

NB: Commentare sempre i risultati ottenuti

- 1.
2. In un'azienda sono attivi due processi produttivi a ciclo continuo (*I* e *II*) per la produzione di due differenti tipi di viti. Al fine di migliorare la qualità dei propri prodotti, l'azienda decide di inserire, al termine del processo produttivo, un dispositivo che sia in grado di testare la qualità di ciascuna vite prodotta e scarti quelle difettose. Al fine di scegliere quale dispositivo adottare tra i tre disponibili sul mercato (*A*, *B* e *C*), ciascun dispositivo viene applicato ad ogni processo produttivo per 5 giorni. Sulla produzione giornaliera si verifica infine la percentuale *X* di viti difettose che il dispositivo ha correttamente scartato.
La seguente tabella riporta le percentuali medie di viti correttamente scartate:

Dispositivo	Processo produttivo		$\bar{X}_{j..}$
	<i>I</i>	<i>II</i>	
<i>A</i>	83,4	93,4	88,4
<i>B</i>	77,8	86,8	82,3
<i>C</i>	83,6	74,8	79,2
$\bar{X}_{.k.}$	81,6	85,0	83,3

Dopo aver specificato il modello da utilizzare soffermandosi in particolare sul significato del parametro di interazione, si verifichino ad un livello di significatività $\alpha = 0,10$ le seguenti ipotesi, sapendo che la *devianza nei gruppi* è 1124,00:

- a) il tipo di dispositivo influenza significativamente la percentuale media di viti difettose individuate;
 b) il processo produttivo influenza significativamente la percentuale media di viti difettose individuate;
 c) vi è *interazione* tra dispositivo e processo produttivo.
3. In occasione dell'indagine multiscope *Aspetti della vita quotidiana*, agli individui intervistati (di almeno 14 anni) è stato chiesto il livello di soddisfazione su alcuni aspetti della vita quotidiana. Il livello di soddisfazione è stato misurato sulla scala qualitativa ordinale: *per niente*, *poco*, *abbastanza* e *molto*.
La seguente tabella riporta, per le regioni indicate, le percentuali standardizzate di intervistati che hanno espresso come livello di soddisfazione *abbastanza* o *molto* in merito a:

	Regione	Z_1	Z_2	Z_3	Z_4
X_1 situazione economica;	Lombardia	0,945	1,091	1,222	0,327
X_2 salute;	Emilia-Romagna	1,057	0,587	1,431	0,957
X_3 relazioni familiari;	Lazio	-0,172	-0,836	-2,408	-1,484
X_4 relazioni con amici.	Basilicata	-0,898	-1,932	-1,710	-0,343
	Calabria	-1,479	-2,347	0,524	-0,736
	Sicilia	-2,004	-1,458	0,035	-0,933

Fonte: Istat - Annuario statistico italiano 2003

La matrice *D* delle *distanze euclidee* tra le sei regioni, considerate nello stesso ordine di presentazione della tabella è:

$$D = \begin{bmatrix} 0 & 0,841 & 4,628 & 4,646 & 4,395 & 4,265 \\ 0,841 & 0 & 4,923 & 4,661 & 4,328 & 4,367 \\ 4,628 & 4,923 & 0 & 1,875 & 3,626 & 3,165 \\ 4,646 & 4,661 & 1,875 & 0 & 2,378 & 2,200 \\ 4,395 & 4,328 & 3,626 & 2,378 & 0 & 1,159 \\ 4,265 & 4,367 & 3,165 & 2,200 & 1,159 & 0 \end{bmatrix}.$$

- a) Tracciare il dendrogramma riferito alle sei regioni avvalendosi del *metodo del legame medio*;
 b) individuare e descrivere la partizione ben strutturata e minimale.
4. Sulle unità di una popolazione si sono rilevate cinque variabili quantitative. Dalla matrice *R* di correlazione tra le cinque variabili, si sono ricavati gli autovalori λ_j ($j = 1, \dots, 5$) e la matrice *A* dei corrispondenti autovettori normalizzati.

$$\begin{matrix} \lambda_1 = 2,487 \\ \lambda_2 = 1,266 \\ \lambda_3 = 0,665 \\ \lambda_4 = 0,531 \\ \lambda_5 = 0,051 \end{matrix} \quad A = \begin{bmatrix} 0,609 & -0,114 & -0,182 & 0,133 & 0,752 \\ -0,560 & -0,057 & 0,492 & -0,260 & 0,611 \\ -0,400 & -0,326 & -0,800 & -0,276 & 0,130 \\ -0,139 & -0,765 & 0,161 & 0,604 & -0,071 \\ -0,368 & 0,541 & -0,241 & 0,688 & 0,200 \end{bmatrix}.$$

- a) Si stabilisca, mediante opportuni criteri, il numero di componenti principali da mantenere nell'analisi;
 b) si valuti, ricorrendo al calcolo delle comunalità, l'idoneità della scelta effettuata.