

Risposta multipla

- Una variabile non inclusa fra i regressori
 - A distorce sempre le stime che si ottengono mediante OLS
 - B distorce le stime che si ottengono se è correlata con la Y
 - C per distorce le stime che otteniamo deve essere correlata con gli altri regressori e non con la Y
 - D distorce i coefficienti che otteniamo se è correlata sia con X che con Y
- quali dei seguenti modelli è utilizzabile per la stima dell'elasticità della domanda (Q) al prezzo (P)
 - A $\log(Q_i) = \beta_0 + \beta_1 P_i + \epsilon$
 - B $\log(Q_i) = \beta_0 + \beta_1 P_i + \beta_2 \log(P_i) + \epsilon$
 - C $\log(Q_i) = \beta_0 + \beta_1 \log(P_i) + \epsilon$
 - D $\log(Q_i) = \beta_0 + \beta_1 P_i + \beta_2 P_i^2 + \epsilon$
- quali delle seguenti formule corrisponde all' R^2 di una regressione?
 - A $R^2 = 1 - \frac{\sum_{i=1}^n u_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
 - B $R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
 - C $R^2 = \left[\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2} \right]^{-1}$
 - D tutte le precedenti sono vere
- Il p -value di una regressione lineare:
 - A nel caso sia negativo segnala una riduzione del valore predetto
 - B indica con che probabilità il coefficiente non è statisticamente diverso da zero
 - C è il valore in probabilità della variabile indipendente
 - D se moltiplicato per β_1 restituisce il livello di confidenza del coefficiente
- Errori eteroschedastici in una regressione:
 - A rendono le stime meno precise ma distorte
 - B riducono la distorsione nella stima dei coefficienti
 - C non distorcono i coefficienti ma rendono meno affidabili gli errori standard
 - D rende i coefficienti di regressione distorti verso il basso
- Nel modello $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, il termine ϵ_i rappresenta:
 - A l'effetto di un aumento unitario di X su Y
 - B l'errore di misurazione nel regressore i
 - C la componente casuale che non può essere spiegata dal modello
 - D la variabilità spiegata dalla variabile dipendente

Risposta multipla

- In una regressione con effetto fisso individuale (dati panel)
 - Il coefficiente dell'effetto fisso individuale è sempre distorto
 - riesce a catturare la variabilità dovuta a variabili omesse se queste non variano nel tempo a livello individuale
 - coglie trend temporali altrimenti inspiegabili
 - distorce i coefficienti che otteniamo se gli effetti fissi sono correlati con Y
- quali dei seguenti modelli è utilizzabile per la stima dell'elasticità della domanda (Q) al prezzo (P)
 - $\log(Q_i) = \beta_0 + \beta_1 \log(P_i) + \epsilon$
 - $\log(Q_i) = \beta_0 + \beta_1 P_i + \epsilon$
 - $\log(Q_i) = \beta_0 + \beta_1 P_i + \beta_2 P_i^2 + \epsilon$
 - $\log(Q_i) = \beta_0 + \beta_1 P_i + \beta_2 \log(P_i) + \epsilon$
- quali delle seguenti formule corrisponde all' R^2 di una regressione?
 - $R^2 = 1 - \frac{\sum_{i=1}^n u_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
 - $R^2 = \frac{\sum_{i=1}^n (\hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
 - $R^2 = \left[\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (\bar{y})^2} \right]^{-1}$
 - tutte le precedenti sono vere
- Il l'errore standard di un coefficiente di una regressione lineare:
 - nel caso sia negativo segnala una riduzione del valore predetto
 - è pari alla probabilità che il coefficiente non sia statisticamente diverso da zero
 - è il valore in probabilità della variabile indipendente
 - è una misura della precisione con cui il coefficiente è stimato
- Errori omoschedastici in una regressione:
 - rendono le stime meno precise ma distorte
 - sono una delle assunzioni necessarie per minimizzare l'errore nella stima dei coefficienti
 - non distorcono i coefficienti ma rendono meno affidabili gli errori standard
 - rende i coefficienti di regressione distorti verso il basso
- Nel modello $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, il termine β_0 rappresenta:
 - l'effetto di un aumento unitario di X su Y
 - l'errore di misurazione nel regressore i
 - il valore atteso di Y quando $X = 0$
 - la componente casuale che non può essere spiegata dal modello

Risposta aperta (scrivere in stampatello solo negli spazi predisposti)

1. Il grafico sotto riporta l'andamento del tasso di cambio fra sterlina e euro fra fine giugno e inizio luglio:



- Quali caratteristiche ha questa serie storica?
In questa serie storica si vede chiaramente la presenza di una rottura strutturale in corrispondenza del referendum sulla Brexit. Si nota anche un trend discendente immediatamente successivo.
- Quali problemi potrebbero sorgere nel caso di previsioni basate su un modello autoregressivo?
La presenza di rotture strutturali e di trend rende inaffidabile l'uso di un modello autoregressivo per fare previsioni riguardo al valore futuro della sterlina.

2. un modello con effetti fissi di area di residenza spiega la probabilità che uno spettatore televisivo si sia sintonizzato su una partita di calcio nella scorsa stagione

variabile	coefficiente	errore standard	t	p - value
costante	0.4398	0.1366	3.2188	0.0012
uomo	0.2313	0.0609	3.7930	0.0002
single	-0.0751	0.0426	-1.7623	0.0782
uomo × single	0.1201	0.0607	1.9775	0.0482
area=Nord	-0.0342	0.0141	-2.4207	0.0159
area=Sud	0.0817	0.0239	3.4213	0.0007

- quali coefficienti sono significativi al 99%?
Il primo (costante), il secondo e l'ultimo.
- come si interpreta la costante?
Il valore atteso se tutti i regressori assumono valore zero. In questo caso la probabilità per una donna sposata/fidanzata che vive al centro Italia
- perchè la variabile "area=centro" non è inclusa nel modello?
L'inclusione porterebbe all'impossibilità di stimare il modello per via della multicollinearità perfetta fra la costante e le tre variabili di residenza. L'effetto di risiedere al centro è catturato dalla costante.

3. Sempre riguardo alla stima del modello dell'esercizio 2

- come si interpreta il coefficiente delle variabili interagite *uomo* e *single*?
E' l'aumento di probabilità di guardare la partita per gli uomini single. Ci mostra che essere single ha un effetto differente sulla probabilità a seconda del sesso.
- cosa catturano i coefficienti fissi per area di residenza?
Gli effetti fissi regionali catturano la variabilità che si associa a tutti quei fattori che sono fissi nel tempo ma variano fra aree geografiche.

4. • Come si può risolvere il problema della presenza da valori difformi (outlier) in un'analisi di regressione?

La sensibilità della stima dei minimi quadrati ad outlier può essere corretta utilizzando stimatori meno sensibili a valori estremi come lo stimatore delle minime deviazioni assolute (LAD).