

Esame Finale

Introduzione all'Econometria

26 maggio 2024

Nome : _____

Cognome: _____

Matricola: _____

Rispondete a tutte le domande negli appositi spazi. Motivare sempre le risposte.
--

Domanda 1

Siete stati assunti da una società di data science per studiare i fattori che spiegano l'accettazione delle domande per ottenere la carta di credito Canicattì Express.

Avete a disposizione un dataset contenente 1319 osservazioni su 7 variabili:

variabile	description
<i>approvata</i>	= 1 se la domanda e' stata accettata
<i>eta</i>	eta' in anni
<i>reddito</i>	reddito annuale (in 10,000 Euro)
<i>spesecc</i>	Spese medie mensili su altre carte di credito (in EURO)
<i>proprietario</i>	= 1 se l'individuo e' proprietario della sua dimora
<i>IVA</i>	= 1 se l'individuo e' titolare di partita IVA
<i>numcc</i>	Numero di carte di credito attive

La prima analisi è modello di probabilità lineare (gli errori standard, in parentesi, sono robusti alla presenza di eteroschedasticità):

$$\begin{aligned}
 \textit{accepted} = & 0.8093 + 0.0305 \log(\textit{reddito}) + 0.1723 \textit{proprietario} - 0.2539 \textit{IVA} \\
 & \quad (0.0927) \quad (0.0488) \quad (0.0689) \quad (0.1175) \\
 & + 0.1420 \log(\textit{reddito}) \times \textit{IVA} - 0.0594 \log(\textit{reddito}) \times \textit{proprietario} \\
 & \quad (0.0647) \quad (0.0579) \\
 & + 0.0027 \textit{numcc} + 0.0002 \textit{spesecc} - 0.0109 \textit{eta} + 0.0001 \textit{eta}^2 \\
 & \quad (0.0019) \quad (0.00007) \quad (0.0049) \quad (0.0001)
 \end{aligned}$$

Domanda 1 (Punti: 2)

La variabile *proprietario* risulta statisticamente associata alla probabilità che la domanda sia accettata? Se si, qual è la migliore interpretazione che e' possibile dare del suo effetto su tale probabilità?

.....

.....

.....

.....

.....

.....

.....

.....

Domanda 1 (Punti: 2)

Qual è l'interpretazione del coefficiente di *spesecc*? Visto che il coefficiente è molto piccolo è corretto concludere che quanto il richiedente spende su altre carte è un fattore economicamente poco rilevante?

.....
.....
.....
.....
.....
.....
.....
.....

Domanda 1 (Punti: 2)

Qual è l'associazione fra reddito e probabilità che la domanda sia accettata? È possibile testare che questa associazione non vari con alcune caratteristiche del richiedente?

.....
.....
.....
.....
.....
.....
.....
.....

Domanda 1 (Punti: 2)

Qual è la probabilità che la domanda di un professore di liceo con le seguenti caratteristiche (i) ha 30 anni; (ii) possiede un reddito di 10.000 euro; (iii) vive in affitto; (iv) non possiede nessuna carte di credito, venga accolta?

.....
.....
.....
.....
.....
.....
.....
.....

Domanda 1 (Punti: 2)

Si prendano in considerazione due individui. Il primo —individuo A — ha le stesse caratteristiche del professore della domanda 3. Il secondo — individuo B — ha le stesse caratteristiche ma è un commercialista. L'affermazione “*Le probabilità con le quali le richieste di carta di credito dei due individui vengano accettate sono identiche, almeno dal punto di vista statistico*” è supportata dai risultati del modello?

.....
.....
.....
.....
.....
.....
.....
.....

Domanda 2

Uno studio finanziato dalle Nazioni Unite ha come obiettivo quello di comprendere l'effetto dell'istruzione sui livelli di natalità con particolare riferimento ai paesi in via di sviluppo. Lo studio usa dati su donne dello Botswana raccolti nel 1988. Il dataset contiene informazioni circa il numero di figli per donna, anni di istruzione e altre variabili concernenti lo status religioso, sociale ed economiche delle donne. In particolare, la variabile `Frsthalf` è una variabile uguale a 1 se la donna è nata fra Luglio e Dicembre `children` è il numero di figli, `educ` sono gli anni di istruzione, `age` è l'età della donna.

Domanda 1 (Punti: 2)

La seguente regressione è stimata:

$$educ = 9.6929 - 0.8523 \times frsthalf - 0.1080 \times age - 0.0005 \times age^2$$

(0.5412) (0.1132) (0.0402) (0.0007)

Cosa ci dice questa regressione circa la validità e la “debolezza” di `Frsthalf` come strumento per `educ`?

.....
.....
.....
.....
.....
.....
.....
.....
.....

Domanda 1 (Punti: 2)

I risultati della regressione che usa `frsthalf` come strumento (potete assumere lo strumento sia valido)

$$children = 3.3878 - 0.1715 \times educ + 0.3236 \times age - 0.0027 \times age^2$$

(0.5449) (0.0524) (0.0202) (0.0004)

Quel è l'intervallo di confidenza per l'effetto di un anno addizionale di istruzione in termini di numero di figli?

.....
.....
.....
.....
.....
.....

Domanda 1 (Punti: 2)

E' possibile testare che la relazione fra numero di figli ed eta' della donna sia effettivamente non lineare?

