

NB: Commentare sempre i risultati ottenuti

1. Considerando il modello dell'analisi della varianza a due criteri di classificazione si fissi l'attenzione sullo stimatore di massima verosimiglianza di $\alpha_j = \mu_j - \mu$ e se ne ricavi il valore atteso, la varianza e la distribuzione di probabilità, commentando adeguatamente.
2. Il centro di ricerca di un'azienda che produce carburanti per autoveicoli vuole verificare se l'autonomia media di un motore (espressa in Km percorsi) dipende dal tipo di benzina impiegato (A, B e C) e/o dall'additivo ad essa abbinato (I, II, III e IV).
A tale scopo si considerano complessivamente 36 motori identici (3 per ogni associazione tra benzina e additivo) e se ne valuta l'autonomia (in Km) con 10 litri di carburante.
La seguente tabella riporta alcune informazioni della tabella di ANOVA ottenuta con l'esperimento.

Fonte di variabilità	S.S. (Devianza)	Gradi Di Libertà			
Benzina	60,15				
Additivo	17,53				
Interazione	85,21				
Errore	63,54				
<i>Totale</i>					

Dopo aver specificato il modello lineare utilizzato nell'analisi della varianza a due criteri di classificazione ed il significato dei parametri in esso contenuti, si completi la tabella ANOVA e si verifichino le consuete ipotesi utilizzando un livello di significatività $\alpha = 0,05$ commentando i risultati ottenuti.

3. La seguente tabella riporta i dati relativi ai laureati dell'anno 1998 classificati in base alla condizione occupazionale nell'anno 2001 (fonte ISTAT). In particolare per cinque tipi di corsi di laurea (Scientifico, Chimico-Farmaceutico, Insegnamento, Psicologico, Agrario) sono disponibili le seguenti informazioni: $X_1 = \text{lavoratori}$; $X_2 = \text{in cerca di lavoro}$; $X_3 = \text{non in cerca di lavoro}$. La seguente tabella riporta i dati standardizzati:

Tipi di Corsi di laurea	Z_1	Z_2	Z_3
Scientifico	1,313	-0,178	1,455
Chimico-Farmaceutico	0,814	-1,211	0,952
Insegnamento	0,042	1,170	-0,917
Psicologico	-0,725	1,137	-0,860
Agrario	-1,444	-0,917	-0,630

La matrice D delle *distanze di manhattan* tra i tipi di corsi di laurea, considerati nello stesso ordine di presentazione della tabella è:

$$D = \begin{bmatrix} 0 & 2,0353 & 4,9913 & 5,6687 & 5,5810 \\ & 0 & 5,0212 & 5,6986 & 4,1327 \\ & & 0 & 0,8576 & 3,8604 \\ & & & 0 & 3,0028 \\ & & & & 0 \end{bmatrix}.$$

- a) Tracciare il dendrogramma riferito ai tipi di corsi di laurea avvalendosi del *metodo del centroide*;
 - b) suggerire una opportuna partizione, giustificando la scelta;
 - c) descrivere la partizione individuata al punto precedente.
4. Sulle unità di una popolazione si sono rilevate quattro variabili quantitative. Dalla matrice R di correlazione tra le quattro variabili, si sono ricavate la matrice diagonale Λ degli autovalori e la matrice A dei corrispondenti autovettori normalizzati.

$$\Lambda = \begin{bmatrix} 1,8825 & 0 & 0 & 0 \\ 0 & 1,5051 & 0 & 0 \\ 0 & 0 & 0,5391 & 0 \\ 0 & 0 & 0 & 0,0732 \end{bmatrix} \quad A = \begin{bmatrix} 0,0320 & 0,7125 & 0,6517 & -0,2581 \\ 0,7082 & 0,0371 & 0,1932 & 0,6780 \\ 0,6501 & -0,3345 & 0,0648 & -0,6792 \\ -0,2735 & -0,6156 & 0,7306 & 0,1112 \end{bmatrix}.$$

- a) Si stabilisca, mediante opportuni criteri, il numero di componenti principali da mantenere nell'analisi;
- b) si determinino le correlazioni tra le variabili originarie e le componenti principali mantenute e se ne fornisca la rappresentazione grafica commentando.