

Corsi di Laurea Specialistica in Scienze Ambientali  
Corso di Metodi probabilistici e statistici per l'analisi dei dati a.a. 2003-2004  
Prova scritta del 02/04/2004

1. In valori in tabella CRIME.DAT si riferiscono a dati forniti da agenzie governative di 47 stati americani e relativi al 1960, per lo studio di come la presenza di criminalità possa dipendere da altre variabili. Le variabili si riferiscono a

$y$	Rapporto di crimini (ogni milione di ab.)
$x_1$	Numero di maschi di età 14-24 per 1000 abitanti
$x_2$	Indicatore di Stato (0 = Nord, 1 = Sud)
$x_3$	Scolarizzazione (media pesata del numero di anni per persona, età > 25)
$x_4$	Spesa procapite per la sicurezza (anno 1960)
$x_5$	Spesa procapite per la sicurezza (anno 1959)
$x_6$	Forza lavoro maschile (ogni 1000 ab.) con età 14-24
$x_7$	Numero di maschi ogni 1000 femmine
$x_8$	Popolazione dello Stato (in centinaia di migliaia)
$x_9$	Numero di non bianchi (ogni 1000 ab.)
$x_{10}$	Rapporto di disoccupaz. maschile (età 14-24) ogni 1000 ab.
$x_{11}$	Rapporto di disoccupaz. maschile (età 25-39) ogni 1000 ab.
$x_{12}$	Stipendio mediano (in decine di USD)
$x_{13}$	Numero di famiglie (ogni 1000 fam.) con stipendi inferiori a 1/2 la mediana

Mediante un Modello di regressione lineare, studiare la dipendenza del rapporto di crimini dalle alle variabili catalogate. In particolare

- (a) Determinare la matrice di correlazione delle variabili predittrici. Commentare.
  - (b) Determinare una stima corretta per i parametri del modello.
  - (c) Valutare la bontà del modello mediante tutti gli strumenti noti (coeff. determinazione, studio grafico del residuo).
  - (d) Fare un test di significatività ( $\alpha = 0.5$ ) sulle due variabili predittrici  $x_{12}, x_{13}$ , considerate contemporaneamente.
2. Il gene di coleottero *Chaetocnema* ha tre specie: Concinna, Heikertingeri e Heptapotamica. La tabella FLEABEETLES.DAT riporta le misure della larghezza ed angolo del "guscio" di ogni coleottero delle 3 specie.
- Fare un test MANOVA per il confronto delle medie delle tre popolazioni:
- (a) Riportare i dati delle tre specie in un diagramma di dispersione. Commentare.
  - (b) Calcolare le matrici di covarianza e commentare sulle ipotesi usuali dell'analisi.
  - (c) Valutare l'ipotesi di uguale media mediante la Lambda di Wilks ( $\alpha = 0.10$ ).