MODELLO LINEARE

Sia $x_1,...,x_n$ un insieme di valori per ciascuno dei quali sia Y_i un'osservazione della v.c. $Y(x_i)$. Pertanto le n coppie $(x_1,Y_1),...,(x_n,Y_n)$ costituiscono un insieme di osservazioni tali che:

1. $E(Y_i)=\mu(x_i)=\beta_0+\beta_1 x_i$;
2. $Var(Y_i)=\sigma^2$.

Il modello lineare può essere anche specificato come segue:

$$Y_i=\mu(x_i)+\varepsilon_i=\beta_0+\beta_1 x_i+\varepsilon_i$$

dove ε_i è sono v.c. tali che

$$E(\varepsilon_i)=0; \quad Var(\varepsilon_i)=\sigma^2.$$

Distinguiamo due casi:

<u>Caso A</u>: le v.c. Y_i sono normali e congiuntamente indipendenti.

<u>Caso B</u>: le v.c. Y_i sono a due a due non correlate ossia

$$Cov(Y_i,Y_j)=0 \text{ per ogni } i\neq j=1,...,n.$$

Caso A

Stimatori di massima verosimiglianza

Siano $y_1,...,y_n$ le osservazioni delle v.c. $Y_1,...,Y_n$. La funzione di verosimiglianza è

$$L\left(y_1,...,y_n;\beta_0;\beta_1;\sigma^2\right)=p\left(y_1\right)\cdot...\cdot p\left(y_n\right)=$$

$$=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}\frac{\left(y_1-\beta_0-\beta_1x_1\right)^2}{\sigma^2}}\cdot...\cdot\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}\frac{\left(y_n-\beta_0-\beta_1x_n\right)^2}{\sigma^2}}=$$

$$=\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}}e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{\left(y_i-\beta_0-\beta_1x_i\right)^2}{\sigma^2}}.$$

Passando ai logaritmi:

$$\log L\left(y_1,...,y_n;\beta_0;\beta_1;\sigma^2\right)=$$

$$=\frac{n}{2}\log\left(\frac{1}{2\pi\sigma^2}\right)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i-\beta_0-\beta_1x_i\right)^2=$$

$$=-\frac{n}{2}\log\left(2\pi\right)-\frac{n}{2}\log\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i-\beta_0-\beta_1x_i\right)^2.$$

Le derivate parziali sono

$$\frac{\partial \log L}{\partial \beta_0} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2\left(y_i - \beta_0 - \beta_1 x_i\right)(-1);$$

$$\frac{\partial \log L}{\partial \beta_1} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2\left(y_i - \beta_0 - \beta_1 x_i\right)(-x_i);$$

$$\frac{\partial \log L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}\left(y_i - \beta_0 - \beta_1 x_i\right).$$

Uguagliando a zero le derivate si ottiene

$$\begin{cases} \sum_{i=1}^{n}\left(y_i - \beta_0 - \beta_1 x_i\right) = 0 \\ \sum_{i=1}^{n}\left(y_i - \beta_0 - \beta_1 x_i\right)x_i = 0 \\ -n\sigma^2 + \sum_{i=1}^{n}\left(y_i - \beta_0 - \beta_1 x_i\right) = 0 \end{cases}$$

Risolvendo le prime due equazioni rispetto a β_0 e β_1 si ottengono gli stimatori di massima verosimiglianza:

$$\hat{\beta}_1 = \frac{\sum\left(y_i - \overline{y}\right)\left(x_i - \overline{x}\right)}{\sum\left(x_i - \overline{x}\right)^2},$$

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x}.$$

Dalla terza equazione si ricava lo stimatore $\tilde{\sigma}^2$ per σ^2

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2.$$

Ricaviamo ora i valori attesi di $\hat{\beta}_0$ e $\hat{\beta}_1$.

$$\hat{\beta}_1 = \frac{1}{Dev\left(x\right)}\sum\left(Y_i - \overline{Y}\right)\left(x_i - \overline{x}\right) =$$

$$= \sum Y_i \frac{\left(x_i - \overline{x}\right)}{Dev\left(x\right)} - \frac{\overline{Y}}{Dev\left(x\right)}\sum\left(x_i - \overline{x}\right) =$$

$$= \sum Y_i \frac{\left(x_i - \overline{x}\right)}{Dev\left(x\right)}.$$

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{x} = \frac{1}{n} \sum Y_i - \sum_{i=1}^{n} \frac{\left(x_i - \overline{x}\right) \cdot \overline{x}}{Dev\left(x\right)} Y_i =$$

