

**SCUOLA DI SCIENZE POLITICHE**  
**CdL: LM-88 SOCIOLOGIA E RICERCA SOCIALE**

**Compito di**  
**Metodi Statistici per la Ricerca Sociale**  
**11 Febbraio 2015**

Studente: \_\_\_\_\_

Matricola: \_\_\_\_\_

**Esercizi**

1. Per valutare gli effetti a lungo termine dei programmi prescolastici per i bambini in condizioni economiche disagiate, su un gruppo di soggetti provenienti da famiglie in difficoltà economiche si è rilevato se hanno frequentato una scuola materna pubblica all'età di 3-4 anni e l'eventuale necessità di assistenza sociale nell'età adulta:

Scuola materna	Necessità di assistenza sociale	
	No	Si
No	12	49
Si	24	38

- (a) Calcolare la proporzione di soggetti che hanno avuto necessità di assistenza sociale nell'età adulta tra coloro che hanno frequentato una scuola materna pubblica e coloro che non hanno frequentato una scuola materna pubblica
- (b) Verificare se questi dati rappresentano un'evidenza significativa a favore dell'ipotesi secondo cui frequentare la scuola materna riduce la necessità futura di ricorrere ai servizi sociali al livello di significatività del 5%
- (c) Calcolare gli odds a favore della necessità di assistenza sociale nell'età adulta tra coloro che hanno frequentato una scuola materna pubblica e coloro che non hanno frequentato una scuola materna pubblica
- (d) Calcolare il rapporto degli odds e costruire un intervallo di confidenza per il rapporto degli odds al livello di confidenza del 95%.
2. Per valutare l'impatto della pubblicità nell'industria del tabacco, una ricerca ha studiato la relazione tra denaro investito in pubblicità di una particolare marca di sigarette e preferenze espresse nei fumatori adulti. La seguente tabella mostra i dati raccolti

Marca	Pubblicità in milioni di euro	Preferenza per la marca (in percentuale)
1	75	23.5
2	43	6.7
3	35	4.8
4	21	4.5
5	4	3.0

- (a) Calcolare il coefficiente di correlazione lineare tra denaro investito in pubblicità di una particolare marca di sigarette e preferenze espresse nei fumatori adulti
- (b) Determinare le stime dei minimi quadrati dei coefficienti della retta di regressione che pone le preferenze ( $Y$ ) in funzione del denaro investito in pubblicità ( $X$ )

- (c) Sottoporre a test l'ipotesi  $H_0 : \beta_1 = 0$  versus  $H_1 : \beta_1 \neq 0$  al livello di significatività dell'1%
- (d) Stimare la percentuale media di adulti che esprimerebbero una preferenza verso una marca di sigarette che investe in pubblicità 30 milioni di euro e costruire il relativo intervallo di confidenza al livello di confidenza del 99%.
3. In un campione di 32 pazienti vengono rilevati la pressione del sangue (variabile risposta  $Y$ ), l'indice di Quetelet ( $100 \times \text{peso} / \text{altezza}^2$ ) (variabile  $X_1$ ), una variabile binaria  $X_2$  che vale 1 se l'individuo è un fumatore e 0 altrimenti. Si sono stimati tre modelli di regressione:

$$\begin{array}{ll} \text{Modello Esteso} & M_e : Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i1} x_{i2} + \epsilon_i \\ \text{Modello ridotto 1} & M_{r1} : Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i \\ \text{Modello ridotto 2} & M_{r2} : Y_i = \beta_0 + \beta_1 x_{i1} + \epsilon_i \end{array}$$

Le tavole di analisi della varianza relative ai tre modelli sono

Modello esteso					
Fonte		$SQ$	$GdL$		
Regressione		4184.1071			
Residua					
Totale		6425.969			

  

Modello ridotto 1			Modello ridotto 2		
Fonte	$SQ$	$GdL$	Fonte	$SQ$	$GdL$
Regressione			Regressione	3537.9453	
Residua	2305.6027		Residua		
Totale			Totale		

- (a) Completare le tre tavole di analisi della varianza
- (b) Confrontare il modello esteso e il modello nullo al livello di significatività del 5%. Specificare le ipotesi e la regione critica, calcolare il valore osservato della statistica test e decidere se rifiutare o meno l'ipotesi nulla.
- (c) Confrontare il modello esteso e il modello ridotto 1 al livello di significatività del 5%. Specificare le ipotesi e la regione critica, calcolare il valore osservato della statistica test e decidere se rifiutare o meno l'ipotesi nulla.

## Domande Teoriche

- In uno studio clinico finalizzato a valutare se la somministrazione di morfina (per via orale) prima di un intervento chirurgico possa migliorare il decorso post-operatorio è stato rilevato il livello di dolore postoperatorio a 4 ore dalla fine dell'operazione su un campione di pazienti suddiviso in due gruppi: al gruppo I è stata somministrata della morfina prima dell'intervento e al gruppo II è stato somministrato un placebo. Nell'ipotesi che le popolazioni siano normali con uguale varianza, un test di ipotesi sulla differenza tra le due medie,  $H_0 : \mu_1 - \mu_2 = 0$  versus  $H_1 : \mu_1 - \mu_2 \neq 0$ , ha fornito un  $p$ -value pari a 0.03. Si può concludere che i dati mostrano evidenza contro  $H_0$  al livello di significatività del 5%? E al livello di significatività del 1%? (Giustificare le risposte)
- In uno studio sulla senilità si osservano una variabile binaria che assume valore 1 per soggetti che presentano sintomi di senilità e il punteggio ottenuto a un particolare test (il punteggio WAIS = Wechsler Adult Intelligence Scale). Si noti che valori maggiori della variabile WAIS indicano capacità intellettive migliori. Specificare un modello logistico che pone la presenza dei sintomi di senilità in funzione della variabili WAIS. Scrivere sia l'espressione del logit delle probabilità sia l'espressione della probabilità di successo. Interpretare il coefficiente relativo alla variabile WAIS in termini di odds.