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

Esercizio 1.1. Avendo ricevuto numerosi reclami circa la lentezza delle code alle casse di un grande supermercato, il responsabile del personale ribatte che il tempo di attesa alle casse è legato al numero di pezzi acquistati dal cliente, più che all'efficienza degli addetti alle casse. Per dimostrarlo, per 8 clienti raccoglie i seguenti dati

y_i	45	120	125	140	150	155	170	180
x_i	5	11	14	21	27	32	40	46

dove y_i rappresenta il tempo di attesa (in secondi) per il cliente i ($i = 1, \ldots, 8$) e x_i rappresenta il corrispondente numero di pezzi acquistati. Si consideri il seguente modello di regressione lineare semplice:

$$y_i = \beta_1 + \beta_2 x_i + \epsilon_i \qquad i = 1, \ldots, 8$$

con $\epsilon_1, \ldots, \epsilon_8$ realizzazioni indipendenti ed identicamente distribuite di variabili $N(0, \sigma^2)$.

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 Si calcoli il livello di significatività osservato per l'ipotesi di nullità su β_2.

5 Si costruisca un intervallo di confidenza per β_2 con livello di confidenza $1 - \alpha = 0.95$.

6 Cosa si può dire circa l'affermazione del responsabile del personale?

Esercizio 1.2. Per verificare l'effetto della vitamina C sull'accrescimento dei maiali, a 30 maiali sono state somministrate dalla nascita dosi diverse di acido ascorbico. Ad una età prefissata è stata poi misurata la lunghezza media dei denti (usata come misura della crescita). (V. unità T di Statistica Descrittiva). I dati sono i seguenti:

Dose (mg)	Lunghezza (mm)
0.5	8.2, 9.4, 9.7, 9.7, 10, 14.5, 15.2, 16.5, 17.6, 21.5
1.0	14.5, 19.7, 20.0, 21.2, 23.3, 23.6, 25.2, 25.8, 26.4, 27.3
2.0	22.4, 23.0, 24.5, 24.8, 25.5, 26.4, 26.4, 27.3, 29.4, 30.9

1 Indicata con Y la lunghezza media dei denti e con x la dose, si consideri il seguente modello di regressione lineare semplice:

$$y_i = \beta_1 + \beta_2 x_i + \epsilon_i \qquad i = 1, \dots, 30$$

con $\epsilon_1, \dots, \epsilon_{30}$ realizzazioni indipendenti ed identicamente distribuite di variabili $N(0, \sigma^2)$. Si calcoli la stima di massima verosimiglianza di (β_1, β_2).

2 Si derivi la distribuzione esatta dello stimatore $\hat{\beta}_1$ e dello stimatore $\hat{\beta}_2$.

3 Si verifichi l'ipotesi di nullità di β_2 facendo ricorso alla distribuzione esatta di $\hat{\beta}_2$.

4 Si calcoli il livello di significatività osservato per il test di cui al punto precedente.

5 Si verifichi l'ipotesi di nullità di β_2 facendo ricorso alla statistica $W_P(\beta_2)$ ed utilizzando la sua distribuzione approssimata.

Esercizio 1.3. Si consideri il seguente modello di regressione lineare semplice:

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

con $\epsilon_1, \dots, \epsilon_n$ realizzazioni indipendenti ed identicamente distribuite di variabili $N(0, \sigma^2)$. Siano $e_i(Y) = Y_i - \hat{Y}_i$, $i = 1, \dots, n$, i residui di regressione. Si dimostri che:

$$V(e_i(Y)) = \sigma^2 \left(1 - \frac{1}{n} - \frac{(x_i - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}\right).$$