$$= \sum_{i=1}^{n} \left[\frac{1}{n} - A\left(x_i - \overline{x}\right) \right] Y_i$$

dove

$$A = \frac{\overline{x}}{Dev\left(x\right)}.$$

Allora

$$\bullet E\left(\hat{\beta}_1\right) = \sum \frac{\left(x_i - \overline{x}\right)}{Dev\left(x\right)} E\left(Y_i\right) =$$

$$= \sum \frac{\left(x_i - \overline{x}\right)}{Dev\left(x\right)} \left(\beta_0 + \beta_1 x_i\right) =$$

$$\frac{\beta_0}{Dev\left(x\right)} \sum \left(x_i - \overline{x}\right) + \frac{\beta_1}{Dev\left(x\right)} \sum \left(x_i - \overline{x}\right) x_i =$$

$$= \frac{\beta_1}{Dev\left(x\right)} Dev\left(x\right) = \beta_1.$$

Quindi $\hat{\beta}_1$ è uno stimatore non distorto di β_1.

$$\bullet E\left(\hat{\beta}_0\right)=\sum_{i=1}^{n}\left[\frac{1}{n}-A\left(x_i-\bar{x}\right)\right]E\left(Y_i\right)=$$

$$=\sum_{i=1}^{n}\left[\frac{1}{n}-A\left(x_i-\bar{x}\right)\right]\left(\beta_0+\beta_1 x_i\right)=$$

$$=\beta_0+\beta_1\bar{x}-A\beta_0\sum\left(x_i-\bar{x}\right)-A\beta_1\sum\left(x_i-\bar{x}\right)=$$

$$=\beta_0+\beta_1\bar{x}-A\beta_1 Dev\left(x\right)=$$

$$=\beta_0+\beta_1\bar{x}-\frac{\bar{x}}{Dev\left(x\right)}\beta_1 Dev\left(x\right)=\beta_0$$

Quindi $\hat{\beta}_0$ è uno stimatore non distorto di β_0.

$$\bullet Var\left(\hat{\beta}_1\right)=\sum_{i=1}^{n}\frac{\left(x_i-\bar{x}\right)^2}{\left[Dev\left(x\right)\right]^2}Var\left(Y_i\right)=\frac{\sigma^2}{nVar\left(x\right)}.$$

$$\bullet Var\left(\hat{\beta}_0\right)=\sum_{i=1}^{n}\left[\frac{1}{n}-A\left(x_i-\bar{x}\right)\right]^2 Var\left(Y_i\right)=$$

$$=\sigma^2\sum_{i=1}^{n}\left[\left(\frac{1}{n}\right)^2+A^2\left(x_i-\bar{x}\right)^2-2\frac{A\left(x_i-\bar{x}\right)}{n}\right]=$$

$$=\frac{\sigma^2}{n}+\sigma^2 A^2 Dev\left(x\right)-2\sigma^2\frac{A}{n}\sum\left(x_i-\bar{x}\right)=$$

$$=\frac{\sigma^2}{n}+\sigma^2\frac{\bar{x}^2}{\left[Dev(x)\right]^2}Dev(x)=$$

$$=\frac{\sigma^2}{n}+\frac{\sigma^2\bar{x}^2}{nVar(x)}=\frac{\sigma^2}{n}\left[1+\frac{\bar{x}^2}{Var(x)}\right]=$$

$$=\frac{\sigma^2}{n}\left\{1+\frac{1}{\left[\frac{\sigma(x)}{\bar{x}}\right]^2}\right\}=\frac{\sigma^2}{n}\left\{1+\left[\frac{1}{C.V.(x)}\right]^2\right\}=$$

$$=\frac{\sigma^2}{n}+\frac{\sigma^2}{n}\frac{\bar{x}^2}{Var(x)}=Var\left(\bar{Y}\right)+\bar{x}^2Var\left(\hat{\beta}_1\right).$$

Questo significa che la v.c. $\bar{Y}$ è in correlata con la v.c. $\hat{\beta}_1$.

Si può inoltre dimostrare che

$$Cov\left(\hat{\beta}_0,\hat{\beta}_1\right)=-\frac{\sigma^2\cdot\bar{x}}{nVar(x)}.$$

Poiché $\hat{\beta}_0$ e $\hat{\beta}_1$ sono somme ponderate di v.c. normali indipendenti hanno entrambi distribuzione normale. Si dimostra infine che la coppia $\left(\hat{\beta}_0, \hat{\beta}_1\right)$ ha una distribuzione bivariata normale.

