

UNIVERSITA' CATTOLICA DEL S. CUORE - MILANO
 LAUREA IN ECONOMIA E COMMERCIO
 LAUREA IN SCIENZE STATISTICHE ED ECONOMICHE - D.U. IN STATISTICA
Prova scritta di STATISTICA del 3.2.2000 (Tema 175)

1) Da un'indagine sulla qualità dell'aria in una data provincia lombarda si è rilevata, per un certo intervallo di tempo, la concentrazione media giornaliera di anidride solforosa (SO₂), variabile **A**, i cui valori sono stati classificati secondo le seguenti modalità: a₁="sotto il livello di attenzione"; a₂="tra il livello di attenzione e di allarme"; a₃="sopra il livello di allarme". La seguente tabella riporta le frequenze relative delle modalità della variabile **A**, separatamente per i giorni ventosi (**v**) e non ventosi (**nv**), che costituiscono le modalità della variabile **X**:

	a₁	a₂	a₃
v	0.77	0.23	0
nv	0.15	0.55+(q/100)	0.30-(q/100)

1.1) Confrontare, tramite opportuni indici di posizione e di mutabilità, le distribuzioni di **A** nei giorni ventosi e non ventosi.

1.2) Sapendo poi che il 30% dei giorni considerati sono stati ventosi, ricostruire la tabella delle frequenze relative congiunte della variabile doppia (**X**, **A**) e dire, calcolando un opportuno indice normalizzato, se esiste connessione tra le due variabili.

1.3) Si sostituiscano alle modalità della variabile **A** i seguenti intervalli (chiusi a destra) dei valori effettivamente rilevati (in µg/mc), a₁=1-5, a₂=5-8, a₃=8-13. Rappresentare graficamente la funzione di regressione di **A** condizionata ad **X**; calcolare quindi il suo indice di adattamento.

2) Nella seguente tabella sono riportati i valori puntuali della variabile concentrazione media giornaliera di SO₂ (**Y**) rilevati nei giorni ventosi e della variabile velocità del vento (in m/sec):

velocità	31	28	23	30	25	16	20	18	15	10
Y	2.4	3.3	4.6	2.5	6.2	5.4	5.9	4.9	6.3	9.3

Detta **Z** la variabile ottenuta dalla trasformazione $Z = \text{velocità} \cdot (q+1)$, si calcolino i parametri dei seguenti modelli di regressione:

$$\text{I) } Y^* = a + bZ \qquad \text{II) } Y^* = c + d/Z$$

Si dica, mediante il confronto di opportuni indici, quale modello è migliore.

3) Sia **X** la variabile che rappresenta i redditi mensili di 5 soggetti. Sapendo che il reddito totale è di 15 milioni e che i valori delle intensità cumulate relative sono (0.1; 0.233; 0.433; 0.667; 1):

- a) ricostruire i redditi effettivi dei 5 soggetti;
- b) rappresentare graficamente la curva di Lorenz e calcolare il rapporto di concentrazione.

4) Siano **X**, **Y** e **Z** tre variabili statistiche le cui covarianze sono rispettivamente cov(**XY**)=12.923, cov(**YZ**)=-6.318, cov(**XZ**)=-9.899. Sapendo che la varianza di **X** è 20.25, che la varianza di **Z** è 4.84 e che la varianza residua del piano di regressione di **Y** in funzione di **X** e di **Z** è 2.897, si determini:

- a) il valore della varianza di **Y**;
- b) l'indice di bontà del modello.

5) Siano **L** e **A** due barche a vela in competizione per un trofeo che prevede 5 gare. Sapendo che la probabilità che si rompa l'unico spinnaker che le due barche hanno a disposizione in ogni gara è rispettivamente 0.4 per la barca **L** e 0.5 per la barca **A**, calcolare la probabilità che nelle 5 gare si rompano 3 spinnaker ad **A** e uno a **L**.