

Corsi di Laurea Magistrale in Scienze per l'Ambiente
Corso di Analisi Statistica Multivariata a.a. 2009-2010
Prova scritta del 02/07/2010 ore 8.30

1. La tabella T9-12.DAT contiene i risultati di una valutazione dell'attività e di un test attitudinale su un gruppo di $n = 50$ venditori di una azienda. Le variabili si riferiscono a: x_1 : Crescita nelle vendite; x_2 : Capacità di vendita; x_3 : Vendite a nuovi acquirenti; x_4 : Creatività; x_5 : capacità meccaniche; x_6 : capacità di astrazione; x_7 : abilità matematiche. Dopo aver fatto una trasformazione **logaritmica**:
 - (a) Calcolare la matrice di correlazione e commentare i valori trovati;
 - (b) Supponendo una distribuzione normale della popolazione, fare un test di fattibilità di un'analisi fattoriale per tali variabili. L'ipotesi di normalità è ragionevole?
 - (c) Fare un'analisi fattoriale ed interpretare i fattori determinati. Stimare i fattori per l'ultima osservazione.

2. Si considerino le $n = 50$ osservazioni relative alla famiglia di *iris versicolor* del data set di R. Fisher (file IRIS1). Le due variabili rappresentano la lunghezza del sepalo (x_1) e del petalo (x_2).
 - (a) Valutare la normalità dei dati con le tecniche note ed eventualmente usare trasformazioni per migliorare la normalità.
 - (b) Determinare la regione (ellisse) di confidenza al 95% e 99% per la *media* delle due variabili. Determinare inoltre gli intervalli simultanei (T^2) di confidenza al 95% e 90% per la media delle due variabili. Riportare i grafici in un diagramma di dispersione.
 - (c) Si consideri quindi anche il campione di $n = 50$ individui della famiglia di *iris virginica* (IRIS2). Confrontare le matrici di covarianza. Fare un test di uguaglianza delle medie, con un livello di significatività del 5% e dell' 1%.