Studiamo ora la distribuzione campionaria di

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2.$$

Si dimostra che

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2 \sim \chi^2_{(n-2)}.$$

Pertanto

$$E\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2\right] = (n-2),$$

$$E\left[\sum_{i=1}^{n}\left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2\right] = \sigma^2 (n-2),$$

$$E\left[\frac{1}{(n-2)}\sum_{i=1}^{n}\left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2\right] = \sigma^2.$$

Indichiamo allora con $\hat{\sigma}^2$ lo stimatore corretto

$$\hat{\sigma}^2 = \frac{1}{(n-2)}\sum_{i=1}^{n}\left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2.$$

La v.c.

$$\frac{(n-2)\hat{\sigma}^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2 \sim \chi^2_{(n-2)}$$

Si può inoltre dimostrare che $\hat{\sigma}^2$ è indipendente sia da $\hat{\beta}_0$ sia da $\hat{\beta}_1$.

Intervalli di confidenza

1) *Intervallo di confidenza per* σ^2.

Poiché $\frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{(n-2)}$, allora

$$P\left[\chi^2_{\frac{\alpha}{2}} \leq \frac{(n-2)\hat{\sigma}^2}{\sigma^2} \leq \chi^2_{1-\frac{\alpha}{2}}\right] = 1-\alpha,$$

da cui ricaviamo l'intervallo di confidenza per σ^2:

$$\boxed{\frac{(n-2)\hat{\sigma}^2}{\chi^2_{1-\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{(n-2)\hat{\sigma}^2}{\chi^2_{\frac{\alpha}{2}}}}.$$

2) *Intervallo di confidenza per* β_0.

Poiché $\hat{\beta}_0$ ha una distribuzione normale di media β_0 e varianza $\frac{\sigma^2}{n}\left[1+\frac{\bar{x}^2}{Var(x)}\right]$, la v.c.

$$Z=\frac{\hat{\beta}_0-\beta_0}{\sqrt{\frac{\sigma^2}{n}\left[1+\frac{\bar{x}^2}{Var(x)}\right]}}\sim N(0;1).$$

Poiché

$$U=\frac{(n-2)\hat{\sigma}^2}{\sigma^2}\sim\chi^2_{(n-2)},$$

e U e Z sono indipendenti, il rapporto

$$T=\frac{Z}{\sqrt{\frac{U}{n-2}}}=\frac{\frac{\hat{\beta}_0-\beta_0}{\sqrt{\frac{\sigma^2}{n}\left[1+\frac{\bar{x}^2}{Var(x)}\right]}}}{\sqrt{\frac{(n-2)\hat{\sigma}^2}{\sigma^2(n-2)}}}=\frac{\hat{\beta}_0-\beta_0}{\sqrt{\frac{\hat{\sigma}^2}{n}\left[1+\frac{\bar{x}^2}{Var(x)}\right]}}$$

è distribuito come una v.c. **t** con *(n-2)* g.d.l.
Pertanto l'intervallo di confidenza di β_0 risulta

$$\boxed{\hat{\beta}_0 \pm t_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{\sigma}^2}{n}\left[1+\frac{\bar{x}^2}{Var(x)}\right]}}$$

3) *Intervallo di confidenza per* $\mu(x) = \beta_0 + \beta_1 x.$

Consideriamo come stimatore di $\mu(x)$

$$\hat{\mu}(x) = \hat{\beta}_0 + \hat{\beta}_1 x =$$

$$= \sum_{i=1}^{n}\left[\frac{1}{n} - A\left(x_i - \bar{x}\right)\right]Y_i + \left[\sum_{i=1}^{n} Y_i \frac{x_i - \bar{x}}{Dev(x)}\right]x =$$

$$= \sum_{i=1}^{n}\left[\frac{1}{n} - \frac{\bar{x}}{Dev(x)}\left(x_i - \bar{x}\right) + \frac{x_i - \bar{x}}{Dev(x)} \cdot x\right]Y_i =$$

$$= \sum_{i=1}^{n}\left[\frac{1}{n} + \frac{\left(x - \bar{x}\right)\left(x_i - \bar{x}\right)}{Dev(x)}\right]Y_i.$$

