Lista 3

Exercício 1. *Suponha que $y = a + bx + \varepsilon$ satisfaça as hipóteses de Gauss-Markov. Você tem uma amostra de 500 indivíduos e sabe que*

$$\sum X_i = 3000, \quad \sum Y_i = 1400, \quad \sum X_iY_i = 18000, \quad \sum X_i^2 = 66000, \quad e \quad \sum Y_i^2 = 7200$$

a) Ache b^{MQO}

b) Você é capaz de obter o R^2 esse caso? E o S^2?

c) Como você testaria $H_0 : b < 0,5$?

Resposta:

a) Note que

$$\hat{\beta} = (X'X)^{-1}X'Y \quad ; \quad X = \begin{bmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_N \end{bmatrix} \quad \Longrightarrow \quad X'X = \begin{bmatrix} N & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix}$$

$$\begin{aligned} (X'X)^{-1} &= \frac{1}{N\sum X_i^2 - (\sum X_i)^2} \begin{bmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & N \end{bmatrix} \\ &= \frac{1}{500.66000 - (3000)^2} \begin{bmatrix} 66000 & -3000 \\ -3000 & 500 \end{bmatrix} \\ X'Y &= \begin{bmatrix} \sum Y_i \\ \sum X_iY_i \end{bmatrix} = \begin{bmatrix} 1400 \\ 18000 \end{bmatrix} \end{aligned}$$

Logo,

$$\hat{\beta} = \begin{bmatrix} 1,60 \\ 0,20 \end{bmatrix}$$

b)

$$\begin{aligned} \hat{V}(Y) &= \frac{7200}{500} - \left(\frac{1400}{500}\right)^2 = 14,4 - 7,84 = 6,56 \\ \hat{V}(\hat{Y}) &= \hat{\beta}^2\hat{V}(X) = 0,2^2 \left[\frac{66000}{500} - \left(\frac{3000}{500}\right)^2\right] = 3,84 \end{aligned}$$

Logo,

$$R^2 = \frac{\hat{V}(\hat{Y})}{\hat{V}(Y)} = \frac{3,84}{6,56} = 0,5854$$

$$\hat{V}(Y) = \hat{V}(\hat{Y}) + \hat{V}(\hat{\varepsilon})$$
$$\hat{V}(\hat{\varepsilon}) = \hat{\sigma}^2 = 6,56 - 3,84 = 2,72$$

Então,

$$S^2 = \frac{N}{N-2}\hat{\sigma}^2 = \frac{500}{498}2,72 = 2,73$$

c) (estou vendo a resposta com o professor)

□

Exercício 2. *Suponha que um pesquisador regrida a nota média de uma amostra de 100 salas de aula de 9º ano contra o tamanho da sala. O resultado foi*

$$Y = \underset{(20,4)}{520,4} - \underset{(2,21)}{5,82}T + \varepsilon,$$

onde em parênteses há o erro padrão. A média do tamanho de turma é $21,4$ *e* $S^2 = 1$. *Você sabe ainda que* $\sum X_i^2 = 215$.

a) Se uma sala tem 22 estudantes, qual sua predição para a nota média da turma? Qual o intervalo de confiança (95%) em torno dessa predição?

b) Construa um intervalo de confiança de 95% para b_1, *a inclinação de sua reta de regressão.*

c) Calcule o p-valor para um teste de $H_0 : b_1 = 0$. *Você rejeita a nula a 5%? A 1%?*

d) Qual a média de Y*?*

Resposta:

a) Valor predito para turma com $T = 22$ alunos:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1.T \ = \ 520,4 - 5,82.22 \ = \ 329,36$$

Intervalo de confiança:

$$IC_{95\%} = \hat{Y} \pm t_{(98)}^{95\%}\sqrt{S^2}\sqrt{\frac{1}{N} + \frac{(X-\bar{X})^2}{\sum(X-\bar{X})^2}}$$
$$= 329,36 \pm 1,98\left(1\sqrt{\frac{1}{100} + \frac{0,6^2}{1}}\right)$$
$$= [328,16 \ ; \ 330,56]$$

b)

$$\frac{\hat{\beta}_1 - \beta_1}{S_{b_1}} \sim t_{N-2}$$

Então,

$$IC_{5\%} \ = \ \hat{\beta}_1 \ \pm \ t_{(N-2;\ 97,5\%)} \ . \ S_{b_1} \ = \ [-10,24 \ ; \ -1,4]$$

c)

$$t^* = \frac{\hat{\beta}_1}{S_{b_1}} = \frac{-5,82}{2,21} = -2,633$$

Logo,

$$Pr(\tau_{98} < -2,633) = 0,016 \implies \text{p-valor} = \frac{0,016}{2} = 0,008$$

Então, rejeitamos $H_0 : \beta_1 = 0$ para $\alpha = 0,05$ e $\alpha = 0,01$.

d) Note que $\bar{T} = 21,4$, pois sabemos que $\hat{Y}(\bar{T}) = \bar{Y}$ (a reta de regressão passa por $\bar{Y}, \bar{T}$). Portanto,

$$\bar{Y} = \hat{Y}(\bar{T}) = 520,4 - 5,82 \times 21,4 = 395,85$$

□

Exercício 3. *Considere o modelo $\boldsymbol{y}_i = \boldsymbol{x}'_{1i}\boldsymbol{\beta} + \boldsymbol{x}'_{2i}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_i$, onde $\boldsymbol{x}_1$ é um vetor contendo K elementos, $\boldsymbol{x}_2$ contém M elementos, e todas as variáveis já estão em desvios com respeito à média (por exemplo, $y_i = y_{0i} - \bar{y}_0$, para alguma variável primária y_0), de modo que o modelo não contém intercepto. Suponha que $E[\varepsilon|x_1, x_2] = 0$, e que $E[\varepsilon^2|x_1, x_2] = \sigma^2 I$.*

a) Suponha que $\boldsymbol{x}'_1\boldsymbol{x}'_2 = 0$. Mostre que nesse caso os estimadores $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'_1\boldsymbol{X}_1)^{-1}\boldsymbol{X}'_1\boldsymbol{y}$, e $\hat{\boldsymbol{\gamma}} = (\boldsymbol{X}'_2\boldsymbol{X}_2)^{-1}\boldsymbol{X}'_2\boldsymbol{y}$ são ambos os melhores estimadores lineares não-viesados para $(\boldsymbol{\beta}, \boldsymbol{\gamma})$.

b) Caso $\boldsymbol{x}'_1\boldsymbol{x}_2 \neq 0$, sua resposta acima se mantém? Justifique e caracterize os respectivos vieses caso os encontre.

c) Novamente suponha que $\boldsymbol{x}'_1\boldsymbol{x}_2 \neq 0$. Seja $\boldsymbol{M}_2 = \boldsymbol{I} - \boldsymbol{x}_2(\boldsymbol{x}'_2\boldsymbol{x}_2)^{-1}\boldsymbol{x}'_2$. Qual a interpretação para M_2? Mostre que sua estimativa de MQO para β_1 pode ser escrita como: $\beta_1 = (\boldsymbol{X}'_1\boldsymbol{M}_2\boldsymbol{X}_1)^{-1}\boldsymbol{X}'_1\boldsymbol{M}_2\boldsymbol{y}$. Compute a matriz de variâncias e covariâncias de $\hat{\beta}_1$.

d) Ainda no caso anterior, e particularizando para o caso onde x_1 e x_2 contenham apenas uma variável cada, expresse a variância de $\hat{\beta}_1$ como função da correlação entre x_1 e x_2, r_{12}. O que acontece quando $r_{12} \to 1$?

