

Scuola di Scienze Politiche
Corso di Laurea LM-88 SOCIOLOGIA E RICERCA SOCIALE
Compito di Metodi Statistici per la Ricerca Sociale
16 Gennaio 2017

Parte applicativa

Esercizio 1

La seguente tabella mostra la distribuzione di frequenza congiunta ottenuta classificando un campione di cittadini di un certo comune classificato rispetto al livello di soddisfazione relativa ai servizi sociali presenti sul territorio e al livello di istruzione

Livello di istruzione	Livello di Soddisfazione		
	Insoddisfatto	Soddisfatto	Molto Soddisfatto
No Diploma	18	60	40
Diploma o più	48	72	40

- (a) Costruire la distribuzione di frequenza relativa condizionata del livello di soddisfazione dato un livello di istruzione pari al diploma o più elevato
- (b) Calcolare la frequenza teorica di indipendenza per cittadini senza diploma e livello di soddisfazione pari a molto soddisfatto. Calcolare il corrispondente residuo (ossia la corrispondente contingenza)
- (c) Verificare l'ipotesi che livello di soddisfazione e livello di istruzione siano indipendenti. (i) Scrivere l'ipotesi nulla e l'ipotesi alternativa; (ii) Scrivere la regione critica al livello di significatività del 10%; (iii) Prendere una decisione sapendo che il valore della statistica test χ^2 è $\chi^{2,oss} = 9.25$

Esercizio 2

In uno studio sui comportamenti di giovani adulti, su un campione di $n = 28$ soggetti di età compresa tra i 20 e i 40 anni si sono rilevate le seguenti variabili Y = Numero di libri letti in un anno; X_1 = età e X_2 = Titolo di Studio (1 = Obbligo o meno, 2 = Diploma, 3 = Laurea o più). I seguenti risultati sono relativi alle stime dei minimi quadrati di un modello di regressione che pone il numero di libri letti in un anno, Y , in funzione delle variabili esplicative età (scartata dalla media pari a 30 anni), e Titolo di Studio (per cui è considerata come categoria di riferimento la scuola dell'obbligo o meno):

Variabile	Coeff.	E.S.
Costante	1.92	0.37
Età - 30	0.16	0.05
Livello di istruzione		
Diploma	1.39	0.28
Laurea o più	1.32	0.36
$SST = 136.96$	$SSE = 36.61$	

- (a) Scrivere l'equazione di regressione stimata
- (b) Stimare il numero medio di libri letti da un soggetto di 35 anni laureato
- (c) Stimare la varianza del modello
- (d) Confrontare il modello stimato con il modello nullo. (i) Scrivere l'ipotesi nulla e l'ipotesi alternativa; (ii) Scrivere la regione critica al livello di significatività del 1%; (iii) Calcolare il valore della statistica test e prendere una decisione

Esercizio 3

La seguente tabella mostra alcune statistiche descrittive del punteggio al test di ingresso (misurato su scala 0-100) alla scuola di medicina in tre Atenei: Ateneo A , Ateneo B e Ateneo C .

Ateneo	Punteggio al test di ingresso		
	Dimensione del campione	Media campionaria	Varianza campionaria
Ateneo A	20	71	90.25
Ateneo B	24	74	96.04
Ateneo C	29	66	81.00

- (a) Calcolare la devianza entro i gruppi e la devianza tra gruppi
- (b) Verificare l'ipotesi che il punteggio medio al test di ingresso sia lo stesso nei tre atenei. (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
- (c) Il p-valore relativo al test al punto (b) è 0.01013. Commentare tale risultato.
- (d) Costruire un intervallo di confidenza al livello di confidenza del 97% per la differenza tra il punteggio medio nell'Ateneo B e il punteggio medio nell'Ateneo C utilizzando la correzione di Bonferroni.

Domande di Teoria

Domanda 1

Si consideri i risultati mostrati nell'esercizio 2.

- (a) Interpretare il coefficiente relativo all'età e il coefficiente relativo a un livello di istruzione pari al diploma.
- (b) Calcolare l'indice di determinazione lineare e interpretarlo
- (c) Definire il coefficiente di correlazione multipla. Calcolare quindi il coefficiente di correlazione multipla per il dati dell'esercizio 2.

Domanda 2

Si considerino le seguenti misure di associazione tra due caratteri binari:

$$\text{Differenza tra proporzioni: } \pi_2 - \pi_1 \quad \text{Rischio relativo: } RR = \frac{\pi_2}{\pi_1} \quad \text{Odds-ratio: } OR = \frac{\pi_2/(1-\pi_2)}{\pi_1/(1-\pi_1)}$$

- (a) Per ciascuna misura definire l'insieme dei valori che può assumere.
- (b) Specificare il valore che tali misure assumono nel caso di assenza di associazione
- (c) Si supponga che in uno studio sull'abbandono degli studi universitari si voglia confrontare la proporzione degli abbandoni tra studenti appartenenti a famiglie a basso reddito che ricevono una borsa di studio (popolazione 2) e studenti appartenenti a famiglie a basso reddito che non ricevono alcuna borsa di studio (popolazione 1). L'odds ratio stimato su un campione di tali studenti è risultato $\hat{OR} = 0.75$. Interpretare tale valore.