

LM–88 SOCIOLOGIA E RICERCA SOCIALE

Compito di Metodi Statistici per la Ricerca Sociale 13 Febbraio 2014

☐ 6 CFU ☐ 9 CFU

Studente: _____

Matricola: _____

Esercizi

1. Un campione di 320 operai sono stati classificati per età e per aver avuto un incidente sul lavoro nell'ultimo anno. I risultati sono mostrati nella seguente tabella:

Età	Incidente	
	No	Si
< 50	190	10
≥ 50	111	9

- (a) Sottoporre a test l'ipotesi nulla che la proporzione di incidenti sul lavoro sia la stessa nei gruppi di età al livello di significatività del 5%
- (b) Calcolare l'odds ratio considerando come probabilità di successo la probabilità che il soggetto abbia avuto un incidente sul lavoro. Costruire quindi un intervallo di confidenza per l'odds ratio al livello di confidenza del 95%.
2. Al fine di determinare l'affidabilità del consumo nominale (in km per litro) delle macchine, sei diverse macchine sono state guidate per uno stesso percorso ottenendo i seguenti risultati

Macchina	Consumo	Consumo
	nominale	effettivo
<i>A</i>	20	16
<i>B</i>	25	19
<i>C</i>	30	25
<i>D</i>	30	22
<i>E</i>	25	18
<i>F</i>	15	12

- (a) Calcolare il coefficiente di correlazione lineare tra consumo nominale e consumo effettivo
- (b) Stimare la retta di regressione che pone il chilometraggio effettivo (Y) in funzione del chilometraggio nominale (X)
- (c) **(SOLO PER ESAME DA 9 CFU)** Trovare un intervallo di confidenza al 90% per il coefficiente di regressione β_1

3. I seguenti dati si riferiscono al tasso di criminalità in un campione di comuni Italiani (dati fittizi)

Gruppo	Numerosità del gruppo	Media Campionaria	Varianza Campionaria
j	n_j	\bar{x}_j	s_j^2
Nord	50	85	81.00
Centro	50	70	81.90
Sud-Isole	50	97	82.81
Totale	150	84	203.6228

- (a) Calcolare la devianza tra gruppi e la devianza entro gruppi
- (b) Verificare l'ipotesi che non ci sia differenza tra i tassi di criminalità medi nelle diverse zone geografiche al livello di significatività del 5%.

Domande Teoriche

1. Si consideri un modello di regressione logistica per la probabilità di trovare lavoro per soggetti in cerca di prima occupazione. Come variabili esplicative si consideri l'età, il genere e il titolo di studio (con categorie nessun titolo o scuola dell'obbligo, diploma, laurea o più). Specificare il modello di regressione logistica.
2. Si consideri un modello di regressione lineare multipla. Definire il concetto di coefficiente di correlazione multipla e la sua relazione con l'indice di determinazione lineare multipla.
3. **(SOLO PER ESAME DA 9 CFU)** Si fornisca una descrizione dei modelli di durata (caratteristiche generali, dati, ambiti di applicazione, esempi). Spiegare cosa significa "censura a destra" e definire la funzione di sopravvivenza e la funzione di rischio (hazard function).