Resposta: Sejam:

$$X'X = \begin{bmatrix} x'_1x_1 & x'_1x_2 \\ x'_2x_1 & x'_2x_2 \end{bmatrix} \qquad \text{e} \qquad \hat{\theta} = \begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix}$$

Na derivação de MQO, vimos que

$$(X'X)\hat{\theta} = X'y$$

$$\begin{bmatrix} x'_1x_1 & x'_1x_2 \\ x'_2x_1 & x'_2x_2 \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} x'_1y \\ x'_2y \end{bmatrix}$$

que pode ser reescrito como

$$\begin{cases} x'_1x_1\hat{\beta} + x'_1x_2\hat{\gamma} &= x'_1y \\ x'_2x_1\hat{\beta} + x'_2x_2\hat{\gamma} &= x'_2y \end{cases} \qquad (*)$$

Resolvendo para $\hat{\beta}$ segue que

$$\hat{\beta} = (x'_1x_1)^{-1}x'_1y - (x'_1x_1)^{-1}x'_1x_2\hat{\gamma}$$

a) $x_1'x_2' = 0 \implies \hat{\beta} = (x_1'x_1)x_1'y$. Como temos, pelo enunciado, o ambiente de Gauss-Markov satisfeito, então $\hat{\theta}' = [\hat{\beta} \quad \hat{\gamma}]$ é BLUE.

Sob $X'X = 0$, $\tilde{\beta} = \hat{\beta} \implies \tilde{\beta}$ é BLUE (para $\hat{\gamma}$, a demonstração é igual)

b) Usando $(*)$ temos:

$$\begin{cases} \hat{\beta} = (x_1'x_1)^{-1}x_1'y - (x_1'x_1)^{-1}x_1'x_2\hat{\gamma} \\ \hat{\gamma} = (x_2'x_2)^{-1}x_2'y - (x_2'x_2)^{-1}x_2'x_1\hat{\beta} \end{cases}$$

Logo,

$$\begin{aligned} \hat{\beta} &= (x_1'x_1)^{-1}x_1'y - (x_1'x_1)^{-1}x_1'x_2 \;.\; [(x_2'x_2)^{-1}x_2'y - (x_2'x_2)^{-1}x_2'x_1\hat{\beta}] \\ &= \left[(x_1'x_1)^{-1}x_1' - (x_1'x_1)^{-1}x_1'x_2(x_2'x_2)^{-1}x_2'\right] y + (x_1'x_1)^{-1}x_1'x_2(x_2'x_2)^{-1}x_2'x_1\hat{\beta} \end{aligned}$$

$$\begin{aligned} \hat{\beta} - (x_1'x_1)^{-1}x_1'x_2(x_2'x_2)^{-1}x_2'x_1\hat{\beta} &= (x_1'x_1)^{-1}x_1'\left[I - x_2(x_2'x_2)^{-1}x_2'\right] y \\ (x_1'x_1)[\hat{\beta} - (x_1'x_1)^{-1}x_1'x_2(x_2'x_2)^{-1}x_2'x_1\hat{\beta}] &= (x_1'x_1)(x_1'x_1)^{-1}x_1'\left[I - x_2(x_2'x_2)^{-1}x_2'\right] y \\ x_1'x_1\hat{\beta} - x_1'x_2(x_2'x_2)^{-1}x_2'x_1\hat{\beta} &= x_1'\left[I - x_2(x_2'x_2)^{-1}x_2'\right] y \\ x_1'\left[I - x_2(x_2'x_2)^{-1}x_2'\right] x_1\hat{\beta} &= x_1'\left[I - x_2(x_2'x_2)^{-1}x_2'\right] y \\ x_1'M_2x_1\hat{\beta} &= x_1'M_2y \\ \hat{\beta} &= (x_1'M_2x_1)^{-1}x_1'M_2y, \end{aligned}$$

em que $M_2 = I - x_2(x_2'x_2)^{-1}x_2'$ é a matriz de projeção no núcleo do espaço-coluna de X_2. Perceba, portanto, que $\hat{\beta}^{MQO}$ explora a variação em X_1 após a remoção da variação em X_2.

Seja $\tilde{x}_1 \equiv M_2x_1$ e $\tilde{y} \equiv M_2y$, então $\hat{\beta}^{MQO} = (\tilde{x}_1'\tilde{x}_1)^{-1}\tilde{x}_1'\tilde{y}$.

Logo,

$$\begin{aligned} \hat{\beta} &= (x_1'M_2x_1)^{-1}x_1'M_2y \;=\; \tilde{\beta} - (x_1'x_1)^{-1}x_1'x_2\hat{\gamma} \\ \tilde{\beta} &= (x_1'x_1)^{-1}x_1'y \end{aligned}$$

Como inicialmente Gauss-Markov é satisfeito, temos $E(\hat{\beta}) = \beta$ e $E(\hat{\gamma}) = \gamma$. Logo,

$$E(\tilde{\beta}|X) = \beta + (x_1'x_1)^{-1}x_1'x_2\gamma$$

em que $(x_1'x_1)^{-1}x_1'x_2\gamma$ é o viés de variável omitida.

c) $cov(\hat{\beta}) = cov[(x_1'M_2x_1)^{-1}x_1'M_2\varepsilon] = \sigma^2(x_1'M_2x_1)^{-1}$

d)

$$\begin{aligned} V(\hat{\beta}) &= \sigma^2\left[x_1'x_1 - \frac{(x_1'x_2)^2}{x_2'x_2}\right]^{-1} \\ &= \sigma^2\left[x_1'x_1\left(1 - \frac{(x_1'x_2)^2}{(x_1'x_1)(x_2'x_2)}\right)\right]^{-1} \\ &= \sigma^2(x_1'x_1)^{-1}(1-\rho_{12}^2)^{-1} \\ &= \frac{\sigma^2}{NV(x_1)(1-\rho_{12}^2)} \end{aligned}$$

Logo,

$$\hat{V}(\hat{\beta}) = \frac{\hat{\sigma}^2}{N\hat{V}(x_1)(1-r_{12}^2)} \qquad \text{e} \qquad \lim_{|\rho_{12}|\to 1} V(\hat{\beta}) = +\infty$$

□

Exercício 4. *Suponha que seu modelo verdadeiro seja* $\boldsymbol{y}_i = \boldsymbol{x}_i\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i$*. No entanto, o econometrista estimou* $\boldsymbol{y}_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \boldsymbol{z}_i'\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_i$*. Sejam* $\hat{\boldsymbol{\beta}}$ *e* $\tilde{\boldsymbol{\beta}}$ *os estimadores de MQO para* β *no modelo correto e modelo mal especificado, respectivamente. Considere ainda que* $E[\varepsilon|\boldsymbol{X},\boldsymbol{Z}] = 0$ *e* $E[\varepsilon\varepsilon'|\boldsymbol{X},\boldsymbol{Z}] = \sigma^2\boldsymbol{I}$*.*

a) $\tilde{\beta}$ *é viesado? Justifique*

b) Compute a matriz de variâncias e covariâncias para $[\tilde{\boldsymbol{\beta}} \quad \tilde{\boldsymbol{\gamma}}]'$*. Compare o bloco correspondente a* $\tilde{\boldsymbol{\beta}}$ *com matriz de variâncias e covariâncias de* $\hat{\boldsymbol{\beta}}$*. Quando elas são idênticas?*

c) Se você souber apenas que $E[\varepsilon|\boldsymbol{X}] = 0$*, como sua resposta (a) se altera?*

Resposta:

a) Retornando ao raciocínio do exercício anterior:

$$\begin{aligned}
\tilde{\beta} &= \beta + (x'M_z x)^{-1}x'M_z\varepsilon \\
E(\tilde{\beta}) &= E_{X,Z}\left[E(\tilde{\beta}|X,Z)\right] \\
&= \beta + E_{X,Z}\left[E\left((x'M_z x)^{-1}x'M_z\varepsilon\right)\right] \\
&= \beta + E_{X,Z}[(x'M_z x)^{-1}x'M_z \cancelto{0}{E(\varepsilon|X,Z)}] = \beta
\end{aligned}$$

