

Corso di Laurea Specialistica in Scienze Ambientali
Università di Bologna
Metodi Probabilistici e Statistici per l'Analisi dei Dati
Programma del corso. a.a. 2006-2007

Testo di riferimento: *Applied Multivariate Statistical Analysis*, R. A. Johnson e D. W. Wichern, Prentice Hall, V edizione, 2002.

- L'organizzazione dei dati. Statistica descrittiva multivariata.
- Misure di dissimilarità (distanze). Misure di similarità e associazione. Analisi di Clusters con metodi di aggregazione.
- Multidimensional scaling classico.
- Vettori e Matrici Casuali. Vettore media e Matrice di Varianza/Covarianza. Combinazioni lineari di vettori casuali e loro media e varianza. Campioni casuali e valori attesi per il vettore media e la Matrice di Covarianza campionarie. Varianza Generalizzata e Varianza totale campionarie.
- La distribuzione normale multivariata. Funzione di densità e proprietà fondamentali. La distribuzione campionaria della media e della matrice di covarianza. Comportamento per grandi campioni. Test sull'ipotesi di normalità. Trasformazioni per quasi-normalità.
- La Statistica T^2 di Hotelling. Regioni di confidenza e intervalli di confidenza simultanei per il vettore media di una distribuzione normale multivariata. Il caso di grandi campioni.
- Confronto di coppie di trattamenti. Confronto di vettori di media tra due popolazioni. Caso di uguale matrice di covarianza. Caso di matrici di covarianza diverse. Confronto di medie di varie popolazioni multivariate: test sul Λ di Wilks e test di Bartlett.
- Separazione e classificazione per due popolazioni, con valutazione della probabilità a priori e dei costi. Classificazione di due popolazioni normali multivariate con matrici di covarianza uguali. La funzione di discriminazione di Fisher.
- Modello classico di Regressione lineare multivariata. La Stima col metodo dei minimi quadrati. Regione di confidenza per i parametri di regressione. Valutazione del modello mediante i valori stimati. Stime del valore atteso e previsione di una nuova osservazione.
- Componenti Principali. Variabili standardizzate. Analisi della variabilità dei dati mediante l'uso delle componenti principali. Studio grafico.
- Il modello fattoriale ortogonale. Il metodo dei fattori (o componenti) principali per la stima dei pesi. Valutazione della bontà della stima. Rotazione dei fattori. Valutazione dei fattori.