

Politecnico di Milano
Facoltà di Ingegneria dei Sistemi
IV APPELLO DI STATISTICA APPLICATA
13 Settembre 2007

©I diritti d'autore sono riservati. Ogni sfruttamento commerciale non autorizzato sarà perseguito.

Nome e cognome:

Numero di matricola:

Problema 1

A livello del suolo, l'atmosfera terrestre è costituita principalmente da tre gas: Azoto (N_2), Ossigeno (O_2) e Argon (Ar). 100 provette sono state raccolte in altrettanti siti dislocati in nord Italia. Le concentrazioni rilevate (espresse in valori percentuali) sono riportate nel file `atmosfera.txt`.

- a) Vi è evidenza statistica del fatto che in nord Italia la concentrazione media dei tre gas sia differente da quella planetaria ovvero rispettivamente 78.08%, 20.95% e 0.93%?
- b) Utilizzando la disuguaglianza di Bonferroni, costruite quattro intervalli di confidenza globale 95% per la concentrazione dei tre gas e della loro somma.
- c) Ricorrendo ancora alla disuguaglianza di Bonferroni, si può concludere, commettendo un errore del I tipo con probabilità al più 5%, che almeno una delle tre concentrazioni medie sia inferiore a 1.2%?

Problema 2

Siano X_1 ed X_2 due variabili congiuntamente normali rispettivamente di media 1 e 2, varianza 1 e 4 e correlazione lineare 0.5.

- a) Si costruisca un intervallo simmetrico di previsione per X_1 di probabilità 90% noto che $2X_1 + X_2 = 5$.
- b) Si costruisca un intervallo simmetrico di previsione per X_2 di probabilità 90% noto che $2X_1 + X_2 = 5$.

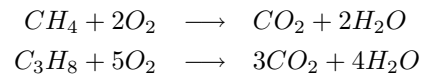
Problema 3

Per due campioni casuali indipendenti costituiti rispettivamente da 50 uomini sposati (`men.txt`) e da 50 donne sposate (`women.txt`) sono state registrate due variabili: anni di matrimonio ed età al momento del matrimonio. Da precedenti analisi sociologiche si può assumere che le matrici di covarianza relative alle due popolazioni siano uguali.

- a) Si esegua un test bivariato per la differenza delle medie.
- b) Dopo aver proiettato i dati lungo la direzione individuata dalla prima componente principale, si esegua un test univariato per la differenza delle medie delle proiezioni.
- c) Dopo aver proiettato i dati lungo la direzione individuata dall'analisi discriminante di Fisher, si esegua un test univariato per la differenza delle medie delle nuove proiezioni.
- d) Si confrontino criticamente i p -value dei test eseguiti ai punti (b) e (c).

Problema 4

In un laboratorio di chimica è stato realizzato un esperimento per confermare che la combustione di una miscela di metano (CH_4) e propano (C_3H_8) si sviluppa principalmente attraverso le due reazioni:



Durante l'esperimento diverse quantità note di metano (CH_4) e propano (C_3H_8) sono state fatte reagire in un ambiente saturo di ossigeno (O_2) e privo di anidride carbonica (CO_2). A reazione terminata è stata rilevata la quantità di anidride carbonica (CO_2) prodotta. I dati sono riportati nel file `chimica.txt`. Le quantità sono tutte espresse in numero di moli (1 *mole* = $6.022 \cdot 10^{23}$ *molecole*).

- Utilizzando il numero di moli di CH_4 e di C_3H_8 come regressori ed il numero di moli di CO_2 come risposta, si individui e si stimi un opportuno modello lineare. (Si mantenga l'intercetta nel modello e si scelga invece se eliminare dal modello il termine di interazione ed uno o più regressori. Si verifichino inoltre le assunzioni necessarie per l'inferenza sui parametri del modello.)
- Si costruiscano degli intervalli di confidenza globale 90% per i parametri β_i del modello di regressione individuato al punto (a).
- Si può concludere, commettendo un errore del I tipo con probabilità al più 1%, che la combustione non si sviluppa attraverso le due reazioni sopracitate?
- Si costruisca un intervallo di confidenza 90% per la varianza σ^2 dell'errore.