Statistica Computazionale
Esempio di compito (1)

È richiesta una relazione contenente il codice R, debitamente commentato, necessario per risolvere i problemi posti. Le procedure usate devono essere giustificate a parole. Anche i grafici, se presenti, utilizzati nella discussione devono essere inseriti nell'elaborato.

Dove richiesto, devono essere usate 10000 replicazioni (Monte Carlo o bootstrap).

Non è possibile consultare alcun materiale, tranne l'"help" di R.

Il tempo a disposizione è di 135 minuti.

Esercizio 1

Si consideri l'integrale

$$I(K) = \int_0^K e^{-x\,(0.5+|\sin(x)|)}\,dx$$

a. Si ottenga un'approssimazione di $I(10)$ via Monte Carlo, quantificando l'errore di approssimazione.

b. Si usi la distribuzione esponenziale di media 1 (con densità pari a e^{-x}) per proporre, via campionamento per importanza, un'approssimazione per $I(\infty)$. [Si ricordi che con $\texttt{rexp}(B)$ si generano B valori dalla densità esponenziale di media 1]

c. Si definisca la funzione di densità

$$f(x) = \frac{1}{I(\infty)}\,e^{-x\,(0.5+|\sin(x)|)}\,, \qquad x > 0\,.$$

Si usi la distribuzione esponenziale di media 2 (con densità pari a $0.5\,e^{-0.5\,x}$) per costruire un algoritmo di generazione di tipo accetto/rifiuto dalla distribuzione con densità f.

Esercizio 2

Il file `esercizio_2.dat` contiene osservazioni relative a due variabili x e y, per le quali si considera un modello lineare $Y_i = \beta_0 + \beta_1\ X_i + \epsilon_i$, $i = 1, \ldots, n$, dove si assume che $\epsilon_1, \ldots, \epsilon_n$ siano i.i.d. con distribuzione $\mathcal{N}(0, \sigma^2)$. Si dubita però sia dell'assunzione di normalità che di quella di linearità.

a. Si scriva il codice R necessario per ottenere delle repliche bootstrap (bootstrap non parametrico) degli stimatori dei minimi quadrati di β_0 e β_1.

b. Si valutino, via bootstrap, distorsione e varianza dello stimatore dei minimi quadrati di β_1.

c. Si fornisca un i.c. bootstrap di grado di fiducia 0.99 per β_1 e lo si usi per verificare l'ipotesi $H_0 : \beta_1 = 0$ contro $H_1 : \beta_1 \neq 0$. Si faccia la stessa cosa con un grado di fiducia 0.9. Si commentino i risultati.

Esercizio 3

Il file `esercizio_3.dat` contiene un campione $x_1, \ldots, x_n$ che si ritiene generato da variabili casuali i.i.d. con distribuzione beta, avente funzione di densità

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)}\, x^{a-1}\,(1-x)^{b-1}\ , \qquad x \in [0,1]\ ,$$

dove $a > 0$ e $b > 0$ sono i parametri di interesse.

a. Si scriva una funzione `R` che calcola la funzione di log-verosimiglianza in corrispondenza di due generici valori di a e b.

b. Si disegni un grafico della funzione di log-verosimiglianza per $(a, b) \in [0, 10]^2$ e se ne commenti l'andamento: quali valori suggerisce il grafico per la stima di massima verosimiglianza?

c. Si ottenga la stima di massima verosimiglianza di (a, b) impiegando la funzione nlm e adottando una opportuna riparametrizzazione. Per la scelta dei valori di partenza, ci si aiuti con la rappresentazione grafica ottenuta al punto b.

d. Si ottenga un intervallo di confidenza con grado di fiducia 0.95 per a utilizzando la distribuzione asintotica dello stimatore di massima verosimiglianza. [Può essere utile la funzione fdHess nella biblioteca nlme]