

Scuola di Scienze Politiche
Corso di Laurea LM-88 SOCIOLOGIA E RICERCA SOCIALE
Compito di Metodi Statistici per la Ricerca Sociale
07 Settembre 2016

Parte applicativa

Esercizio 1

La seguente tabella mostra alcune statistiche descrittive del numero di ore lavorate calcolate su un campione di lavoratori autonomi di due paesi, paese *A* e paese *B*

Paese	Numero di ore lavorate		
	Dimensione del campione	Media campionaria	Varianza campionaria
Paese A	21	42	144.0
Paese B	25	48	149.5

Si supponga che le ore lavorate dai lavoratori autonomi nei due paesi abbiano distribuzione Normale con uguale varianza $\sigma_A^2 = \sigma_B^2 = \sigma^2$.

- (a) Calcolare una stima della varianza σ^2
- (b) Verificare l'ipotesi che i lavoratori autonomi del paese A lavorino in media lo stesso numero di ore dei lavoratori autonomi del paese B. (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) In base al risultato ottenuto al punto (b) stabilire se il p -valore per il test al punto (b) sarebbe stato minore/uguale/maggiore di 0.05.
- (d) Costruire un intervallo di confidenza al livello di confidenza del 95% per la differenza tra le medie. Interpretare l'intervallo di confidenza ottenuto e confrontarlo con il risultato del test al punto (b).

Esercizio 2

Su un campione di 6 lavoratori autonomi di un certo paese si è rilevato il numero di ore lavorate (in media) alla settimana e il numero di anni di attività:

Anni di attività	12	18	8	22	21	15
Ore Lavorate	42	40	56	38	39	49

- (a) Calcolare il coefficiente di correlazione tra gli anni di attività e le ore lavorate
- (b) Stimare i coefficienti di regressione di un modello di regressione che pone le ore lavorate (variabile Y) in funzione degli anni di attività (variabile X)
- (c) Stimare la varianza degli errori del modello di regressione
- (d) Stimare il numero medio di ore che in media lavoratori autonomi che hanno iniziato la loro attività da 20 anni dedicano al lavoro alla settimana. Costruire quindi un intervallo di confidenza al livello di confidenza del 90% per il numero medio di ore che in media lavoratori autonomi che hanno iniziato la loro attività da 20 anni dedicano al lavoro alla settimana

Esercizio 3

Su un campione di 56 liberi professionisti autonomi che operano come avvocati, commercialisti imprenditori o medici in un certo paese si è stimato un modello di analisi della varianza che pone il numero di ore lavorate (in media) alla settimana in funzione del tipo di attività (Avvocato; Commercialista; Imprenditore, Medico) e della zona in cui operano (Nord; Centro, Sud-Isole)

- (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	825			—
Totale	2750			—

- (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 1% e prendere una decisione.
- (c) Confrontare il modello in esame con un modello ridotto in cui viene eliminata la variabile zona di residenza sapendo che la somma dei quadrati dei residui del modello ridotto è $SQE_r = 912.5$. (i) Scrivere l'ipotesi nulla e l'ipotesi alternativa; (ii) Calcolare il valore della statistica test e prendere una decisione sapendo che il p -valore per tale test è 0.0082.

Domande di Teoria

Domanda 1

Su un campione di liberi professionisti che operano come avvocati, commercialisti imprenditori o medici in un certo paese si è stimato un modello di analisi della covarianza che pone il numero di ore lavorate (in media) alla settimana in funzione del tipo di attività (Avvocato; Commercialista; Imprenditore, Medico) e del numero di anni di attività. La seguente tabella riporta le stime dei coefficienti del modello:

Variabile	Stima
Costante	47.7
Tipo di attività (<i>Avvocato</i>)	
Commercialista	−0.2
Imprenditore	1.3
Medico	3.2
Anni di attività	−0.5

- (a) Scrivere l'equazione di regressione stimata
- (b) Interpretare i coefficienti di regressione del modello
- (c) Come cambierebbe l'interpretazione del coefficiente relativo alla variabile anni di attività se nel modello fosse presente l'interazione tra tipo di attività e anni di attività?

Domanda 2

La seguente tabella mostra i risultati di un modello di regressione logistica stimato su campione di clienti di età non inferiore ai 18 anni di un certo ristorante. La variabile risposta, Y , è il giudizio espresso dai clienti sulla qualità del servizio ($Y = 1$ se il giudizio è positivo; $Y = 0$ se il giudizio è negativo) e la variabile esplicativa è l'età del cliente.

Variabile	Stima
Costante	2.8
Età	−0.06

- (a) Interpretare il segno del coefficiente della variabile età
- (b) Interpretare il coefficiente della variabile età in termini di odds
- (c) Stimare la probabilità di esprimere un giudizio positivo per un cliente di età uguale a 40 anni