Estimador é não-viesado.

b) $\hat{\beta}$ é mais eficiente do que $\tilde{\beta}$ se e somente se $V(\tilde{\beta}) - V(\hat{\beta})$ for positiva-definida $\iff \sigma^{-2}[V(\tilde{\beta}) - V(\hat{\beta})]$ positiva-definida $\iff (x'M_2x)^{-1} - (x'x)^{-1}$ for positiva-definida

$$\begin{aligned}
V(\hat{\beta}) &= \sigma^2(x'x)^{-1} \\
V(\tilde{\beta}) &= \sigma^2(x'M_z x)^{-1} \\
&= \sigma^2(x'(I-P_z)x)^{-1} \\
&= \sigma^2(x'x - x'P_z x)^{-1}
\end{aligned}$$

Lema: A positiva-definida $\implies C'AC$ positiva-definida, em que C tem posto completo

Prova: $x'Ax > 0$ para todo x não-nulo. Seja $Cx = x$. Logo, $V'C'ACV$ para todo V não-nulo $\implies C'AC$ positiva-definida.

Lema: $A - B$ positiva-definida $\implies C'(A-B)C$ positiva-definida

Lema: $B - I$ positiva-definida e B positiva-definida $\implies I - B^{-1}$ positiva-definida

Prova: $B = C\Lambda C'$; $CC' = I$; $C' = C^{-1}$; $B - I = C(\Lambda - I)C'$;
$B^{-1} = C'^{-1}\Lambda^{-1}C^{-1}$; $I = CC' = (C^{-1})'C^{-1}$

Como $\Lambda - I$ é positiva-definida, todos os auto-valores λ de Λ são maiores que 1, caso contrário haveria algum $V \neq 0$ para o qual $V'(\Lambda - I)V < 0$

$$\Lambda^{-1} = \begin{bmatrix} \lambda_1^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_N^{-1} \end{bmatrix} \quad \Longrightarrow \quad \lambda_i^{-1} < 1, \forall i$$

$\Longrightarrow$ $I - \Lambda$ positiva-definida $\Longrightarrow$ $I - B^{-1}$ idem.

Proposição: $B - A$ positiva-definida $\Longrightarrow$ $A^{-1} - B^{-1}$ positiva-definida.

Prova:

$$\begin{aligned}
B - A \text{ pos-def} &\Longrightarrow A^{-1/2}(B - A)(A^{-1/2})' \text{ pos-def} \\
&\Longrightarrow A^{-1/2}B(A^{-1/2})' - A^{-1/2}A^{1/2}(A^{1/2})'(A^{-1/2})' \text{ pos-def} \\
&\Longrightarrow A^{-1/2}B(A^{-1/2})' - I \text{ pos-def} \\
&\Longrightarrow I - [A^{-1/2}B(A^{-1/2})']^{-1} \text{ pos-def} \\
&\Longrightarrow I - (A^{1/2})'B^{-1}A^{1/2} \text{ pos-def} \\
&\Longrightarrow (A^{1/2})'[(A^{1/2})']^{-1}(A^{1/2})^{-1}A^{-1/2} - (A^{1/2})'B^{-1}A^{1/2} \text{ pos-def} \\
&\Longrightarrow (A^{1/2})'\left([(A^{1/2})']^{-1}(A^{1/2})^{-1} - B^{-1}\right)A^{1/2} \text{ pos-def} \\
&\Longrightarrow (A^{1/2})'\left(A^{-1} - B^{-1}\right)A^{1/2} \text{ pos-def} \\
&\Longrightarrow [(A^{1/2})']^{-1}(A^{1/2})'\left(A^{-1} - B^{-1}\right)A^{1/2}(A^{1/2})^{-1} \text{ pos-def} \\
&\Longrightarrow A^{-1} - B^{-1} \text{ pos-def}
\end{aligned}$$

Finalmente,

$$\begin{aligned}
V(\hat{\beta}) < V(\tilde{\beta}) &\Longleftrightarrow \sigma^2(x'x)^{-1} < \sigma^2(x'M_zx)^{-1} \\
&\Longleftrightarrow (x'x)^{-1} < (x'M_zx)^{-1} \\
&\Longleftrightarrow x'x < x'M_zx = x'(I - P_z)x \\
&\Longleftrightarrow x'x < x'x - x'P_zx \\
&\Longleftrightarrow -x'P_zx < 0 \\
&\Longleftrightarrow x'P_zx \geq 0 \\
&\Longleftrightarrow P_z \geq 0 \qquad \text{(verdade, pois } P_z \text{ é positiva-semidefinida)}
\end{aligned}$$

c) $\tilde{\beta} = \beta + (x'M_zx)^{-1}x'M_z$, em que

- $E[(x'M_zx)^{-1}x'M_z\varepsilon] =$ viés
- $M_z = I - P_z \Longrightarrow$ viés $= 0$, caso $P_zX = 0$.

□

Exercício 5. *Suponha que $y = a + bx + \varepsilon$. Você tem uma base de dados com duas colunas para x, uma onde essa variável é medida em unidades e outra onde é medida em dezenas. Você experimenta estimar seu modelo com ambas as variáveis.*

a) Como os valores obtidos para R^2, S^2, $\hat{b}$, $\hat{a}$ se comparam nos dois exercícios?

b) Como as estatísticas de teste para a hipótese $H_0 : b = 0$ se comparam entre os dois exercícios?

c) Como o p-valor de $\hat{b}$ e o intervalo de confiança de 95% para b se comparam entre os dois exercícios?

Resposta:

a) Seja $x_1 = 10x_0$, então

$$\hat{b}_0 = \frac{\widehat{cov}(y, x_0)}{\widehat{V}(x_0)}$$

$$\hat{b}_1 = \frac{\widehat{cov}(y, 10x_0)}{\widehat{V}(10x_0)} = \frac{10\widehat{cov}(y, x_0)}{10^2\widehat{V}(x_0)} = \frac{\hat{b}_0}{10} \tag{1}$$

$$\hat{a}_1 = \bar{y} - \hat{b}_1\bar{x}_1 = \bar{y} - \frac{\hat{b}_0}{10}10\bar{x}_0 = \bar{y} - \hat{b}_0\bar{x}_0 = \hat{a}_0 \tag{2}$$

Para todo i:

$$\begin{aligned} \hat{\varepsilon}_{1i} &= y_i - \hat{a}_1 - \hat{b}_1 x_{1i} \\ &= y_i - \hat{a}_0 - \frac{\hat{b}_0}{10}10.x_{0i} && \text{(por 1 e 2)} \\ &= y_i - \hat{a}_0 - \hat{b}_0 x_{0i} \\ &= \hat{\varepsilon}_{0i} && (3) \end{aligned}$$

