

Politecnico di Milano
Facoltà di Ingegneria dei Sistemi
APPELLO DI STATISTICA APPLICATA
10 Settembre 2010

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

Nome e cognome:
Numero di matricola:

Problema 1

Un allevatore di cavalli dell'East Sussex possiede 54 purosangue inglesi, 36 purosangue arabi e 18 cavalli di altre razze minori. Sapendo che in Inghilterra si ha una media di un purosangue inglese e di un purosangue arabo ogni tre cavalli, l'allevatore vuole costruire una regola di classificazione per le razze dei cavalli basata sull'altezza al garrese [cm] e sul peso [kg].

- a) Dopo aver verificato le assunzioni necessarie per l'utilizzo di un'analisi discriminante quadratica, si costruisca un classificatore utilizzando i dati disponibili all'allevatore (data set `horse.txt`) e con l'aiuto del calcolatore si riportino qualitativamente su di un grafico le regioni di assegnazione.
- b) Si stimi l'AER del classificatore mediante il calcolo dell'APER.
- c) Si stimino le probabilità di appartenenza ai tre gruppi per un cavallo di altezza al garrese 155 e di peso 500.

Problema 2

Nel file `pound.txt` sono riportati i tassi di cambio sterlina/euro utilizzati da 24 banche del Regno Unito nella settimana dal 30 Agosto al 5 Settembre 2010.

- a) Dopo aver inquadrato il problema nell'ottica di un problema di misure ripetute, si esegua un test al 5% per la verifica dell'ipotesi di costanza temporale del tasso di cambio medio.
- b) Si forniscano 6 intervalli di confidenza globale 95% per gli incrementi giornalieri medi del tasso di cambio.
- c) Si commentino i 6 intervalli precedentemente calcolati.

Problema 3

Nel file `extra.txt` sono riportate le spese [£] di rappresentanza del primo ministro inglese e del suo vice nei 12 mesi del 2009. Assumendo tali dati realizzazioni indipendenti da una normale bivariata:

- a) si costruisca una regione ellissoidale di confidenza 90% per la medie delle due spese di rappresentanza;
- b) vi è evidenza statistica del fatto che il primo ministro spenda in media più del doppio di quanto spende il suo vice?
- c) si costruisca un intervallo di confidenza 90% per la media della somma delle due spese di rappresentanza?

Problema 4

Un noto supermercato inglese ha raccolto mediante un sondaggio alcuni dati relativi alla spesa mensile (maggio 2010) di 100 suoi clienti abituali. In particolare nel file `tesco.txt`, sono riportati: spesa mensile [£] in generi di prima necessità, stipendio mensile [£], status single/non-single, sesso uomo/donna.

- a) Si stimino i parametri di un modello di regressione lineare che assuma una dipendenza lineare (variabile a seconda dello status e del sesso) tra lo stipendio mensile e la media della spesa mensile. Non si assumano interazioni tra lo status e il sesso.
- b) Utilizzando il modello (a), si eseguano tre test (ciascuno a livello 5%) per la significatività dello stipendio mensile, dello status e del sesso.
- c) Sulla base dei tre precedenti test si proponga un modello ridotto.
- d) Se possibile, si riduca ulteriormente il modello (c) e si tracci su di un grafico la curva di regressione associata al modello ridotto.
- e) Sulla base del modello (d) si forniscano due intervalli di confidenza globale 95% per la media e la varianza della spesa mensile di un cliente uomo, single, con uno stipendio mensile pari a 1200 sterline.