

**Politecnico di Milano**  
**Facoltà di Ingegneria dei Sistemi**  
APPELLO DI STATISTICA APPLICATA  
14 febbraio 2011

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

*Nome e cognome:*

*Numero di matricola:*

### **Problema 1**

Uno statistico milanese ha raccolto i dati relativi ad alcuni suoi spostamenti in metropolitana nell'ultimo anno. Nel file `MM.txt` sono riportati i tempi [sec] di percorrenza della tratta Caiazzo-Loreto, i tempi [sec] di attesa in banchina alla fermata di Loreto e i tempi [sec] di percorrenza della tratta Loreto-Piola. Assumendo un modello normale tri-variato per le osservazioni raccolte:

- a) si costruiscano sei intervalli di confidenza globale 90% per le tre medie e le tre varianze;
- b) si esegua un test al 10% per confermare/smentire l'ipotesi secondo la quale il tempo di percorrenza medio della tratta Caiazzo-Loreto coincide con quello della tratta Loreto-Piola;
- c) l'ATM afferma che in media i treni stazionano a Loreto per non più di 30 secondi. Possiamo smentire tale affermazione?

## Problema 2

L'ATM sta valutando la possibilità di inserire nei tornelli dei lettori ottici in grado di misurare le dimensioni [mm] dei biglietti vidimati. Al fine di costruire un opportuno software, vi chiede di costruire una regola di classificazione che minimizzi il numero atteso di errori. Partendo dalle misure relative a 100 biglietti regolari (file `true.txt`) e a 100 biglietti contraffatti (file `false.txt`) ed essendo noto che lo 0.5% dei biglietti in circolazione risulta contraffatto:

- a) si costruisca un'opportuna regola di classificazione (in particolare, se ne verifichino, per quanto possibile, le assunzioni e si riportino su di un grafico qualitativo le regioni di classificazione);
- b) si calcoli l'APER e se ne discuta il valore;
- c) sulla base della regola individuata al punto (a) come verrà classificato un biglietto di lunghezza 85.5 mm e larghezza 55.0 mm? Con che probabilità il biglietto sarà contraffatto?

### Problema 3

Dal 1 Gennaio 2010 sono quotate in piazza affari tre titoli azionari: PoliLeo, PoliBov e PoliPub. I valori [€] dei tre titoli alla chiusura giornaliera delle contrattazioni possono essere modellizzati come una normale trivariata di media  $\mu = \begin{pmatrix} 10 \\ 10 \\ 20 \end{pmatrix}$  e matrice di covarianza  $\Sigma = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & -1 \\ -1 & -1 & 1 \end{pmatrix}$ . A seguito di alcuni disguidi tecnici, nella giornata odierna il valore delle azioni PoliPub non è stato trasmesso alla PoliBank:

- a) si calcoli la probabilità che il valore delle azioni PoliPub sia sotto ai 18 €;
- b) si calcoli la probabilità che il valore delle azioni PoliPub sia sotto ai 18 € sapendo che il valore delle azioni PoliLeo è uguale a 11 €;
- c) si calcoli la probabilità che il valore delle azioni PoliPub sia sotto ai 18 € sapendo che il valore delle azioni PoliLeo è uguale a 11 € ed il valore delle azioni PoliBov è uguale a 12 €;

## Problema 4

La Poli&Oil ha affettuato tre pozzi esplorativi ed ha raccolto dati di temperatura [Celsius] al variare della profondità [m] nei tre pozzi (file `BL1.txt`, `BL2.txt` e `BL3.txt`). Assumendo una dipendenza lineare affine (che possa variare da pozzo a pozzo) tra la temperatura media e la profondità:

- a) si stimino col metodo dei minimi quadrati i 7 parametri di un opportuno modello di regressione lineare;
- b) si esegua un test per provare la dipendenza lineare tra la temperatura media e la profondità;
- c) ipotizzando che pozzi diversi ma relativi allo stesso giacimento presentino la stessa dipendenza lineare, si eseguino quattro test per verificare le seguenti ipotesi:
  - c1) il pozzo BL1 ed il pozzo BL2 pescano dallo stesso giacimento;
  - c2) il pozzo BL1 ed il pozzo BL3 pescano dallo stesso giacimento;
  - c3) il pozzo BL2 ed il pozzo BL3 pescano dallo stesso giacimento;
  - c4) tutti e tre pescano dallo stesso giacimento;
- d) sulla base dei precedenti test si individui il numero di giacimenti e si proponga il conseguente modello ridotto;
- e) sulla base del modello (d) si forniscano degli intervalli di confidenza globale 90% per la media e la varianza della temperatura dei tre pozzi a 200 m di profondità.