Logo, usando (3),

$$S_1^2 = \frac{1}{N-2}\sum_{i=1}^{N}\varepsilon_{1i}^2 = \frac{1}{N-2}\sum_{i=1}^{N}\varepsilon_{0i}^2 = S_0^2$$

$$R_1^2 = 1 - \frac{\sum_{i=1}^{N}\varepsilon_{1i}^2}{SQT} = 1 - \frac{\sum_{i=1}^{N}\varepsilon_{0i}^2}{SQT} = R_0^2$$

b)

$$var(\hat{b}_1) = \frac{\sigma^2}{\sum_{i=1}^{N}(x_{1i} - \bar{x}_1)^2} = \frac{\sigma^2}{\sum_{i=1}^{N}(10x_{0i} - 10\bar{x}_0)^2} = \frac{\sigma^2}{10^2\sum_{i=1}^{N}(x_{0i} - \bar{x}_0)^2} = \frac{1}{10^2}var(\hat{b}_0)$$

$$SD(\hat{b}_1) = \frac{1}{10}SD(\hat{b}_0)$$

$$t_1 = \frac{\hat{b}_1}{SD(\hat{b}_1)} = \frac{\hat{b}_0/10}{SD(\hat{b}_0)/10} = \frac{\hat{b}_0}{SD(\hat{b}_0)} = t_0$$

c)

$$IC_1^{95\%} = \left[\hat{b}_1 \pm 1,96.SD(\hat{b}_1)\right] = \left[\frac{\hat{b}_0}{10} \pm 1,96.\frac{SD(\hat{b}_0)}{10}\right]$$
$$= \frac{1}{10}\left[\hat{b}_0 \pm 1,96.SD(\hat{b}_0)\right] = \frac{1}{10} IC_0^{95\%}$$

$$\tau_1 = \frac{\hat{b}_1}{SD(\hat{b}_1)} = \frac{\hat{b}_0/10}{SD(\hat{b}_0)/10} = \frac{\hat{b}_0}{SD(\hat{b}_0)} = \tau_0 \sim t_{N-2}$$

Logo, $Pr[\tau_0 < T] = Pr[\tau_1 < T], \forall T$:

$$\implies \begin{cases} Pr(T \geq \tau_0 | H_0) = Pr(T \geq \tau_1 | H_0) \\ Pr(T \leq \tau_0 | H_0) = Pr(T \leq \tau_1 | H_0) \end{cases} \qquad \text{(unicaudal)}$$

ou

$$\implies \begin{cases} 2 \times \min\left[Pr(T \geq \tau_0 | H_0);\ Pr(T \leq \tau_0 | H_0)\right] \\ 2 \times \min\left[Pr(T \geq \tau_1 | H_0);\ Pr(T \leq \tau_1 | H_0)\right] \end{cases} \qquad \text{(bicaudal)}$$

□

Exercício 6. *Considere os seguintes modelos:*

$$y_i = a + bx_i + \epsilon_i \tag{1}$$
$$x_i = \alpha + \beta y_i + \varepsilon_i \tag{2}$$

a) *Compare b^{MQO} e β^{MQO}. O que ocorre se y e x possuírem variância (amostral) unitária nesse caso? Se o sinal de b^{MQO} for positivo, o que você aprende sobre o sinal de β^{MQO}?*

b) *Compare o R^2 dos dois exercícios.*

Resposta:

a)

$$\hat{b} = \frac{\widehat{cov}(X,Y)}{\widehat{V}(X)}; \quad \hat{\beta} = \frac{\widehat{cov}(X,Y)}{\widehat{V}(Y)} \quad \implies \quad \hat{b}\hat{\beta} = \frac{\widehat{cov}^2(X,Y)}{\widehat{V}(X)\widehat{V}(Y)} = \hat{\rho}_{XY}^2$$

em que

$$0 \leq \hat{b}\hat{\beta} \leq 1 \quad \implies \quad |\hat{b}| \leq \frac{1}{|\hat{\beta}|}.$$

Além disso:

$$\text{sign}(\hat{b}) = \text{sign}(\widehat{cov}(X,Y)) = \text{sign}(\hat{\beta})$$

Se $\hat{V}(X) = \hat{V}(Y) = 1$, então $\hat{b} = \hat{\beta}$.

b)

$$
\begin{aligned}
R_1^2 = \frac{SQE}{SQT} &= \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} = \frac{\sum_{i=1}^N [\hat{a} + \hat{b}x_i - \hat{a} - \hat{b}\bar{x}]^2}{\sum_{i=1}^N (y_i - \bar{y})^2} = \hat{b}^2 \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \\
&= \frac{\widehat{cov}^2(X,Y)}{\widehat{V}^2(X)} \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} = \frac{\widehat{cov}^2(X,Y)}{\widehat{V}^2(X)} \frac{\widehat{V}(X)}{\widehat{V}(Y)} = \frac{\widehat{cov}^2(X,Y)}{\widehat{V}(X)\widehat{V}(Y)}
\end{aligned}
$$

Analogamente,

$$
R_2^2 = \frac{\widehat{cov}^2(X,Y)}{\widehat{V}(X)\widehat{V}(Y)} = R_1^2
$$

□

Exercício 7. *Suponha que o modelo verdadeiro seja $y = \alpha + \beta \ln(x) + \varepsilon$, e que $E[\varepsilon|x] = 0$. No entanto, você erroneamente estimou $y = a + bx + u$.*

a) $E[u|x] = 0$? Justifique

b) b^{MQO} é consistente? Justifique.

<u>*Resposta*</u>:

a)

$$
\begin{aligned}
u &= y - a - bx = (\alpha + \beta \ln(x) + \varepsilon) - a - bx \\
&= (\alpha - a) + [\beta \ln(x) - bx] + \varepsilon \\
E[u|x] &= (\alpha - a) + \beta \ln(x) - bx \neq 0
\end{aligned}
$$

b)

$$
\begin{aligned}
\hat{b}^{MQO} &= \frac{\widehat{cov}(Y,X)}{\widehat{V}(X)} = \frac{\widehat{cov}(\alpha + \beta \ln X + \varepsilon,\ X)}{\widehat{V}(X)} = \beta \frac{\widehat{cov}(\ln X, X)}{\widehat{V}(X)} + \frac{\widehat{cov}(\varepsilon, X)}{\widehat{V}(X)} \\
\text{plim}(\hat{b}^{MQO}) &= \beta \frac{\widehat{cov}(\ln X, X)}{\widehat{V}(X)} \neq \beta
\end{aligned}
$$

□

Exercício 8. *Você agora é um analista em uma sala de situação (por exemplo em uma crise de covid), onde novas informações chegam instantaneamente. Sua amostra, portanto, cresce constantemente $(N, N+1, ...)$. Os parâmetros de interesse para responder a uma pergunta com significado econômico são β, em um modelo $y = \boldsymbol{x}'\boldsymbol{\beta} + \varepsilon$, onde $E[\varepsilon|\boldsymbol{X}] = 0$.*

a) Mostre que

$$\hat{\boldsymbol{\beta}}^{MQO}_{(N+1)} = \hat{\boldsymbol{\beta}}_{(N)} + \left[\frac{1}{1+\boldsymbol{x}'_{N+1}(\boldsymbol{X}'_N\boldsymbol{X}_N)^{-1}\boldsymbol{x}_{N+1}}\right](\boldsymbol{X}'_N\boldsymbol{X}_N)^{-1}\boldsymbol{x}_{N+1}\left(y_{N+1} - \boldsymbol{x}'_{N+1}\hat{\boldsymbol{\beta}}^{MQO}_{(N)}\right)$$

onde $\boldsymbol{X}_n$ *representa a matriz* $X = \begin{bmatrix} 1 & x_{11} & \dots & x_{K1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \dots & x_{Kn} \end{bmatrix}$ *para as primeiras n observações, e* x_i *representa o vetor* $[1 \quad x_{1i} \quad \dots \quad x_{Ki}]'$ *para o i-ésimo indivíduo.*

b) Seja $\hat{y}_N(\boldsymbol{x}_0) = \boldsymbol{x}'_0(\boldsymbol{X}'_N\boldsymbol{X}_N)^{-1}\boldsymbol{X}'_N\boldsymbol{y}_N$ *sua melhor previsão linear para y caso* $\boldsymbol{x} = \boldsymbol{x}_0$*, e sua estimação tenha partido de uma amostra original de tamanho N. Mostre que se* $\hat{y}_N(\boldsymbol{x}_{N+1}) = y_{N+1}$*, então* $\hat{\boldsymbol{\beta}}^{MQO}_{(N+1)} = \hat{\boldsymbol{\beta}}^{MQO}_{(N)}$*, e conclua que sua estimação só se altera se novos valores de y não puderem ser perfeitamente preditos a partir das observações anteriores de x e y.*