Essendo $\hat{\mu}(x)$ una somma ponderata delle Y_i, essa segue una distribuzione normale. Inoltre

$$E\left[\hat{\mu}(x)\right]=\sum_{i=1}^{n}(\beta_0+\beta_1 x_i)\left[\frac{1}{n}+\frac{\left(x-\overline{x}\right)\left(x_i-\overline{x}\right)}{Dev(x)}\right]=$$

$$=\beta_0+\beta_1\overline{x}+\beta_0\frac{x-\overline{x}}{Dev(x)}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)+$$

$$+\beta_1\frac{x-\overline{x}}{Dev(x)}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)x_i=$$

$$=\beta_0+\beta_1\overline{x}+\beta_1\frac{x-\overline{x}}{Dev(x)}Dev(x)=\beta_0+\beta_1 x.$$

Quindi $\hat{\mu}(x)$ è uno stimatore corretto di $\mu(x)$.

$$Var\left[\hat{\mu}(x)\right]=\sum_{i=1}^{n}Var(Y_i)\left[\frac{1}{n}+\frac{\left(x-\overline{x}\right)\left(x_i-\overline{x}\right)}{Dev(x)}\right]^2=$$

$$=\sigma^2\sum_{i=1}^{n}\left[\frac{1}{n}+\frac{\left(x-\overline{x}\right)\left(x_i-\overline{x}\right)}{Dev(x)}\right]^2=$$

$$= \sigma^2 \sum_{i=1}^{n} \left\{ \frac{1}{n^2} + \frac{\left(x-\bar{x}\right)^2 \left(x_i-\bar{x}\right)^2}{\left[Dev(x)\right]^2} + \frac{2}{n} \frac{\left(x-\bar{x}\right)\left(x_i-\bar{x}\right)}{Dev(x)} \right\}^2 =$$

$$= \sigma^2 \sum_{i=1}^{n} \left(\frac{1}{n}\right)^2 + \sigma^2 \frac{\left(x-\bar{x}\right)^2}{\left[Dev(x)\right]^2} \sum_{i=1}^{n} \left(x_i-\bar{x}\right)^2 +$$

$$+ \frac{2\sigma^2}{n} \frac{\left(x-\bar{x}\right)}{Dev(x)} \sum_{i=1}^{n} \left(x_i-\bar{x}\right) =$$

$$= \frac{\sigma^2}{n} + \sigma^2 \frac{\left(x-\bar{x}\right)^2}{Dev(x)} = \frac{\sigma^2}{n} + \sigma^2 \frac{\left(x-\bar{x}\right)^2}{nVar(x)} = \frac{\sigma^2}{n} \left[1 + \frac{\left(x-\bar{x}\right)^2}{Var(x)}\right].$$

Deriva che

$$Z = \frac{\hat{\mu}(x) - (\beta_0 + \beta_1 x)}{\sqrt{\frac{\sigma^2}{n}\left[1 + \frac{\left(x-\bar{x}\right)^2}{Var(x)}\right]}} \sim N(0;1)$$

$$T = \frac{\hat{\mu}(x) - (\beta_0 + \beta_1 x)}{\sqrt{\frac{\hat{\sigma}^2}{n}\left[1 + \frac{(x - \bar{x})^2}{Var(x)}\right]}} \sim t_{(n-2)}$$

L'intervallo di confidenza per $\mu(x) = \beta_0 + \beta_1 x$ è

$$\boxed{\hat{\mu}(x) \pm t_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{\sigma}^2}{n}\left[1 + \frac{(x - \bar{x})^2}{Var(x)}\right]}}.$$

Verifiche di ipotesi

1) $H_0 : \beta_0 = 0$ contro $\beta_0 \neq 0$.
Basiamo il test sulla statistica

$$T = \frac{\hat{\beta}_0 - 0}{\sqrt{\frac{\hat{\sigma}^2}{n}\left[1 + \frac{\bar{x}^2}{Var(x)}\right]}} \underset{H_0}{\sim} t_{(n-2)}$$

Respingiamo H_0 per valori $|T| > t_{1-\frac{\alpha}{2}}$.

2) $H_0 : \beta_0 \geq 0$ contro $\beta_0 < 0$: respingiamo H_0 se $T < -t_{(1-\alpha)}$.

3) $H_0 : \beta_0 \leq 0$ contro $\beta_0 > 0$: respingiamo H_0 se $T > t_{\left(1-\frac{\alpha}{2}\right)}$.

4) $H_0 : \beta_1 = 0$ contro $\mathrm{H}_1 : \beta_1 \neq 0$: utilizziamo la statistica

$$T = \frac{\hat{\beta}_1 - 0}{\sqrt{\frac{\hat{\sigma}^2}{n} \frac{1}{Var(x)}}} \underset{H_0}{\sim} t_{(n-2)},$$

respingiamo H_0 per valori $|T| > t_{1-\frac{\alpha}{2}}$.

