

**Politecnico di Milano**  
**Facoltà di Ingegneria dei Sistemi**  
APPELLO DI STATISTICA APPLICATA  
16 Luglio 2010

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

*Nome e cognome:*

*Numero di matricola:*

### **Problema 1**

Il 28 è un tram che ogni mezz'ora attraversa la città di Lisbona dal quartiere dell'Alfama a quello del Barrio Alto. Nel file `tram.txt` sono riportati i tempi (in minuti trascorsi dalle ore 15:00) di passaggio presso la fermata di Largo do Chiado per le corse delle 15:00, delle 15:30 e delle 16:00 relativi al mese di Giugno.

- a) Dopo aver introdotto opportune assunzioni, eseguite un test di livello 5% per verificare che le medie dei tempi di passaggio coincidano con quelle dichiarate.
- b) Utilizzando la disuguaglianza di Bonferroni, fornite 6 limiti inferiori di confidenza globale 95% per le tre medie e le tre varianze relative ai tre tempi di passaggio e commentatene i valori.

## Problema 2

Un giovane ingegnere portoghese vuole costruire una macchina in grado di distinguere in modo automatico due specie di sardine (la sardina atlantica e la sardina iberica) sulla base della lunghezza [cm] della sardina. Per fare ciò, si pone come primo obiettivo quello di costruire un classificatore partendo dalle misurazioni relative a 500 sardine atlantiche e 500 sardine iberiche (file `sardine.txt`).

- a) Si eseguano due test per la verifica dell'ipotesi di normalità delle due popolazioni, un test per l'uguaglianza delle medie, un test per l'uguaglianza delle varianze.
- b) Sulla base dei precedenti test e sapendo che il 75% delle sardine pescate appartengono alla specie atlantica mentre il 25% alla specie iberica, si costruisca un classificatore che minimizzi il numero di sardine misclassificate e se ne riportino i parametri.
- c) Si stimi l'AER del classificatore in modo analitico utilizzando le densità di probabilità stimate.

### Problema 3

Nel file `datium.txt` sono riportati rispettivamente il numero di dobloni incassati tramite dazi nel porto di Belém, il numero di dobloni utilizzati per finanziare la costruzione del Monastero dos Jeronimos e il numero di dobloni utilizzati per costruire la Torre di Belém, nei 40 anni trascorsi tra il 1706 e il 1745. Si assumano i dati realizzazioni indipendenti di una normale trivariata.

- a) Fornite tre intervalli  $T^2$ -simultanei di confidenza globale 90% per le tre medie.
- b) Possiamo smentire l'ipotesi dello storico Vasco de Moura secondo il quale, in quel periodo, il 20% degli introiti forniti dai dazi del porto di Belém veniva utilizzato per finanziare la costruzione del monastero ed il 10% per finanziare la costruzione della torre?
- c) Aggiungete agli intervalli costruiti al punto (a) altri due intervalli  $T^2$ -simultanei a supporto del precedente test.

## Problema 4

Nle file `aurum.txt` sono riportate le libbre d'oro trasportate dai 352 vascelli provenienti dalle Americhe che hanno fatto scalo presso il porto di Lisbona nell'anno 1754. Nello stesso file sono riportate per ogni vascello anche la nazionalità (Portogallo, Spagna, Francia, Inghilterra) e la lunghezza dello scafo in piedi. Si assuma per la singola nazione una dipendenza al più lineare tra la lunghezza dello scafo ed il carico d'oro medio. Si assuma inoltre la varianza del carico d'oro indipendente sia dalla nazionalità che dalla lunghezza del vascello. In fine, si assumano i 352 carichi d'oro indipendenti e distribuiti normalmente.

- a) Si introduca un modello di regressione lineare coerente con le assunzioni sopra riportate e si forniscano delle stime puntuali dei parametri.
- b) Sulla base del modello (a) si eseguano i seguenti test:
  - un test per verificare l'esistenza di una dipendenza tra il carico medio d'oro e la lunghezza del vascello;
  - un test per verificare l'esistenza di una dipendenza tra il carico medio d'oro e la nazionalità del vascello;
  - sei test per verificare l'uguaglianza, a due a due, della dipendenza tra il carico medio d'oro e la lunghezza del vascello per le diverse nazioni.
- c) Sulla base dei precedenti test si individui un modello ridotto e se ne stimino i parametri.
- d) Si commentino gli  $R^2$  del modello (a) e del modello (c).