Corso di Laurea Specialistica in Scienze Ambientali
Corso di MPS per l'analisi dei dati a.a. 2007-2008
Prova scritta del 27/06/2008

1. I dati nella tabella FISH.F si riferiscono al diametro (in centesimi di pollice) degli anelli di un campione di salmone femmina dopo un anno in acqua dolce (x_1) e dopo un anno in oceano (x_2), per due popolazioni di salmone: Canadese (popolazione 1) e dell'Alaska (popolazione 2). Si vuole valutare se le due popolazioni sono effettivamente distinte.
 - (a) Dopo aver effettuato la trasformazione $\tilde{x} \leftarrow \sqrt[4]{x}$, fare un test di normalità univariata e bivariata delle due popolazioni; **usare le variabili trasformate nel resto dell'esercizio**
 - (b) Fare un confronto delle medie delle due popolazioni, valutando sia la regione di confidenza che gli intervalli simultanei di confidenza per la differenza delle medie, considerando un livello di significatività $\alpha = 0.05$.

2. I valori nella tabella T1-5.DAT si riferiscono ai valori di air-pollution. (Si veda anche la fotocopia)

Mediante un Modello di regressione lineare, studiare la dipendenza del valore di NO_2 (risposta, var. X_5), dal vento (X_1) e dalla radiazione solare (X_2).

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 da solo, con la risposta e con le variabili predittrici). Valutare la presenza di eventuali osservazioni anomale.
- (d) Costruire intervalli di confidenza al 95% per la risposta, corrispondenti a $x_1 = 10$, $x_2 = 80$.
- (e) Fare un test di significatività ($\alpha = 0.05$) sul vento come variabile predittrice.