

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

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

*Nome e cognome:*

*Numero di matricola:*

### **Problema 1**

Il direttore di un cinema di prossima apertura chiede il vostro aiuto per pianificare un possibile orario di programmazione giornaliera da seguirsi nei primi due mesi. Nel file `lunghezza.txt` sono riportate le lunghezze [min] del primo e del secondo tempo degli ultimi 20 film usciti.

- a) Si confermi o smentisca l'ipotesi del direttore secondo la quale il primo tempo dura in media 45 minuti ed il secondo 30 minuti;
- b) si forniscano tre intervalli  $T^2$ -simultanei di confidenza globale 95% per la durata media del primo tempo, del secondo tempo e dell'intero film;
- c) si commentino “in tre righe per il direttore” le conclusioni dell'analisi.

## Problema 2

La Poli-consumatori sta eseguendo un'indagine di mercato riguardante il costo [€] del noleggio di dvd in EU e in USA, e via web o via distributore tradizionale. Nel file `dvd.txt` sono riportati i prezzi di noleggio di 200 dvd con il relativo luogo di noleggio e la relativa modalità di noleggio.

- a) Si costruisca un opportuno modello ANOVA a due fattori; in particolare si valuti tramite un test la possibilità di rimuovere il termine di interazione ed eventualmente si riduca il modello;
- b) Sulla base del modello precedentemente individuato, si eseguano due test per testare l'influenza del fattore luogo e del fattore modalità; si proponga un eventuale modello ridotto;
- c) Sulla base del modello (b) si forniscano quattro intervalli di confidenza globale 95% per il costo medio di un noleggio USA-web, EU-web, USA-trad e EU-trad.

### Problema 3

Nel file `revenue.txt` sono riportati per 100 film selezionati a caso tra i film prodotti negli Stati Uniti nel 2010, il rispettivo incasso [M€] nella prima settimana e nel primo mese di programmazione in USA e in Italia.

- a) Si esegua un'analisi delle componenti principali; in particolare si individuino le componenti principali più importanti e se ne riporti la frazione di varianza totale spiegata;
- b) si interpreti e commenti il significato delle componenti principali selezionate.

## Problema 4

Il Policinema ha in programmazione da 5 settimane due film: “Maths” e “Stats”. Nei file `maths.txt` e `stats.txt` sono riportati gli incassi [k€] delle ultime 5 settimane per i due film. Assumendo un modello per la media del singolo film  $\mu(t) = A + B/t$ , con  $A$  e  $B$  due costanti che dipendono dal film e  $t$  il tempo [settimane] trascorso dall’inizio programmazione, ed un errore additivo normale a media nulla e a varianza costante:

- a) si stimino col metodo dei minimi quadrati i 5 parametri del modello;
- b) si esegua un test per confermare o smentire l’ipotesi secondo la quale le due curve di regressione presentano la stessa derivata;
- c) si esegua un test per confermare o smentire l’ipotesi secondo la quale le due curve di regressione presentano lo stesso asintoto orizzontale per  $t \rightarrow +\infty$ ;
- d) sulla base dei due precedenti test si proponga un modello ridotto e se ne stimino i parametri;
- e) sulla base del modello (d) si forniscano due intervalli di previsione di livello globale 90% per gli incassi dei due film alla sesta settimana di programmazione.