Resposta: Você precisa saber que $X'X = \sum x_i x_i'$ e conhecer (ou derivar) a fórmula de Sherman-Morrison:

$$[A + xx']^{-1} = A^{-1} - \frac{A^{-1}xx'A^{-1}}{1 + x'A^{-1}x}$$

a) Sejam $\underset{K\times K}{A} \equiv \underset{K\times N}{X'_N}\overset{N\times K}{X_N}$, e $\underset{K\times 1}{x_{N+1}} = [1 \quad x_{1,N+1} \quad \dots \quad x_{K,N+1}]'$ o elemento $N+1$ de X_{N+1}, então

$$\underset{K\times N+1}{X'_{N+1}}\overset{N+1\times K}{X_{N+1}} = \underset{K\times K}{A} + \underset{K\times 1}{x_{N+1}}\underset{1\times K}{x'_{N+1}}$$

Da mesma maneira, temos que

$$\underset{K\times N}{X'_{N+1}}\underset{N+1\times 1}{Y_{N+1}} = \underset{K\times N}{X'_N}\overset{N\times 1}{Y_N} + \underset{K\times 1}{x_{N+1}}\underset{1\times 1}{y_{N+1}},$$

em que y_{N+1} é o elemento $N+1$ do vetor Y_{N+1}.

Logo, os estimadores $\hat{\beta}^{MQO}_{N+1}$ e $\hat{\beta}^{MQO}_{N}$ podem ser escritos como

$$\hat{\beta}_N = (X'_N X_N)^{-1}X'_N Y_N = A^{-1}X'_N Y_N \qquad (A \equiv X'_N X_N)$$

$$\hat{\beta}_{N+1} = (X'_{N+1}X_{N+1})^{-1}X'_{N+1}Y_{N+1} = [A + x_{N+1}x'_{N+1}]^{-1}[X'_N Y_N + x_{N+1}y_{N+1}] \tag{1}$$

$$= \left[A^{-1} - \frac{A^{-1}x_{N+1}x'_{N+1}A^{-1}}{1 + x'_{N+1}A^{-1}x_{N+1}}\right][X'_N Y_N + x_{N+1}y_{N+1}]$$

$$= A^{-1}X'_N Y_N + A^{-1}x_{N+1}y_{N+1} - \frac{A^{-1}x_{N+1}x'_{N+1}A^{-1}}{1 + x'_{N+1}A^{-1}x_{N+1}}[X'_N Y_N + x_{N+1}y_{N+1}] \tag{2}$$

Usando (1) em (2), temos que

$$
\begin{aligned}
\hat{\beta}_{N+1} &= \hat{\beta}_N + A^{-1}x_{N+1}y_{N+1} - \frac{A^{-1}x_{N+1}x'_{N+1}A^{-1}}{1+x'_{N+1}A^{-1}x_{N+1}}\left(X'_N Y_N + x_{N+1}y_{N+1}\right) \\
&= \hat{\beta}_N + \frac{A^{-1}x_{N+1}}{1+x'_{N+1}A^{-1}x_{N+1}}\left[(1+x'_{N+1}A^{-1}x_{N+1})y_{N+1} - x'_{N+1}A^{-1}\left(X'_N Y_N + x_{N+1}y_{N+1}\right)\right] \\
&= \hat{\beta}_N + \frac{A^{-1}x_{N+1}}{1+x'_{N+1}A^{-1}x_{N+1}}[y_{N+1} + x'_{N+1}A^{-1}x_{N+1}y_{N+1} - x'_{N+1}A^{-1}X'_N Y_N - x'_{N+1}A^{-1}x_{N+1}y_{N+1}] \\
&= \hat{\beta}_N + \frac{A^{-1}x_{N+1}}{1+x'_{N+1}A^{-1}x_{N+1}}[y_{N+1} - x'_{N+1}A^{-1}X'_N Y_N] \\
&= \hat{\beta}_N + \frac{A^{-1}x_{N+1}}{1+x'_{N+1}A^{-1}x_{N+1}}[y_{N+1} - x'_{N+1}\hat{\beta}_N] && (\hat{\beta}_N = A^{-1}X'_N Y_N) \\
&= \hat{\beta}_N + \frac{(X'_N X_N)^{-1}x_{N+1}}{1+x'_{N+1}(X'_N X_N)^{-1}x_{N+1}}[y_{N+1} - x'_{N+1}\hat{\beta}_N] && (A \equiv X'_N X_N)
\end{aligned}
$$

b) É fácil ver que, se $y_{N+1} - x'_{N+1}\hat{\beta_N} = 0$, logo

$$\hat{\beta}_{N+1} = \hat{\beta}_N.$$

□

Exercício 9. *Seu modelo é $y = a + bx + \varepsilon$, porém $f(\varepsilon) = \lambda^{\cancel{-1}}e^{-\lambda\varepsilon}$.*

a) Calcule $E[\varepsilon|x]$ e $Var[\varepsilon|x]$ nesse caso.

b) Se você estimar a e b por MQO estes estimadores são consistentes? São viesados? Se sim, calcule qual o viés. Se não, demonstre.

Resposta: Havia erro no enunciado, considere que $f(\varepsilon) = \lambda e^{-\lambda\varepsilon}$

a) Como $\varepsilon \perp\!\!\!\perp x$, então

$$
\begin{aligned}
E[\varepsilon|x] &= E[\varepsilon] = \lambda^{-1} \\
V[\varepsilon|x] &= V[\varepsilon] = \lambda^{-2}
\end{aligned}
$$

b) Note que

$$\hat{b} = b + \frac{\sum(x_i - \bar{x})\varepsilon_i}{\sum(x_i - \bar{x})^2}$$

Logo,

$$
\begin{aligned}
E(\hat{b}|x) &= b + \frac{E[\sum(x_i - \bar{x})\varepsilon_i|x]}{\sum(x_i - \bar{x})^2} \\
&= b + \frac{1}{\sum(x_i - \bar{x})^2}\sum(x_i - \bar{x})E[\varepsilon_i|x] \\
&= b + \frac{1}{\sum(x_i - \bar{x})^2}\cancelto{0}{\sum(x_i - \bar{x})}\lambda^{-1} \\
&= b \implies \hat{b} \text{ é não-viesado}
\end{aligned}
$$

Já para a temos:

$$\hat{a} = \bar{y} - \hat{b}\bar{x} = a + (b - \hat{b})\bar{x} + \bar{\varepsilon}$$

Portanto,

$$\begin{aligned} E[\hat{a}|x] &= a + \cancelto{0}{E[(b-\hat{b})\bar{x}|x]} + E[\bar{\varepsilon}|x] \\ &= a + \lambda^{-1} \implies \hat{a} \text{ é viesado} \end{aligned}$$

□

Exercício 10. *Suponha que você esteja interessado(a) em estimar a proporção de pobres (isto é, a proporção de pessoas com renda y abaixo de uma linha de pobreza, L) no Brasil a partir da PNAD. Ou seja, seu parâmetro de interesse é* $\theta = F_y(L) = Pr(Y < L)$*. Você tem acesso a uma base de dados com observações (iid) de renda para* N *indivíduos.*

a) *Seu primeiro estimador é* $T = \frac{1}{N}\sum_{i=1}^{N} d_i$*, onde* d *é uma variável dummy que indica se a renda está abaixo da linha da pobreza, ou seja,* $d_i = 1 \iff y_i < L$*. Encontre* $E(T)$ *e* $Var(T)$*.* T *é não-viesado?* T *é consistente?*

b) *Qual a distribuição assintótica de* T*?*

c) *Suponha agora que adicionalmente você saiba que* $y \sim N(\mu, \sigma^2)$*, com* σ^2 *conhecida. Nesse caso, você agora sabe que* $\theta = \Phi\left(\frac{L-\mu}{\sigma}\right)$*, onde o símbolo* Φ *denota a função de distribuição Normal-Padrão. Considere o estimador alternativo* $U = \Phi\left(\frac{L-\bar{y}}{\sigma}\right)$*.* U *é não-viesado?* U *é consistente?*

d) *Qual a distribuição assintótica de* U*?*

e) *Qual estimador você prefere? Dica:* $\frac{\phi^2(z))}{\Phi(z)[1-\Phi(z)]} < 0,64$ *para qualquer valor de* z*.*

