Compiti per casa 1 – Probabilità e Statistica 2
Claudio Agostinelli – 2013–2014
`claudio@unive.it`

Istruzioni:

1

Siano $Y_1, Y_2, \cdots, Y_n$, n variabili aleatorie indipendenti $N(\mu, \sigma^2)$ e T_1, T_2 e T_3 tre stimatori della media μ:

$$T_1 = \frac{1}{n-2}\sum_{i=1}^{n} w_i Y_i + \frac{w_n Y_n}{n-2} \qquad T_2 = \frac{2(Y_2 + Y_{n-1})}{n} \qquad T_3 = Y_{n-1} \ ,$$

e $\sum_{i=1}^{n-1} w_i = n-2$, $\sum_{i=1}^{n-1} w_i^2 = (n-2)^2$, $w_n = 1$.

1. Calcolare la distorsione dei tre stimatori;
2. Calcolare la varianza dei tre stimatori;
3. Calcolare l'errore quadratico medio dei tre stimatori;
4. Dire per ognuno dei tre stimatori se è consistente.
5. Con R (cioè scrivere il codice R appropriato per eseguire quanto richiesto):
   - Consideriamo $\mu = 0$ e $\sigma^2 = 1$ e dimensioni campionarie: $n = 3, 5, 10, 25, 50, 100, 200$;
   - Per ognuno degli stimatori considerati e per ogni dimensione campionaria, eseguendo $R = 1000$ replicazioni Monte Carlo, valutare la distorsione, la varianza, l'errore quadratico medio e l'efficienza relativa di T_1 rispetto a T_2 e T_3.

2

Un'industria automobilistica acquista degli pneumatici per vettura (tipo $C1$ e di class A) da due diverse aziende. Si supponga che l'indice di resistenza al rotolamento di uno pneumatico, si veda `www.pneumaticisottocontrollo.it/resistenzaRotolamento.html` abbia distribuzione normale con media μ e varianza σ^2. Si hanno a disposizione due campioni casuali semplici di dimensione n_1 e n_2, relativi rispettivamente alla prima e alla seconda azienda. Siano $\bar{Y}_1$ e $\bar{Y}_2$ le rispettive medie campionarie. Si considerino tre diversi stimatori per l'indice di resistenza media degli pneumatici:

$$T_1 = \bar{Y}_1, \quad T_2 = \bar{Y}_2, \quad T_3 = \frac{n_1 \bar{Y}_1 + n_2 \bar{Y}_2}{n_1 + n_2}.$$

1. Sono tutti stimatori corretti?
2. Determinare le varianze dei tre stimatori.
3. Dire, sulla base dei risultati dei punti precedenti, quale dei tre stimatori scegliereste.
4. con R:
   - Si considerino i due campioni di numerosità $n_1 = 25$ e $n_2 = 100$ disponibili nel file allegato `Homework-ps2-1-A.Rdata`
   - Caricare i data set usando la funzione `load`;
   - Per ciascun campione calcolare le tre stime;
   - Utilizzando in ogni caso $\hat{\sigma}^2 = \sum_{i=1}^{n} (x_i - T_j)^2/n$ come stima della varianza ignota σ^2 calcolare una stima della probabilità che uno pneumatico fornito non sia di classe A.
5. con R:
   - Supponiamo $n_1 = 25$ e $n_2 = 100$;
   - Consideriamo una griglia di valori per i due parametri $\mu = 3, 4, 5, 6$ e $\sigma^2 = 0.02, 0.05, 0.1$;
   - Per ogni coppia di parametri (μ, σ^2) e ogni stimatore, si eseguano $n = 500$ replicazioni Monte Carlo;
   - si rappresenti graficamente i risultati utilizzando i diagrammi a scatola e baffi (i più coraggiosi possono usare la funzione `bwplot` disponibile nella libreria `lattice`).

3

Sia $X = (X_1, \ldots, X_n)$ un c.c.s. estratto da una popolazione con funzione di probabilità

$$f(x;\theta) = \left(\frac{\theta}{2}\right)^{|x|} (1-\theta)^{1-|x|},$$

con $x = -1, 0, 1$ e $0 \leq \theta \leq 1$.

1. Trovare uno stimatore con il metodo dei momenti e verificarne la consistenza.

2. Calcolare la stima per θ in corrispondenza del campione osservato $(-1, -1, 0, 1, -1)$.

3. con R:

   - scrivere una funzione di R di nome `dtri` con parametri `x` e `theta`, vettorializzata nel parametro `x`, che calcoli la densità $f(x;\theta)$;
   - calcolare (possibilmente analiticamente, altrimenti utilizzando la funzione `integrate`) l'integrale $F(x;\theta) = \int_{-\infty}^{x} f(z;\theta)$ e scrivere una funzione in R di nome `ptri` con caratteristiche come sopra;
   - tracciare un grafico della densità e della funzione di ripartizione.

4

Le votazioni degli esami A e B possono essere rappresentate da due v.c. normali di parametri (μ, σ_1^2), (μ, σ_2^2). Al primo appello si sono presentati un campione casuale semplice di n_1 studenti in A e di n_2 studenti in B, ottenendo per il primo gruppo una media $\overline{x}$ mentre per l'altro una media $\overline{y}$. Per stimare il voto medio complessivo viene proposto lo stimatore

$$T = a\overline{X} + (1-a)\overline{Y}$$

con $0 \leq a \leq 1$.

1. Verificare se lo stimatore è non distorto.

2. Si calcoli il valore di a che rende minima la varianza dello stimatore T nell'ipotesi che il voto nell'esame in A sia indipendente da quello in B.

5

Sia $(X_1, \ldots, X_n)$ un campione casuale semplice da una distribuzione $\mathcal{N}(\mu, \sigma^2)$.

1. Dire quale degli stimatori di μ

$$T_1 = \frac{\sum_{i=1}^n X_i}{n} = \bar{X}, \qquad T_2 = 2\bar{X} - X_1$$

   è preferibile in termini di errore quadratico medio.

2. con R:

   - scrivere due funzioni, una per l'errore quadratico medio di T_1 e una per l'errore quadratico medio di T_2 (attenzione, la funzione avrà come argomenti: la dimensione del campione, la media e la varianza);
   - per $n = 1, 2, 3, 10$, $\mu = -1, 0, 1$ e $\sigma^2 = 1$, utilizzando le funzioni al punto precedente, ottenere i grafici dell'andamento dell'errore quadratico medio per entrambi gli stimatori.

6

Sia X_1, X_2, X_3 un c.c.s. estratto da una v.c. X di valor medio w e varianza w^2, con $x > 0$ e $w > 0$. Confrontare in termini di errore quadratico medio i due stimatori di w:

$$T_1 = \frac{2X_1 + X_2 + 2X_3}{5} \quad \text{e} \quad T_2 = \frac{2X_1 + 2X_3}{4}.$$