

Università di Pavia

Econometria

2008-2009

Esercizi 3

Maggio, 2009

1. Sia $\hat{\beta}$ il vettore delle stime OLS di una regressione di y su \mathbf{X} e \mathbf{c} un qualunque vettore ($K \times 1$) diverso da $\hat{\beta}$. Provate che la differenza nella somma dei residui al quadrato sia

$$(\mathbf{y} - \mathbf{X}\mathbf{c})'(\mathbf{y} - \mathbf{X}\mathbf{c}) - (\mathbf{y} - \mathbf{X}\hat{\beta})'(\mathbf{y} - \mathbf{X}\hat{\beta}) = (\mathbf{c} - \hat{\beta})'\mathbf{X}'\mathbf{X}(\mathbf{c} - \hat{\beta})$$

e che $(\mathbf{c} - \hat{\beta})'\mathbf{X}'\mathbf{X}(\mathbf{c} - \hat{\beta}) > 0$.

2. Considerate la regressione OLS di y su K variabili, \mathbf{X} (inclusa la costante). Considerate un insieme alternativo di regressori, $\mathbf{Z} = \mathbf{X}\mathbf{P}$, dove \mathbf{P} è una matrice non singolare. Così ogni colonna di \mathbf{Z} è una combinazione delle colonne di \mathbf{X} .
- (a) Provate che i vettori dei residui nella regressione di y su \mathbf{X} e di y su \mathbf{Z} sono identici.
- (b) Come possiamo rappresentare il cambio di unità di misura delle variabili indipendenti?
- (c) Cambia il fit della regressione?
3. Supponete che il modello che abbia generato i dati sia:

$$y_t = \beta x_t + \epsilon_t \quad t = 1, \dots, N$$

con $\epsilon_t \sim WN(0, \sigma^2)$ ed x_t non stocastico.

Il modello non ha costante. Confrontate la varianza dello stimatore OLS di β calcolato senza il termine costante nella regressione con quella che si avrebbe se il modello includesse la costante.

4. Considera il modello

$$y_t = \beta_1 + \beta_2 x_{t2} + \epsilon_t \quad t = 1, \dots, N \quad (1)$$

dove \mathbf{X} e ϵ sono indipendenti. Scrivi un codice in GRETL che permetta di valutare attraverso un esercizio Monte Carlo, con S simulazioni, la correttezza degli stimatori OLS di β_1 e β_2 .

Considera i seguenti scenari alternativi per definire il setup del Monte Carlo:

- $N = 50, 100, 500$;
- $S = 100, 1000, 5000$;

- $\epsilon_t \sim iidN(0, 1)$ e $x_{t2} \sim iidN(0, 2)$. Utilizza un *seed* progressivo che parta da 500 per generare gli ϵ_t , mentre considera sempre la stessa X in ciascuna replica del Monte Carlo;
- ϵ_t distribuito come t -Student con media zero ed 1 grado di libertà e $x_{t2} \sim N(0, 2)$. Anche in questo caso, utilizza un *seed* progressivo che parta da 500 per generare ϵ , mentre considera la stessa X in ciascuna replica del Monte Carlo;
- $\beta_1 = 0$ e $\beta_2 = 2$;
- $\beta_1 = 2$ e $\beta_2 = 0.5$.

Quali conclusioni si possono trarre sulla correttezza degli OLS?