Resposta:

a)

$$\begin{aligned} E(T) &= E\left[\frac{1}{N}\sum_{i=1}^{N} d_i\right] = \frac{1}{N}\sum_{i=1}^{N} E[d_i] = \frac{1}{N}\sum_{i=1}^{N} Pr(d_i = 1) \\ &= \frac{1}{N}\sum_{i=1}^{N} Pr(y_i < L) = \frac{1}{N}\sum_{i=1}^{N}\theta = \frac{N\theta}{N} = \theta \implies \text{não-viesado} \\ V(T) &= v\left[\frac{1}{N}\sum_{i=1}^{N} d_i\right] = \frac{1}{N^2}\sum_{i=1}^{N} V(d_i) = \frac{1}{N^2}\sum_{i=1}^{N}\theta(1-\theta) = \frac{\theta(1-\theta)}{N} \end{aligned}$$

$plim(T) = E(T) = \theta$ (aplicação direta da LGN)

b)

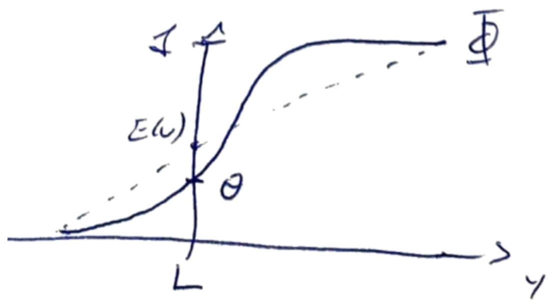
$$\begin{aligned} \sqrt{N}(T-\theta) &\xrightarrow{d} N[0;\ \theta(1-\theta)] \\ T &\xrightarrow{A} N\left[\theta;\ \frac{1}{N}\theta(1-\theta)\right] \end{aligned}$$

c)

$$
\begin{aligned}
plim(U_N) &= plim\left[\Phi\left(\frac{L-\bar{y}}{\sigma}\right)\right] \\
&= \Phi\left(\frac{L-plim(\bar{y})}{\sigma}\right) \qquad\qquad \text{(por Slutsky/mapeamento contínuo)} \\
&= \Phi\left(\frac{L-\mu}{\sigma}\right) = \theta
\end{aligned}
$$

$$
E(U_N) = E\left[\Phi\left(\frac{L-\bar{y}}{\sigma}\right)\right] \gtrless \Phi\left(\frac{L-E(\bar{y})}{\sigma}\right) = \theta, \qquad \text{se } \Phi''\left(\frac{L-E(\bar{y})}{\sigma}\right) \gtrless 0
$$

Em geral, $E(U_N) > \theta$



d) Método Delta:

$$
\begin{aligned}
f(\bar{y}) &\approx f(\mu) + f'(\mu)(\bar{y}-\mu) \\
U_N &\approx \theta + \frac{\partial \Phi}{\partial \mu}(\bar{y}-\mu) \\
U_N - \theta &\approx \frac{\phi}{\sigma^2}\left(\frac{L-\mu}{\sigma}\right)(\bar{y}-\mu) \\
\sqrt{N}(U_N-\theta) &\approx \sqrt{N}\frac{\phi}{\sigma^2}\left(\frac{L-\mu}{\sigma}\right)(\bar{y}-\mu) \overset{A}{\sim} N\left[0;\ \phi^2\left(\frac{L-\mu}{\sigma}\right)\right] \\
U_N &\overset{A}{\sim} N\left[\theta;\ \frac{1}{N}\phi^2\left(\frac{L-\mu}{\sigma}\right)\right]
\end{aligned}
$$

e) U atinge o limite inferior de Cramer-Rao. Outra forma de ver isso é:

<u>Para todo z</u>:

$$
\frac{\phi^2(z)}{\Phi(z)[1-\Phi(z)]} < 0,64
$$

Logo,

$$
\phi^2(z) = 0,64\Phi(z)[1-\Phi(z)] < \Phi(z)[1-\Phi(z)]
$$

Para $z = (L-\mu)/\sigma$:

$$
\phi^2\left(\frac{L-\mu}{\sigma}\right) = Var(U) < \Phi\left(\frac{L-\mu}{\sigma}\right)\left[1-\Phi\left(\frac{L-\mu}{\sigma}\right)\right] = \theta(1-\theta) = Var(T)
$$

$\square$

Exercício 11. *a) Considere o modelo $y = a + \varepsilon$, com $E(\varepsilon) = 0$; $cov(\varepsilon_i, \varepsilon_j) = \delta^2$ e $var(\varepsilon_i) = \sigma^2 + \delta^2$. Calcule os estimadores de MQO e MQG para este caso e suas respectivas variâncias. Compute a eficiência relativa de MQG através de $\frac{var(\hat{a}^{MQG})}{var(\hat{a}^{MQO})}$.*

b) Agora considere o caso em que $cov(\varepsilon_i, \varepsilon_j) = 0$ e $var(\varepsilon_i) = i\sigma^2$. Como fica sua resposta? O que você aprende com estes dois casos?

Resposta:

a)

$$\begin{aligned} var(\varepsilon) &= E(\varepsilon\varepsilon') = \sigma^2 I + \delta^2 \iota\iota' = \Omega \\ \hat{\mu}^{MQO} &= \bar{y} \\ \hat{\mu}^{MQG} &= (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y && \text{(em que } X = \iota) \end{aligned}$$

$$\begin{aligned} var(\hat{y}^{MQO}) &= (X'X)^{-1}X'\Omega X(X'X)^{-1} = \frac{1}{N^2}X'\Omega X = \frac{1}{N^2}\iota'\left[\sigma^2 I + \delta^2\iota\iota'\right]\iota \\ &= \frac{\sigma^2\iota'\iota + \delta^2\iota'\iota\iota'\iota}{N^2} = \frac{\sigma^2 N + \delta^2 N^2}{N^2} = \frac{\sigma^2 + \delta^2 N}{N} \end{aligned}$$

Obtendo Ω^{-1} por chute $\rightarrow \Omega^{-1} = \gamma_0 I + \gamma_1 \iota\iota'$

$$\begin{aligned} \gamma_0 I\sigma^2 I + \gamma_0 I\delta^2\iota\iota' + \gamma_1\iota\iota'\sigma^2 I + \gamma_1\delta^2\iota\iota'\iota\iota' = I \\ \gamma_0\sigma^2 I + \delta^2\gamma_0\iota\iota' + \gamma_1\sigma^2\iota\iota' + N\gamma_1\delta^2\iota\iota' = I \end{aligned}$$

$$\left.\begin{aligned} &\gamma_0\sigma^2 = 1 \\ &\delta^2\gamma_0 + \sigma^2\gamma_1 + N\gamma_1\sigma^2 = 0 \end{aligned}\right\} \quad \begin{aligned} \gamma_0 &= \tfrac{1}{\sigma^2} \\ \gamma_1 &= -\tfrac{\delta^2}{\sigma^2(\sigma^2 + N\delta^2)} \end{aligned}$$

