Università di Pavia

Econometria

2008-2009

Esercizi 4

Maggio, 2009

Istruzioni: I commenti devono essere concisi!

1. Dato il modello di regressione lineare, con K regressori

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

con $E(\boldsymbol{\epsilon}) = \mathbf{0}$ e $E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}') = \sigma^2\mathbf{I}_N$

   (a) Mostrate che l'R_c^2 si può scrivere come:

$$R_c^2 = \frac{\|\mathbf{M}_\iota\mathbf{X}\widehat{\boldsymbol{\beta}}\|^2}{\|\mathbf{M}_\iota\mathbf{y}\|^2} = \frac{\|\mathbf{P}_X\mathbf{M}_\iota\mathbf{y}\|^2}{\|\mathbf{M}_\iota\mathbf{y}\|^2}$$

   (b) Mostrate che

$$Var(\widehat{\beta}_i|\mathbf{X}) = \frac{\sigma^2}{\sum_{t=1}^N (x_{ti} - \overline{x}_i)^2 \left(1 - \frac{\sum_{t=1}^N (\widehat{x}_{ti} - \overline{x}_i)^2}{\sum_{t=1}^N (x_{ti} - \overline{x}_i)^2}\right)}$$

   dove $\frac{\sum_{t=1}^N (\widehat{x}_{ti} - \overline{x}_i)^2}{\sum_{t=1}^N (x_{ti} - \overline{x}_i)^2}$ è l'R_c^2 di una regressione del generico regressore $\mathbf{x}_i$ sugli altri $K-1$ inclusa la costante. (Suggerimento: Per ottenere $Var(\widehat{\beta}_i|\mathbf{X})$ partite dall'espressione dello stimatore OLS di β_i ottenuto applicando il teorema di FWL).

2. Considerate un modello di regressione con solo due variabili esplicative $\mathbf{x}_1$ and $\mathbf{x}_2$ entrambe centrate:

$$\mathbf{y} = \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \boldsymbol{\epsilon}.$$

dove $\mathbf{x}_1$ and $\mathbf{x}_2$ sono non stocastici e $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_N)$. Con $\widehat{\rho}$ indichiamo il coefficiente di correlazione campionario tra $\mathbf{x}_1$ e $\mathbf{x}_2$. Poichè entrambi sono centrati, la correlazione campionaria è

$$\widehat{\rho} \equiv \frac{\sum_{t=1}^N x_{1t}x_{2t}}{\left((\sum_{t=1}^T x_{1t}^2)(\sum_{t=1}^T x_{2t}^2)\right)^{1/2}}$$

   (a) Mostrate che la correlazione tra le stime OLS $\widehat{\beta}_1$ e $\widehat{\beta}_2$ è uguale a $-\widehat{\rho}$.

(b) Quali, se ve ne sono, delle assunzioni di questo modello possono essere abbandonate senza cambiare il risultato?

3. Nel file *advertising.csv* trovate tre variabili S, le vendite di un certo prodotto, P, il prezzo e A le spese in pubblicità, per ciascuna avete un totale di $N = 75$ osservazioni. Stimate il seguente modello:

$$S_t = \beta_1 + \beta_2 P_t + \beta_3 A_t + \epsilon_t \quad t = 1, \ldots, N$$

assumendo che ϵ_t sia $i.i.d.N(0, \sigma^2)$.

(a) La regressione è significativa (significa testare per $\beta_2 = \beta_3 = 0$)?

(b) Testate l'ipotesi nulla che il prezzo non abbia alcun effetto sulle vendite. Commentate il risultato del test.

(c) Includete tra i regressori A_t^2 che permette di investigare l'esistenza di rendimenti decrescenti nelle spese pubblicitarie. Stimate il nuovo modello e commentate i risultati.

(d) Come misurate l'effetto marginale della pubblicità sulle vendite attese?

(e) Testate l'ipotesi nulla che la pubblicità non abbia effetti sulle vendite.

(f) Trovate il livello ottimo di pubblicità tale che l'ultimo euro speso in pubblicità risulti in un euro di vendite addizionali, sotto l'ipotesi che il costo marginale di produrre e vendere un'altra unità di prodotto sia zero.

(g) Testate l'ipotesi che il livello ottimo sia 1900 euro (A è misurato in migliaia di euro).

4. Scrivi una funzione in GRETL che permetta di determinare le osservazioni influenti in un modello di regressione lineare classico. In particolare, tale funzione dovrà restituire i seguenti output:

- I residui, $\hat{\epsilon}_i$, per ciascuna osservazione;
- Gli elementi sulla diagonale principale della matrice di proiezione $\mathbf{P}_X$;
- Il vettore delle differenze tra le stime OLS di $\boldsymbol{\beta}$ ottenute utlizzando l'intero campione e quelle ottenute omettendo, di volta in volta, ciascuna osservazione;

Utilizzando i dati in ESERCIZIO4.gdt e la funzione appena creata, individua l'osservazione più influente nella seguente relazione lineare

$$producgr_t = \beta_1 + \beta_2 invgdp_t + \epsilon_t$$

dove $producgr$ è la crescita del PIL per lavoratore, mentre $invgdp$ è il rapporto tra il livello degli investimenti in capitale fisso e il PIL.