

Corso di Laurea Magistrale in Analisi e Gestione dell'Ambiente  
Corso di Analisi Statistica Multivariata a.a. 2011-2012  
Prova scritta del 19/07/2012

1. I dati in **cicale.txt** (nel sito del corso, di fianco al testo della prova scritta) si riferiscono a variabili di un campione di 3 specie di cicale (si legga il testo). Si considerino le SOLE due famiglie tredecula (44 oss.) e tredecassini (54 oss.), e come variabili le due variabili ottenute come:

$$x_1 = BL/BW \quad x_2 = WL*WW$$

Si vuole valutare se le misure delle due specie appartengono alla stessa popolazione.

- (a) fare un test di normalità univariata e bivariata delle due popolazioni (liv. sign.  $\alpha = 0.05$ );
  - (b) Fornire regione ed intervalli di confidenza al 95% e 99% per la media per entrambe le popolazioni (su stesso grafico);
  - (c) Fare un confronto delle medie delle due popolazioni individuando la regione di confidenza per la differenza delle medie (liv. sign.  $\alpha = 0.05$ );
  - (d) Effettuare un'analisi di discriminanza delle due popolazioni. Qual'è il valore di APER? Commentare.
2. La tabella **T1-5.DAT** (JW) contiene dati monitorati per uno studio di inquinamento dell'aria in vari siti di Los Angeles alle 12.00 di giorni diversi. Si tratta delle variabili  $x_1$ =Wind,  $x_2$ =Rad.Solare,  $x_3$ =CO,  $x_4$ =NO,  $x_5$ =NO2,  $x_6$ =O3,  $x_7$ =HC.

Nel seguito, lavorare sulle *variabili*.

- (a) Calcolare la matrice di correlazione e commentare i valori trovati;
- (b) Effettuare un'analisi di clustering sulle variabili usando almeno due distanze diverse e commentare. Confrontare con una analisi di MultiDimensional Scaling;
- (c) Supponendo una distribuzione normale della popolazione, fare un test di fattibilità di un'analisi fattoriale per tali variabili.
- (d) Fare un'analisi fattoriale ed interpretare i fattori determinati. Stimare i fattori per tutte le osservazioni e commentare.