$$\begin{aligned} var(\hat{y}^{MQG}) &= (X'\Omega^{-1}X)^{-1} = \left[\iota'\left(\gamma_0 I + \gamma_1\iota\iota'\right)\iota\right]^{-1} = \left[\gamma_0\iota'\iota + \gamma_1\iota'\iota\iota'\iota\right]^{-1} \\ &= \left(\gamma_0 N + \gamma_1 N^2\right)^{-1} = \left[\frac{N}{\sigma^2} - \frac{\delta^2 N^2}{\sigma^2(\sigma^2 + N\delta^2)}\right]^{-1} \\ &= \left[\frac{N(\sigma^2 + N\delta^2) - \delta^2 N^2}{\sigma^2(\sigma^2 + N\delta^2)}\right]^{-1} = \left[\frac{N^2\delta^2}{\sigma^2(\sigma^2 + N\delta^2)}\right]^{-1} \\ &= \left[\frac{N^2}{\sigma^2 + N\delta^2}\right]^{-1} = \frac{\sigma^2 + N\delta^2}{N^2} = var(\hat{\mu}^{MQO}) \\ \hat{y}^{MQG} &= (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y = [\iota'(\gamma_0 I + \gamma_1\iota\iota')\iota]^{-1}\iota'[\gamma_0 I + \gamma_1\iota\iota'] \\ &= \left[N\gamma_0 + N^2\gamma_1\right]^{-1}\left[\gamma_0\sum_{i=1}^N y_i + N\gamma_1\sum_{i=1}^N y_i\right] = \frac{\gamma_0 + N\gamma_1}{N(\gamma_0 + N\gamma_1)}\sum_{i=1}^N y_i \\ &= \bar{y} = \hat{y}^{MQO} \end{aligned}$$

b) MQO:

$$\hat{a} = \bar{y}$$
$$V(\hat{a}) = \frac{\sigma^2 I}{N^2}\iota'\Omega\iota = \frac{\sigma^2 I}{N^2}\sum_{i=1}^{N} i = \frac{\sigma^2 I}{N^2}\frac{N(N+1)}{2} = \frac{\sigma^2 I(N+1)}{2N}$$

MQG:

$$\hat{a} = \sum_{i=1}^{N} w_i y_i \quad ; \quad w_i = \frac{i^{-1}}{\sum i^{-1}}$$
$$V(\hat{a}) = \sum_{i=1}^{N} V(w_i y_i) = \sum_{i=1}^{N} w_i^2 V(\varepsilon_i) = \sum_{n=1}^{N} w_i^2 \sigma_i^2 = \frac{\sum_{i=1}^{N}\sigma^2}{(\sum_{i=1}^{N} w_i^2)^2} = \frac{N\sigma^2}{(\sum_{i=1}^{N} w_i^2)^2}$$

A correlação de ε não afetou a estimação de $\hat{a}$, mas a heteroscedasticidade sim.

□

Exercício 12. *Considere o modelo já em desvios com respeito à média: $y = ax_1 + bx_2 + \varepsilon$, onde $E[\varepsilon|x_1, x_2] = 0$, porém $E[\varepsilon^2|x_1, x_2] = \alpha x_1^2 + \beta x_2^2$.*

a) Derive a distribuição assintótica de $\hat{\boldsymbol{\gamma}}^{MQO} = [\hat{a} \quad \hat{b}]'$

b) Como você estimaria a matriz de variâncias e covariâncias assintótica de $\hat{\boldsymbol{\gamma}}^{MQO}$?

c) Como você testaria a hipótese $a/b = 1$?

d) Explique como você construiria um estimador $\hat{\boldsymbol{\gamma}}^{MQGf}$, e derive sua distribuição assintótica.

Resposta:

a)

$$\sqrt{N}(\hat{\gamma} - \gamma) \overset{d}{\sim} N\left(0;\ E(xx')^{-1}E(\varepsilon^2 xx')E(xx')^{-1}\right)$$

Como:

$$E(\varepsilon^2 xx') = E_x\left[(\alpha x_1^2 + \beta x_2^2)xx'\right]$$
$$\equiv E(x'Axxx'), \qquad \text{em que } A \equiv \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}$$

Logo,
$$\sqrt{N}(\hat{\gamma} - \gamma) \overset{d}{\sim} N\left(0;\ E(xx')^{-1}E(x'Axxx')E(xx')^{-1}\right)$$

b) $\underline{1^{\underline{o}} \text{ estágio}}$: MQO

$$\hat{\gamma} = (X'X)^{-1}X'y \quad ; \quad \hat{\varepsilon} = y - \hat{\beta}$$

$\underline{2^{\underline{o}} \text{ estágio}}$: $\hat{\varepsilon}_i^2 = \alpha X_{1i}^2 + \beta X_{2i}^2 + u$, logo

$$\text{MQO:} \qquad \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \left((x^2)'x^2\right)^{-1} (x^2)'\hat{\varepsilon}^2$$

$$\hat{\Omega} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\alpha}x_{1i}^2 + \hat{\beta}x_{2i}^2)x_i^{-2}x_i^{-2}$$

c) $H_0 : \frac{a}{b} = 1 \iff H_0 : a = b \iff a - b = 0$

Seja $r = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \implies H_0 : r'\gamma = 0$

Teste z (assintótico): $r'\gamma \sim N(r'\gamma;\ r'V(\hat{\gamma})r)$

$$z^* = \frac{r'\hat{\gamma}}{\sqrt{r'V(\hat{\gamma}r)}} \sim N(0,1)$$

d) $\underline{1^{\underline{o}} \text{ estágio}}$: MQO

$$\hat{\gamma} = (X'X)^{-1}X'y \quad ; \quad \hat{\varepsilon} = y - \hat{\beta}$$

$\underline{2^{\underline{o}} \text{ estágio}}$: $\hat{\varepsilon}_i^2 = \alpha X_{1i}^2 + \beta X_{2i}^2 + u$, logo

$$\text{MQO:} \qquad \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \left((x^2)'x^2\right)^{-1} (x^2)'\hat{\varepsilon}^2$$

$$\hat{\Omega} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\alpha}x_{1i}^2 + \hat{\beta}x_{2i}^2)x_i^{-2}x_i^{-2}$$

$\underline{3^{\underline{o}} \text{ estágio}}$:

$$\hat{\gamma}^{MQGf} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y \overset{a}{\sim} N[\gamma;\ N^{-1}(X'\hat{\Omega}^{-1}X)^{-1}]$$

□

Exercício 13. *Considere o modelo:* $y = c + ax_1 + bx_2 + \varepsilon$*, onde* $E[\varepsilon|x_1, x_2] = 0$*,* $E[\varepsilon^2|x_1, x_2] = \sigma^2$*. O problema agora é que, para* x_1*, as últimas* $N - J$ *observações estão faltando (missing data). Avalie os seguintes três procedimentos:*

a) Descarte as últimas J *observações de sua amostra. Nesse caso MQO ainda seria não-viesado? Justifique*

b) Suponha que x_1 *possa ser modelado através da relação* $x_1 = \boldsymbol{z}'\boldsymbol{\gamma} + \boldsymbol{u}$*, para um conjunto de variáveis* $\boldsymbol{z}$ *disponível para toda a amostra, e onde* $E[u|z] = 0$*;* $E[\boldsymbol{u}\boldsymbol{u}'|Z] = \delta^2\boldsymbol{I}$*, e* $E[u\varepsilon|\boldsymbol{z}, x] = 0$*. Nesse caso, você poderia, nas primeiras* $N - J$ *observações, estimar* $\hat{\boldsymbol{\gamma}}$ *e posteriormente imputar* $\hat{x}_{1i} = \boldsymbol{z}_i'\hat{\boldsymbol{\gamma}}$ *às últimas* J *observações e desse modo completar a base de dados para poder fazer MQO. Este procedimento resulta em estimativas não-viesadas para (a,b,c)? Seriam elas ao menos consistentes?*

c) *No procedimento anterior, há um erro de medida em parte da amostra, $\eta_i = x_{1i} - \hat{x}_{1i}$. No item anterior já deveríamos ter discutido potencial viés, mas além disso há aqui um problema de heterocedasticidade. Desenvolva um estimador MQGf para lidar com este problema. Este procedimento é assintoticamente eficiente?*

