

Corso di Laurea Magistrale in Analisi e Gestione dell'Ambiente
Corso di Analisi Statistica Multivariata a.a. 2012-2013
Prova scritta del 17/07/2013

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).
 - (a) Fare un diagramma di dispersione dei due campioni e commentare.
 - (b) Fare un test di normalità univariata e bivariata dei due campioni (liv. sign. $\alpha = 0.05$) (non effettuare trasformazioni);
 - (c) Determinare la regione di confidenza (95% e 99%) per la media di ognuna delle due popolazioni. Riportare tutti i grafici sulla stessa figura; il valore $\mu_0 = [130, 340]$ potrebbe essere un buon candidato per la media della seconda popolazione? ($\alpha = 0.05, 0.01$)
 - (d) 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 $\alpha = 0.05$;
 - (e) Effettuare un'analisi di discriminanza delle due popolazioni, allocando il nuovo campione: $\mathbf{x}_0 = [110, 350]$. Qual'è il valore di APER? Commentare.

2. I dati nella tabella T5-1.DAT (JW) si riferiscono a 20 osservazioni sulla traspirazione, riferite a donne in salute. Le variabili riportate sono
 - x_1 : rapporto di sudorazione
 - x_2 : contenuto di sodio
 - x_3 : contenuto di potassio
 - (a) Calcolare la matrice di correlazione e commentare.
 - (b) Effettuare un'analisi della componenti principali: analizzare la varianza spiegata dalle componenti principali ed il coefficiente di correlazione tra le nuove e vecchie variabili. Valutare l'eventuale presenza di osservazioni anomale. Valutare il numero di componenti principali sufficienti per una buona descrizione delle osservazioni.
 - (c) Assumendo che le variabili originarie provenissero da una distribuzione normale, possiamo accettare l'ipotesi che ognuna delle prime due componenti principali provenga da una distribuzione normale?
 - (d) Interpretare i risultati, possibilmente individuando una caratteristica di ognuna delle componenti principali tenute. Riportare le osservazioni su di un grafico delle componenti principali tenute, e commentare.