Modelli Statistici 1	
A.A. 2002-2003	**Homework 2**

Esercizio 2.1. Si consideri il seguente modello di regressione lineare semplice:

$$Y_i = \beta_1 + \beta_2 x_i + \epsilon_i \qquad i = 1, \ldots, n$$

con $\epsilon_1, \ldots, \epsilon_n$ variabili casuali indipendenti ed identicamente distribuite $N(0, \sigma^2)$. Si assuma σ^2 noto. Siano $\hat{\beta}_1$ e $\hat{\beta}_2$ gli stimatori di massima verosimiglianza dei parametri β_1 e β_2. Sia inoltre $\hat{Y}_i$ la variabile casuale

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i,$$

ovvero lo stimatore del valore previsto dal modello in corrispondenza del valore x_i della variabile esplicativa.

1 Si derivi la distribuzione esatta dello stimatore $\hat{Y}_i$. (Suggerimento: si esprima $\hat{Y}_i$ in termini di combinazione lineare delle osservazioni Y_i.)

2 Fissato un valore x_k, e detto $\hat{Y}_k$ il corrispondente valore previsto dal modello, si calcoli $E(\hat{Y}_k)$.

3 Si ottenga un intervallo di confidenza per $E(Y_k)$ con livello esatto $1 - \alpha$.

Esercizio 2.2. L'isola di Vandalusia esporta alberi tropicali. I dati che seguono rappresentano l'estensione della foresta tropicale negli anni dal 1985 al 1995. L'estensione (espressa in migliaia di ettari) misura la dimensione della foresta presente sull'isola al 31 dicembre di ciascun anno.

Area	18.7	16.1	15.7	15.3	14.9	11.2	10.9	8.1	7.7	6.9	5.4
Anno	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995

Si indichi con y_i, $i = 1, \ldots, 10$, l'area misurata nell'anno x_i e si consideri il seguente modello di regressione lineare semplice:

$$y_i = \beta_1 + \beta_2 x_i + \epsilon_i \qquad i = 1, \ldots, 11$$

con $\epsilon_1, \ldots, \epsilon_{11}$ realizzazioni di variabili casuali $N(0, \sigma^2)$ indipendenti.

1 Si calcoli la stima di massima verosimiglianza di (β_1, β_2).

2 Si calcoli la stima non distorta di σ^2.

3 Si verifichino le ipotesi di nullità dei parametri β_1 e β_2, utilizzando in ciascuna verifica un livello di significatività $\alpha = 0.05$.

4 Per i test di cui al punto precedente, si calcoli il livello di significatività osservato.

5 In che anno ci si può attendere che scompaia completamente la foresta?

6 Sulla base dei dati raccolti negli anni 1985–1995, si fornisca una stima per la dimensione della foresta nell'anno 1996.

Esercizio 2.3. Si considerino i dati relativi all'esempio "I cuculi e Darwin" riportati nel lucido 5.3. Si indichi con y_i, $i = 1, \ldots, 31$, la lunghezza media delle uova di cuculo deposte nei nidi altrui e con x_i la specie dell'uccello ospite, definita come segue:

$$x_i = \begin{cases} -1 & \text{se Specie = ``pettirosso''} \\ +1 & \text{se Specie = ``scricciolo''} \end{cases}$$

Si consideri il seguente modello di regressione lineare semplice:

$$y_i = \beta_1 + \beta_2 x_i + \epsilon_i \qquad i = 1, \ldots, 31$$

con $\epsilon_1, \ldots, \epsilon_{31}$ realizzazioni di variabili casuali $N(0, \sigma^2)$ indipendenti.

1 Si ricavi lo stimatore di massima verosimiglianza per (β_1, β_2).

2 Si ricavi la distribuzione esatta di $(\hat{\beta}_1,\ \hat{\beta}_2)$.

3 Che ipotesi statistica corrisponde all'affermazione: in media la lunghezza delle uova di cuculo non cambia al variare della specie ospite (pettirosso o cuculo)?