

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

Parte applicativa

Esercizio 1

A un campione di residenti di un piccolo comune viene chiesto il numero di vani presente nella propria abitazione e se sono soddisfatti della propria abitazione. La seguente tabella mostra la distribuzione di frequenza congiunta:

Soddisfazione	Numero di vani					
	1	2	3	4	5	6
No	22	17	12	6	7	5
Si	8	9	10	14	10	16

- (a) Costruire la distribuzione di frequenza relativa condizionata del livello di soddisfazione dato che un soggetto vive in una abitazione con 6 vani
- (b) Calcolare le frequenze teoriche di indipendenza (usare un'approssimazione a due cifre decimali)
- (c) Verificare se livello di soddisfazione e numero di vani della propria abitazione sono indipendenti: (i) Scrivere l'ipotesi nulla e l'ipotesi alternativa; (ii) Scrivere la regione critica al livello di significatività del 5%; (iii) Prendere una decisione sapendo che il valore osservato della statistica test è 18.64.
- (d) Il p -value per tale test è 0.0022. Commentare tale valore anche sulla base della decisione presa al punto (c)

Esercizio 2

In uno studio sulla relazione tra Y = tasso di criminalità (numero di crimini ogni 1000 abitanti) e X = tasso di povertà (percentuale di popolazione con reddito sotto il livello di povertà) nelle città un certo paese si è considerata un campione di 20 città. Per tale campione si sono osservate le seguenti statistiche:

$$\sum_{i=1}^{20} x_i = 280 \quad \sum_{i=1}^{20} y_i = 170 \quad \sum_{i=1}^{20} x_i^2 = 4851 \quad \sum_{i=1}^{20} y_i^2 = 4181 \quad \sum_{i=1}^{20} x_i \cdot y_i = 3776.5$$

- (a) Calcolare il coefficiente di correlazione lineare tra tasso di criminalità e tasso di povertà
- (b) Stimare i coefficiente del modello di regressione che pone il tasso di criminalità in funzione del tasso di povertà
- (c) Stimare la varianza degli errori del modello di regressione
- (d) Per una certa città si è osservato un tasso di povertà $x = 12$ e un tasso di criminalità $y = 6$. Calcolare il tasso di criminalità atteso utilizzando il modello di regressione e il residuo di regressione per tale città. Costruire quindi un intervallo di confidenza al 95% per il tasso di criminalità atteso di città con un tasso di povertà $x = 12$.

Esercizio 3

Con riferimento ai dati dell'esercizio 2 si supponga che sia stimato un modello di regressione che pone il Y = tasso di criminalità (numero di crimini ogni 1000 abitanti) in funzione di X_1 = tasso di povertà (percentuale di popolazione con reddito sotto il livello di povertà) X_2 = percentuale di residenti con almeno il diploma e X_3 = percentuale di residenti che vivono in aree urbane.

- Costruire la tavola di analisi della varianza per tale modello sapendo che la somma dei quadrati dei residui è $SSE_e = 275.8$
- Confrontare tale modello con il modello ridotto considerato nell'esercizio 2. (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
- In base al risultato ottenuto al punto (b), il p -valore relativo al test condotto al punto (b) sarebbe stato maggiore, minore o uguale a 0.5? Giustificare la risposta.

Domande di Teoria

Domanda 1

In uno studio sulla soddisfazione su un nuovo servizio comunale, il livello di soddisfazione (misurato con un punteggio su scala 0-100) viene posto in funzione del genere (0=Femmina, 1=Maschi) e del livello di istruzione (1=Obbligo; 2=Diploma; 3=Laurea). La seguente tabella mostra i risultati dell'analisi:

Variabile	Stima	Errore standard
Costante	2.4	0.936
Genere (Femmina)		
Maschio	0.8	0.350
Livello di istruzione (Obbligo)		
Diploma	-0.2	0.075
Laurea	-0.5	0.080

- Scrivere l'equazione di regressione stimata
- Interpretare i coefficienti di regressione
- Scrivere l'ipotesi nulla e l'ipotesi alternativa per il confronto di tale modello con un modello (esteso) in cui è inclusa l'interazione tra genere e livello di istruzione

Domanda 2

In uno studio sulla pratica dell'auto-palpazione come mezzo di prevenzione del cancro al seno viene stimato un modello logit in cui la pratica dell'auto-palpazione (0 = No, 1 = Si) è la variabile risposta (Y) e la partecipazione a un corso organizzato da personale medico su come effettuare l'auto-palpazione ($X_1 = 0 =$ No, $X_1 = 1 =$ Si) e X_2 = l'età sono le variabili esplicative. La seguente tabella mostra i risultati della stima del modello:

Variabile	Coefficiente	Errore standard
Costante	1.28	0.40
Partecipazione al corso	1.17	0.3
Età	0.007	0.005

- Interpretare il segno del coefficiente della variabile partecipazione al corso
- Interpretare il coefficiente della variabile partecipazione al corso in termini di odds
- Calcolare la probabilità di successo stimata per una donna di età uguale a 35 anni che partecipa al corso