.....
.....
.....
.....
.....
.....
.....
.....
.....
.....

Domanda 3

More Guns, Less Crime è un libro di John Lott che esamina come la criminalità cambia quando gli Stati americani modificano le leggi che permettono di detenere armi da fuoco. In particolare, Lott — famoso attivista a favore della libera circolazione della armi — si concentra sulle leggi cosiddette “shall issue”. In uno Stato che ha in vigore la legge “*shall issue*” è molto più semplice l’acquisto di armi da fuoco senza la necessità di un controllo sull’idoneità psicologica dell’acquirente. Usando dati sui 50 stati per il periodo 1977-1994, Lott trova che la legge “*shall issue*” diminuisce il numero di crimini violenti.

Nel 2003, due ricercatori di Yale, Ayres and Donohoue, estendono il dataset includendo il periodo 1995-1999 e studiano l’effetto di questo tipo di leggi sui crimini violenti usando tecniche econometriche. Il dataset contiene le seguenti variabili:

- `state` indica lo Stato
- `year` indica l’anno
- `violent` crimini violenti (incidenti per 100,000 persone)
- `murder` omicidi (incidenti per 100,000 persone)
- `prisoners` numero di carcerati (per 100,000 persone; valori anno precedente)
- `afam` percentuale della popolazione dello Stato che e’ Afro-Americans
- `density` popolazione per miglio quadrato dello superficie dello Stato (diviso per 1000)
- `law =1` se lo Stato ha una legge "shall carry" in vigore nell’anno

Tre diversi modelli sono stimati. Le statistiche per variabile nel dataset sono riportate in Tabella 1.

Tabella 1: Statistiche descrittive

Statistic	N	Mean	St. Dev.	Min	Max
violent	1,173	503.075	334.277	47.000	2,921.800
murder	1,173	7.665	7.523	0.200	80.600
prisoners	1,173	226.580	178.888	19	1,913
afam	1,173	5.336	4.886	0.248	26.980
cauc	1,173	62.945	9.762	21.780	76.526
density	1,173	0.352	1.355	0.001	11.102
law	1,173	0.243	0.429	0	1

1. **[Modello A]** Il primo modello

$$\log(\text{murder}_{it}) = \beta_0 + \beta_1 \text{law}_{it} + u_{it},$$

è stimato con il seguente comando

```
mod1 <- feols(log(murder)~law, data = Guns, vcov = "hetero")
```

2. **[Modello B]** Il secondo modello

$$\log(\text{murder}_{it}) = \beta_0 + \beta_1 \text{law}_{it} + \alpha_i + S_t + u_{it},$$

è stimato con il seguente comando

```
mod2 <- feols(log(murder)~law|state+year, data = Guns, vcov = "hetero")
```

3. **[Modello C]** Il terzo modello

$$\log(\text{murder}_{it}) = \beta_0 + \beta_1 \text{law}_{it} + \beta_2 \text{afam}_{it} + \beta_3 \text{prisoners}_{it} + \beta_4 \text{density}_{it} + \alpha_i + S_t + u_{it},$$

è stimato con due comandi

```
mod3 <- feols(log(murder)~law+afam+prisoners+density|state+year,
              data = Guns, vcov = "hetero")
```

I risultati della stima sono contenuti nella Tabella 2.

Dependent Variable:	log(murder)		
Model:	(1)	(2)	(3)
<i>Variables</i>			
Constant	1.898 (0.0220)		
law	-0.4734 (0.0485)	-0.0335 (0.0270)	-0.0174 (0.0268)
afam			0.0339 (0.0261)
prisoners			-0.0002 (0.0002)
density			-0.6921 (0.1477)
<i>Fixed-effects</i>			
state		Yes	Yes
year		Yes	Yes
<i>Fit statistics</i>			
Observations	1,173	1,173	1,173
R ²	0.08337	0.91256	0.91740
Within R ²		0.00165	0.05700

Heteroskedasticity-robust standard-errors in parentheses

Domanda 1 (Punti: 2)

Nel Modello A, qual è l'interpretazione di $\hat{\beta}_1$? Il coefficiente è statisticamente significativo?

.....

.....

.....

.....

.....

.....

.....

.....

.....

Domanda 1 (Punti: 2)

Nel Modello B, qual è l'interpretazione di $\hat{\beta}_1$? Il coefficiente è statisticamente significativo?

.....
.....
.....
.....
.....
.....
.....
.....

Domanda 1 (Punti: 2)

Nel Modello C, qual è l'interpretazione di $\hat{\beta}_1$? Il coefficiente è statisticamente significativo?

.....
.....
.....
.....
.....
.....
.....
.....

Domanda 1 (Punti: 2)

Quale dei tre modelli è da preferire per poter dare una interpretazione causale alla stima del coefficiente di interesse (β_1). (Fornire una risposta precisa)

.....
.....
.....
.....
.....
.....
.....
.....

Domanda 1 (Punti: 2)

Si consideri il Modello C. Quale assunzione è necessario imporre sugli errori affinché gli errori standard riportati in Tabella 2 siano validi? Nel caso tale assunzione non fosse soddisfatta, come è possibile ottenere gli errori standard appropriati?

.....

.....

.....

.....

.....

.....

.....

.....