

Politecnico di Milano
Facoltà di Ingegneria dei Sistemi

APPELLO DI STATISTICA APPLICATA

Milano, 14 Settembre 2006

Nome e cognome:

Numero di matricola:

©I diritti d'autore sono riservati. Ogni sfruttamento commerciale non autorizzato sarà perseguito.

Problema 1

In America centrale vivono 3 diverse specie di scoiattoli: rosso, grigio e volante. Sono stati misurati rispettivamente peso (hg) e altezza (dm) di 30 esemplari, 10 per ogni specie. In seguito sono riportate la media e la matrice di covarianza dei tre gruppi:

$$M_R = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad M_G = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad M_V = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$
$$S_R = \begin{pmatrix} 0.5 & -0.25 \\ -0.25 & 2 \end{pmatrix} \quad S_G = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad S_V = \begin{pmatrix} 0.5 & 0.25 \\ 0.25 & 1 \end{pmatrix}$$

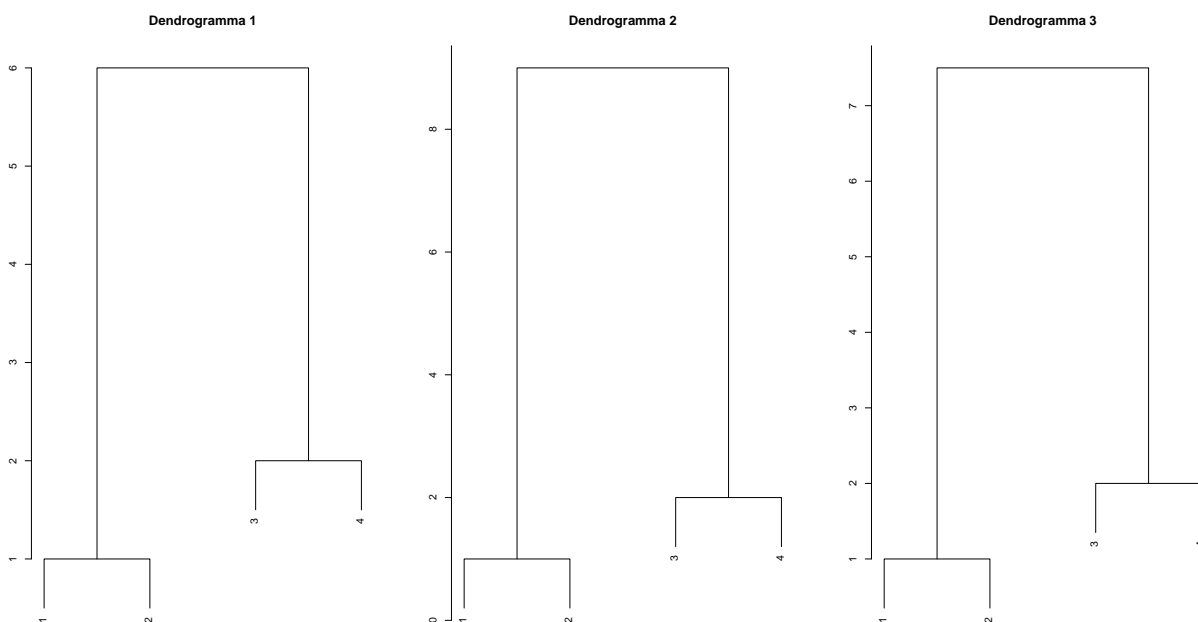
- (a) Si individuino e si rappresentino le tre regioni di classificazione R_R , R_G e R_V utilizzando un'analisi discriminante lineare basata sulla regola di Bayes. Si suppongano identiche le probabilità a priori di appartenenza alle tre specie.
- (b) Si commenti il significato e la validità dell'analisi effettuata al punto precedente.
- (c) Si calcolino le probabilità di appartenenza alle tre diverse specie per un generico scoiattolo che pesi 3 hg e sia alto 2 dm .

Problema 2

Una rivista specializzata in telefonia vuole confrontare le prestazioni (durata in standby e durata in conversazione) di quattro tipi diversi di mini-batterie usa e getta:

Batteria	standby (h)	conver. (h)
1	11	2
2	12	2
3	15	5
4	15	7

Per effettuare il confronto vengono implementati tre diversi algoritmi di clustering gerarchico agglomerativo utilizzando per ogni caso la distanza di Manhattan abbinata ad un particolare linkage. I rispettivi dendrogrammi sono riportati in figura.



- (a) Esistono dei cluster? In caso affermativo, quali?
- (b) Si identifichino i linkage utilizzati nei tre algoritmi.
- (c) Per ciascun algoritmo si riporti matrice di dissimilarità e matrice cofenetica.
- (d) Si confronti l'efficacia dei tre algoritmi.

Problema 3

Un'azienda farmaceutica ha realizzato un esperimento clinico su 6 ratti per conoscere l'effetto sulla pressione sanguigna di una nuova sostanza chimica. Ad ogni ratto è stata misurata la pressione appena prima della somministrazione, 8, 16 e 24 ore dopo la somministrazione. In seguito sono riportate media e matrice di covarianza campionarie delle quattro misure:

$$\bar{x} = \begin{pmatrix} 1 \\ 4 \\ -3 \\ 2 \end{pmatrix} \quad S = \begin{pmatrix} 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 & 0.5 \\ 0.5 & 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 \end{pmatrix}$$

- (a) Si esegua un test di livello 5% per provare che la sostanza influenza la pressione sanguigna nell'arco delle prime 24 ore.
- (b) Si mettano in evidenza gli effetti provocati dalla sostanza sulla pressione sanguigna.

Problema 4

In un piccolo paese della Svizzera sono presenti due distributori di carburante: uno Esso ed uno Shell. Entrambi vendono sia benzina a 95 ottani che benzina a 98 ottani.

Un giovane statistico del luogo vuole scoprire a che distributore e con quale benzina debba rifornire la sua auto per cercare di massimizzare il numero di chilometri percorsi con un litro di carburante.

Dopo 8 pieni le rese da lui misurate sono:

km/l	Dist.	Benz.
18	Esso	95
16	Esso	95
20	Esso	98
22	Esso	98
14	Shell	95
16	Shell	95
22	Shell	98
24	Shell	98

- (a) Mediante un'ANOVA two ways si individui a che distributore e con quale benzina è meglio che il giovane rifornisca la sua auto.
- (b) Esiste interazione tra distributore e tipo di benzina?
- (c) Perché il giovane ha scelto fare due pieni allo stesso distributore e con la stessa benzina?