

Scuola di Scienze Politiche
Corso di Laurea LM–88 SOCIOLOGIA E RICERCA SOCIALE
Compito di Metodi Statistici per la Ricerca Sociale
20 Dicembre 2016

Parte applicativa

Esercizio 1

Due comuni limitrofi (comune A e comune B) devono decidere se creare una fusione. Al fine di capire l'opinione dei cittadini sulla possibile fusione viene effettuata un'indagine su un campione di 1000 soggetti iscritti nelle liste elettorali dei due comuni. La seguente tabella mostra la distribuzione di frequenza dell'opinione dei cittadini intervistati classificati per comune di residenza:

Comune	Opinione sulla fusione	
	Non favorevole	Favorevole
Comune A	90	410
Comune B	125	375

- (a) Stimare la proporzione dei cittadini favorevoli nel comune A e nel comune B
- (b) Sottoporre a test l'ipotesi che la proporzione di favorevoli nei due comuni sia la stessa: (i) Scrivere l'ipotesi nulla e l'ipotesi alternativa; (ii) Scrivere la regione critica al livello di significatività del 5%; (iii) Calcolare il valore della statistica test e prendere una decisione. (iv) Calcolare il p-valore e confrontare il risultato con quello ottenuto utilizzando l'approccio di Neyman-Pearson
- (c) Calcolare l'odds a favore di essere favorevole alla fusione nei due comuni
- (d) Calcolare l'odds-ratio e costruire un intervallo di confidenza al livello del 95% per l'odds-ratio nella popolazione

Esercizio 2

Su un campione di 5 donne di un certo paese sono stati rilevati i seguenti dati sul numero di anni di istruzione (variabile X) e il numero di figli (variabile Y)

Anni di istruzione	5	8	10	5	15
Numero di figli	4	3	2	5	1

- (a) Stimare i coefficienti di regressione di un modello di regressione che pone il numero dei figli (variabile Y) in funzione degli anni di istruzione (variabile X)
- (b) Interpretare i coefficienti ottenuti al punto (a)
- (c) Stimare la varianza degli errori del modello di regressione
- (d) Stimare il numero medio di figli di donne con 13 anni di istruzione.. Costruire quindi un intervallo di confidenza al livello di confidenza del 95% per il numero medio di figli di donne con 13 anni di istruzione.

Esercizio 3

Si consideri un campione di $n = 34$ donne di un certo paese. Si supponga di stimare un modello di analisi della varianza a due vie che pone il numero di figli in funzione del titolo di studio (1 =Nessun Titolo, 2 =Scuola Primaria, 3 =Scuola Secondaria, 4 =Diploma, 5 = Laurea o più) e dell'attitudine verso la religione (1 = Religioso; 0 = Non religioso).

- (a) Completare la tavola di analisi della varianza che segue relativa a tale modello:

Fonte di variabilità	Somma dei quadrati	Gradi di libertà	Media dei quadrati	$F - value$
Regressione				
Residua	145.75			–
Totale	276.12			–

- (b) Specificare l'ipotesi nulla e l'ipotesi alternativa a cui si riferisce il valore della statistica F calcolato al punto (a). Scrivere quindi la regione critica al livello di significatività del 5% e prendere una decisione.
- (c) Confrontare il modello in esame con un modello ridotto in cui viene eliminata la variabile titolo di studio sapendo che la somma dei quadrati dei residui del modello ridotto è $SQE_r = 207.47$. (i) Scrivere l'ipotesi nulla e l'ipotesi alternativa; (ii) Calcolare il valore della statistica test; (iii) Definire il p -valore per tale test e prendere una decisione sapendo che il p -valore per tale test è 0.037.

Domande di Teoria

Domanda 1

La seguente tabella riporta le stime dei coefficienti del modello di analisi della varianza a due vie (senza interazione) per i dati dell'esercizio 3 (la categoria di riferimento è scritta tra parentesi):

Variabile	Stima
Costante	5.75
Titolo di studio (Nessun Titolo)	
Scuola Primaria	-0.67
Scuola Secondaria	-1.49
Diploma	-2.86
Laurea o più	-2.50
Religiosità (Non religioso)	0.46

- (a) Scrivere l'equazione di regressione stimata
- (b) Interpretare i coefficienti di regressione relativi alla variabile titolo di studio.
- (c) Come cambierebbe l'interpretazione del coefficiente relativo alla modalità "Scuola primaria" della variabile titolo di studio se nel modello fosse presente l'interazione tra religiosità e titolo di studio?

Domanda 2

La seguente tabella mostra i risultati di un modello di regressione logistica stimato su campione di soggetti iscritti alla liste elettorali di un certo comune. La variabile risposta, Y , è tale che $Y = 1$ se il soggetto è favorevole alla fusione del comune con un altro comune limitrofo; $Y = 0$ se il soggetto non è favorevole alla fusione. Le variabili esplicative sono età e titolo di studio del soggetto (0 = Al massimo la scuola dell'obbligo; 1 = Diploma o più)

Variabile	Stima	
Costante	-0.87	0.246
Età	0.09	0.008
Titolo di studio (Scuola dell'obbligo o meno)		
Diploma o più	0.49	0.107

- (a) Interpretare i coefficienti in termini di odds
- (b) Scrivere l'espressione delle probabilità stimate
- (c) Stimare le probabilità di essere favorevole alla fusione per un soggetto di 35 anni con almeno il diploma e per un soggetto di 35 anni senza il diploma. Confrontare le due probabilità.