja  $e_0 = y - y^*$  um erro de medida na variável dependente.
  - a) Quais são as consequências para os estimadores se o erro de medida não for correlacionado com as variáveis independentes?
  - b) E se o erro de medida for sistematicamente correlacionado com alguma variável dependente?
7. (Mesma questão lista 3) Seja :

$$y = \beta_0 + \beta_1 x_1^* + u \quad \text{onde } x_1^* \text{ é não observada}$$

O erro de medida é tal que  $e_1 = x_1 - x_1^*$ , e assumimos que  $E(e_1) = 0$

- a) Explique porque se  $cov(x_1, e_1) = 0$ , o estimador de MQO continua sendo consistente.
- b) O que é o erro clássico nas variáveis? Mostre que no caso de erro clássico o estimador de MQO é viesado.
- c) Nesse caso de erro clássico, o estimador de MQO é consistente? Explique também o que é o viés de atenuação.

## 9.5 Ausência de Dados, Amostras não aleatórias e observações extremas

- 8. Explique qual tipo de seleção de amostra pode ser um problema para a estimação por MQO.
- 9. Explique como devemos proceder quando tivermos observações extremas (*outliers*) em nossa amostra.

## 9.6 Questão extra

- 10. (Anpec 2005 Q.10) Julgue os itens como verdadeiro ou falso. A respeito do modelo de regressão múltipla:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + e_i$$

em que  $e_i$  tem média zero e variância  $\sigma^2$ , são corretas as afirmativas:

- a) Se os erros são heterocedásticos, ainda assim os testes usuais t e F podem, sem prejuízo algum, ser empregados para se testar a significância dos parâmetros do modelo, caso estes sejam estimados por Mínimos Quadrados Ordinários.
- b) Erros de medida da variável dependente reduzem as variâncias dos estimadores de Mínimos Quadrados Ordinários de  $\hat{\beta}_1$  e  $\hat{\beta}_2$ .
- c) A omissão da variável explicativa relevante,  $x_2$ , para explicar a variável dependente,  $y_i$ , torna a estimativa dos coeficientes  $\beta_0$  e  $\beta_1$  tendenciosa e inconsistente, se somente se, a variável omitida  $X_2$ , for correlacionada com a variável incluída,  $X_1$ .