

LM-88 SOCIOLOGIA E RICERCA SOCIALE

Compito di Metodi Statistici per la Ricerca Sociale 10 Luglio 2014

Studente: _____

Matricola: _____

Esercizi

- Interessa confrontare il prezzo al litro (in euro) dell'olio extra vergine di oliva prodotto in Toscana e in Liguria. Il prezzo al litro (in euro) dell'olio extra vergine di oliva prodotto da un campione di 4 aziende Toscane è (7.2, 9.5, 7.0, 7.5). Il prezzo al litro (in euro) dell'olio extra vergine di oliva prodotto da un campione di 7 aziende della Liguria è risultato in media di 8.5 Euro con varianza campionaria 1.44. Rispondere ai seguenti quesiti supponendo che le popolazioni siano Normali e abbiano uguale varianza.
 - Stabilire con un test al livello di significatività del 10% se il prezzo dell'olio extra vergine di oliva prodotto in Toscana e in Liguria è diverso.
 - Costruire un intervallo di confidenza al 90% per la differenza tra le medie del prezzo al litro (in euro) dell'olio extra vergine di oliva prodotto in Toscana e in Liguria. Confrontare il risultato dell'intervallo di confidenza con il risultato del test ottenuto al punto precedente.
- La seguente tabella mostra i risultati di un'analisi di regressione che pone il consumo di alcool (misurato come il numero di bevande alcoliche che il soggetto ha bevuto nell'ultimo mese) in funzione delle seguenti variabili esplicative: Sesso (0 = Femmina, 1 = Maschio); Stato civile (1 =single; 2 =sposato; 3 =divorziato/separato; 4 = Vedovo) e consumo di alcool tre anni prima. Per le variabili categoriche la categoria di riferimento è scritta in parentesi.

Variabile	Stima	Errore standard
Costante	9.40	1.25
Genere (Femmina)		
Maschio	6.54	0.74
Stato civile (Single)		
Sposato	-3.49	0.72
Div./Sep.	2.38	0.78
Vedovo	-0.92	0.63
Consumo di alcool 3 anni prima	0.31	0.12

$$n = 56 \quad SQT = 1233.09 \quad SQE = 145.39$$

- Scrivere l'equazione di regressione stimata e interpretare i coefficienti di regressione
- Calcolare una stima della varianza degli errori del modello di regressione
- Costruire la tavola di analisi della varianza e sottoporre a test l'ipotesi che tutti i coefficienti di regressione relativi alle variabili esplicative siano nulli al livello di significatività del 5%.

3. Si considerino i dati dell'esercizio precedente. Sapendo che la somma dei quadrati dei residui di un modello ridotto in cui non è incluso lo stato civile è $SQE_r = 251.55$, confrontare il modello esteso (in cui è incluso lo stato civile) con il modello ridotto (in cui non è incluso lo stato civile) utilizzando un livello di significatività pari a 5%.

Domande Teoriche

1. Si supponga di voler studiare l'associazione tra la variabile risposta status occupazionale (1 = occupato; 0 = disoccupato) e le variabili esplicative sesso e anni di istruzione. Specificare un modello di regressione logistica (senza interazione tra variabili esplicative) definendo sia il logit della probabilità di essere occupato che la probabilità di essere occupato.
2. Con riferimento al modello di regressione lineare, definire cosa si intende per diagnostica e analisi dei residui, illustrandone le principali tecniche, anche con riferimento alle assunzioni alla base del modello di regressione.