

Analisi Statistica Multivariata. A.A. 2011-2012
Progetto n.1. Gruppo I. Consegna: 30/5/2012.
I dati per i problemi sono reperibili sul sito del corso

Problema 1.

Si considerino i dati **pollutant** (in *DASL and More Data*). Considerare i 12 settori sorgenti di inquinamento, e gli inquinanti descritti (Carbon Monoxide, PM 2.5 e 10, Lead, Nitrogen.Oxides, Volatile Organic Compounds, Sulfur Dioxide). Fare un'analisi di clustering usando single e complete linkage, considerando almeno tre distanze, commentando sulle scelte fatte. Interpretare i risultati alla luce del problema.

Valutare i raggruppamenti con anche un metodo di Multidimensional scaling, e confrontare i risultati.

Problema 2.

Si considerino i dati degli Hook-billed Kites per i maschi (Tab.6-11) e per le femmine (Tab.5-12), dove x_1 è la lunghezza della coda e x_2 è la lunghezza dell'ala. (si veda anche il libro).

1. Per la matrice di osservazioni relativa ad ogni popolazione considerata, fare uno studio della normalità univariata, bivariata e multivariata. Verificare la presenza di outliers e, nel caso, eliminarli dall'analisi successiva giustificando la scelta. Trasformare eventualmente le variabili, o alcune di esse, per migliorare la normalità dei dati.
2. Determinare la regione e gli intervalli simultanei di confidenza (99% e 95%) per la media, relativamente alla famiglia di maschi.
3. Per le osservazioni provenienti dalle due popolazioni, valutare l'ipotesi di uguale media delle popolazioni, con livello di significatività $\alpha = 0.05$ e $\alpha = 0.01$. Commentare i risultati. In caso di rifiuto, valutare quale delle variabili è più responsabile del rifiuto, facendo il confronto di medie sulle singole variabili.
4. Nel confronto delle medie, determinare intervalli simultanei di confidenza (95%) e commentare su eventuali differenze tra i risultati ottenuti rispetto al test del quesito precedente.
5. Impostare un test di discriminanza per allocare la nuova osservazione

$$\mathbf{x} = [195, 270],$$

anche mediante uno studio grafico.