<u>*Resposta*</u>:

a) Seja $A = 1$ o evento X presente e $A = 0$ o ausente (*missing*). Então,

$$E(\varepsilon|X) = E(\varepsilon|X, A = 1).Pr(A = 1|X) + E(\varepsilon|X, A = 0).Pr(A = 0|X) = 0$$
$$E(\varepsilon|X, A = 1) = -\frac{Pr(A = 0|X)}{Pr(A = 1|X)} E(\varepsilon|X, A = 0)$$

2 condições para $E(\varepsilon|X, A = 1) = 0$:

- $Pr(A = 0|X) = 0$ (não há atrito)
  ou
- $E(\varepsilon|X, A = 1) = E(\varepsilon|X) \implies A \perp\!\!\!\perp \varepsilon|X$ (*missing at random*)

Se $E(\varepsilon|X, A = 1) = 0$, então MQO é não-viesado.

b) Seja

$$\tilde{X} = \begin{bmatrix} X_J \\ \hat{X}_{N-J} \end{bmatrix} \; ; \quad \hat{X}_i' = \begin{bmatrix} 1 & z_i\hat{\gamma} & X_{2i} \end{bmatrix}$$

$$y = \tilde{X}\beta + v \; ; \quad v = \begin{bmatrix} \varepsilon \\ \varepsilon + (X - \hat{X})\beta \end{bmatrix}$$

Note que

$$E(\varepsilon|X) = 0 \not\Longrightarrow E(v|\tilde{X}) = 0,$$

pois seria necessário $E(\varepsilon|X, Z) = 0$. Logo, $\hat{\beta}$ é, em geral, viesado.

Observe, contudo, que

$$plim(\hat{\gamma}) = \gamma \implies \hat{X} \xrightarrow{p} x - u$$

$$\hat{\beta} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'Y = \left(\frac{1}{N}\sum \tilde{x}_i\tilde{x}_i'\right)^{-1}\left(\frac{1}{N}\sum \tilde{x}_i y_i\right)$$
$$= \begin{bmatrix} \frac{X_J'X_J}{N} & \frac{X_J'\hat{X}_{N-J}}{N} \\ \frac{\hat{X}_{N-J}'X_J}{N} & \frac{\hat{X}_{N-J}'\hat{X}_{N-J}}{N} \end{bmatrix}^{-1} \begin{bmatrix} \frac{X_J'}{N} \\ \frac{\hat{X}_{N-J}'}{N} \end{bmatrix} Y,$$

em que:

$$
\begin{aligned}
plim\left(\frac{X_J'X_J}{N}\right) &= E(x_Jx_J') \\
plim\left(\frac{\hat{X}_{N-J}'X_J}{N}\right) &= plim\left(\frac{(X_{N-J}'+u)'X_J}{N}\right) = E(x_{N-J}x_J') + \cancelto{0}{plim\left(\frac{u'x_J'}{N}\right)} \\
plim\left(\frac{\hat{X}_{N-J}'\hat{X}_{N-J}}{N}\right) &= plim\left(\frac{X_{N-J}'X_{N-J}}{N}\right) + plim\left(\frac{u'u}{N}\right) + 0 = E(x_{N-J}x_{N-J}') + \Sigma_v
\end{aligned}
$$

$$
\Longrightarrow \; plim\left(\frac{\tilde{X}'\tilde{X}}{N}\right) = E(xx') + \Sigma_{vv} \;\neq\; E(xx'),
$$

em que

$$
\Sigma_{vv} = \left[\begin{array}{cc} 0 & 0 \\ 0 & \Sigma_v \end{array}\right]
$$

Logo,

$$
\begin{aligned}
plim\left(\frac{\tilde{X}'\tilde{X}}{N}\right)^{-1} &= \left[E(xx') + \Sigma_{vv}\right]^{-1} \\
plim\left(\frac{\tilde{X}'Y}{N}\right) &= \left[\begin{array}{c} E(x_Jy) \\ plim\left(\frac{\hat{x}y}{N}\right) \end{array}\right] = \left[\begin{array}{c} E(x_Jy) \\ E(x_{N-J}y) \end{array}\right] = E(xy)
\end{aligned}
$$

$$
plim\hat{\beta} = \left[E(xx') + \Sigma_{vv}\right]^{-1} E(xy) \;\neq\; \beta
$$

c) Nas primeiras J observações, podemos calcular

$$
\hat{V} = x - \hat{x} \quad \text{e} \quad S_v^2 = \frac{1}{J - K_z} \quad \Longrightarrow \quad \sum_J \hat{v}^2 = \left[\begin{array}{cc} 0 & 0 \\ 0 & I_{N-J} \end{array}\right] S_v^2
$$

Podemos formar

$$
\tilde{\tilde{X}}'\tilde{\tilde{X}} = \tilde{X}'\tilde{X} - \Sigma_{vv}
$$

logo,

$$
\begin{aligned}
\tilde{\tilde{\beta}} &= (\tilde{\tilde{X}}'\tilde{\tilde{X}})^{-1}\tilde{\tilde{X}}Y \;\xrightarrow{p}\; \beta \\
\tilde{\tilde{\varepsilon}} &= Y - X\tilde{\tilde{\beta}} \\
\hat{\Omega} &= \frac{1}{N}\sum \tilde{\tilde{\varepsilon}}_i^2 \tilde{x}_i \tilde{x}_i^2
\end{aligned}
$$

Assim,

$$
\hat{\beta}^{MQGf} = A\tilde{X}'\hat{\Omega}^{-1}Y, \quad \text{em que } A = \tilde{X}'\hat{\Omega}^{-1}\tilde{X} = \hat{\Sigma}_{vv}
$$

□

Exercício 14. *Suponha que seu modelo seja* $y = a + bx + \varepsilon$*, e que as hipóteses populacionais de Gauss-Markov sejam satisfeitas. No entanto, sua amostra foi coletada em um desenho onde, inicialmente, a população original foi dividida em* J *grupos tal que o tamanho do grupo 1 é 1, o tamanho do grupo 2 é 2 e assim por diante até que o* j*-ésimo grupo tem tamanho* j*, de modo que ao final* $\sum_{j=1}^{J} n_j = N$*. Infelizmente, contudo, você tem acesso somente às médias amostrais de cada grupo* $\bar{y}_j$, $\bar{x}_j$.

a) *Como você lidaria com a situação? Simplesmente formaria uma base de dados com* J *entradas e aplicaria MQO? Este procedimento acarretaria em estimadores consistentes? Qual seria a variância de* $b^{MQO(J)}$ *nesse caso?*

b) *Encontre o estimador de MQGf nesse caso.*

Resposta:

a) A regressão agora fica

$$\bar{y}_j = \bar{x}_j' \beta u_j$$

em que $\bar{y}_j = \frac{1}{J}\sum_{k=1}^{J} y_k^{(j)}$, sendo $y_k^{(j)}$ a k-ésima observação no grupo j.

$$\varepsilon_k \sim N(0, \sigma^2) \implies u_j \sim N\left(0;\ \frac{\sigma^2}{j}\right)$$

Logo,

$$\Omega = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1/2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1/J \end{bmatrix} \implies \Omega^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & J \end{bmatrix}$$

MQO é consistente, porém $Avar(\hat{\beta}^{MQO}) = (\hat{X}'\hat{X})^{-1}\bar{X}'\Omega\bar{X}(\hat{X}'\hat{X})^{-1}$

b) $\hat{\beta}^{MQGf} = (\bar{X}'\Omega^{-1}\bar{X})^{-1}\bar{X}'\Omega^{-1}\bar{y}$

□