

NB: Commentare sempre i risultati ottenuti

- Si illustri il modello di analisi della varianza a due criteri di classificazione specificando le ipotesi assunte, il significato dei parametri e i vincoli sugli stessi. Si illustri successivamente la scomposizione della devianza totale necessaria alla costruzione delle statistiche test per la verifica delle consuete ipotesi di nullità dei parametri del modello.
- Con l'obiettivo di verificare quale tra quattro metodi di insegnamento (A, B, C e D) del linguaggio dei segni risulta più efficace, i 20 studenti di una classe vengono suddivisi casualmente in quattro gruppi ciascuno dei quali è sottoposto ad un diverso metodo di insegnamento. Al termine del ciclo di lezioni gli studenti vengono sottoposti al medesimo test di apprendimento. Il test prevede un punteggio tra 0 e 100. La seguente tabella riporta, per ciascun gruppo, i punteggi ottenuti dagli studenti:

| | | | | | | |
|-----|----|----|----|----|----|----|
| A | 75 | 78 | 81 | 79 | 85 | |
| B | 65 | 70 | 72 | 66 | 75 | 78 |
| C | 93 | 92 | 86 | 92 | | |
| D | 75 | 74 | 73 | 76 | 75 | |

Sapendo che la *devianza totale* è 1234 dopo aver specificato il modello da utilizzare:

- si calcolino le stime per gli effetti specifici $\alpha_j = \mu_j - \mu..$ $j = 1, 2, 3, 4$ imputabili ai livelli del fattore *metodo di insegnamento* commentando opportunamente;
- si verifichi (utilizzando un livello di significatività $\alpha = 0,05$) se il metodo di insegnamento influisce significativamente sul punteggio medio ottenuto dagli studenti al test di apprendimento;
- si verifichi, ad un livello di significatività $\alpha = 0,05$, la seguente ipotesi:

$$H_0 : \frac{\mu_1 + \mu_4}{2} = \frac{\mu_2 + \mu_3}{2} \quad \text{contro} \quad H_1 : \frac{\mu_1 + \mu_4}{2} \neq \frac{\mu_2 + \mu_3}{2}$$

interpretando opportunamente il risultato ottenuto.

- La seguente tabella riporta il valore delle esportazioni (X_1) e delle importazioni (X_2) italiane in milioni di euro nei confronti di 6 Paesi nell'anno 2004 (Fonte ISTAT).

| Paese | X_1 | X_2 |
|----------|-------|-------|
| Germania | 38186 | 50695 |
| Svizzera | 11772 | 9333 |
| USA | 22374 | 9993 |
| Spagna | 20244 | 16739 |
| UK | 19400 | 11999 |
| Francia | 34624 | 30753 |

La corrispondente matrice dei dati standardizzati è:

| Paese | Z_1 | Z_2 |
|----------|--------|--------|
| Germania | 1,506 | 1,954 |
| Svizzera | -1,386 | -0,823 |
| USA | -0,225 | -0,778 |
| Spagna | -0,459 | -0,325 |
| UK | -0,551 | -0,644 |
| Francia | 1,116 | 0,615 |

- Applicare la procedura di analisi dei gruppi non gerarchica delle k -medie selezionando come centri iniziali i Paesi *Germania* e *USA*;
 - descrivere la partizione ottenuta con riferimento ai dati originari (non standardizzati).
- Sulle unità di una popolazione si sono rilevate quattro variabili quantitative. Dalla matrice R di correlazione tra le quattro variabili originarie, si sono ricavati gli autovalori λ_j ($j = 1, \dots, 4$) e la matrice A dei corrispondenti autovettori normalizzati:

$$\begin{aligned} \lambda_1 &= 1,779 \\ \lambda_2 &= 1,373 \\ \lambda_3 &= 0,529 \\ \lambda_4 &= 0,319 \end{aligned} \quad A = \begin{bmatrix} -0,635 & 0,128 & 0,599 & 0,471 \\ -0,286 & -0,696 & -0,496 & 0,435 \\ 0,678 & 0,070 & 0,139 & 0,718 \\ -0,235 & 0,703 & -0,613 & 0,272 \end{bmatrix}.$$

- Si stabilisca, mediante opportuni criteri, il numero di componenti principali da mantenere nell'analisi;
- si valuti, ricorrendo al calcolo delle comunalità, l' idoneità della scelta effettuata;
- si rappresentino graficamente le correlazioni tra le variabili originarie e le prime due componenti principali commentando opportunamente sia la posizione delle variabili rispetto alle componenti principali sia la posizione di ciascuna variabile rispetto alle altre.