5) $H_0 : \beta_1 = 2$ contro $\mathrm{H}_1 : \beta_1 > 2$: utilizziamo la statistica

$$T = \frac{\hat{\beta}_1 - 2}{\sqrt{\frac{\hat{\sigma}^2}{n} \frac{1}{Var(x)}}} \underset{H_0}{\sim} t_{(n-2)}$$

e respingiamo H_0 per valori $T > t_{(1-\alpha)}$.

Caso B

In questo caso supponiamo le $Y_1,...,Y_n$ incorrelate. Inoltre $E(Y_i) = \beta_0 + \beta_1 x_i$ e $Var(Y_i) = \sigma^2$. Per ottenere uno stimatore di β_0 e β_1 utilizziamo il metodo dei minimi quadrati.

Definizione (*stimatori a minimi quadrati per* β_0 e β_1)

Siano n coppie di osservazioni (Y_i, x_i), $i = 1,...,n$ che soddisfino i requisiti del caso B. I valori di β_0 e β_1 che minimizzano la quantità

$$\sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 x_i)^2$$

sono chiamati stimatori a minimi quadrati per β_0 e β_1. E' noto che tali stimatori sono

$$\hat{\beta}_1 = \frac{\sum\left(Y_i - \overline{Y}\right)\left(x_i - \overline{x}\right)}{Dev\left(x\right)}, \quad \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{x}.$$

Dopo aver ricavato $\hat{\beta}_0$ e $\hat{\beta}_1$ possiamo calcolare il seguente stimatore per σ^2:

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2.$$

Anche se gli stimatori ottenuti nel caso A sono gli stessi che si ottengono nel caso B, nel caso A essi godono di molte proprietà poiché sono stimatori di massima verosimiglianza. Vediamo ora le proprietà degli stimatori a minimi quadrati.

Teorema di Gauss-Markov

Nel caso B gli stimatori a minimi quadrati $\hat{\beta}_0$ e $\hat{\beta}_1$ sono gli stimatori migliori nella classe degli stimatori lineari non distorti (ossia sono quelli a varianza minore).

Esercizio

X: diminuzione di consumo conseguita, in Watt
Y: durata di accensione, in ore (tolte le 400 iniziali).
Sono stati ottenuti i seguenti risultati:
$\bar{x} = 34 \qquad \bar{y} = 11 \qquad \sum x_i^2 = 37568$
$\sum y_i^2 = 3967 \qquad \sum x_i y_i = 11760.$
Assumendo valide le ipotesi del modello lineare caso A:

a) Stimare i parametri β_0 e β_1
b) Stimare la varianza dello stimatore $\hat{\beta}_1$
c) Determinare l'intervallo di confidenza al 98% per β_1
d) Determinare l'intervallo di confidenza al 90% per $\mu(x)$, in corrispondenza di $x = 30$
e) Verificare l'ipotesi $H_0 : \beta_0 = 22,$ con alternativa bilaterale, volendo commettere

l'errore di prima specie con probabilità del 10%

f) Verificare l'ipotesi $H_0 : \beta_1 \geq -0,3$ con alternativa unilaterale, volendo commettere l'errore di prima specie con probabilità dell'1%

g) Determinare l'intervallo di confidenza all'80% per la varianza di *Y*.

<u>Soluzione</u>

a) $$\hat{\beta}_1 = \frac{Cov(X,Y)}{Var(X)} = \frac{\frac{\sum x_i y_i}{n} - \overline{xy}}{\frac{\sum x_i^2}{n} - \overline{x}^2} = \frac{\frac{11760}{32} - 34 \cdot 11}{\frac{37568}{32} - 34^2} =$$

$$\frac{-6,5}{18} = -0,36\overline{1}$$

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x} = 11 + 0,36\overline{1} \cdot 34 = 23,2\overline{7}$$

$$\hat{y} = 23,2\overline{7} - 0,36\overline{1} \cdot x$$

b) $$Var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{n} \cdot \frac{1}{Var(x)}$$

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2$$

$$\sum\left(y_i-\hat{\beta}_0-\hat{\beta}_1x_i\right)^2=Dev(Y)\left(1-I_d^2\right)=$$

$$=\left(3967-32\cdot 11^2\right)\left(1-\frac{(-6,5)^2}{18\left(\frac{3967-32\cdot 11^2}{32}\right)}\right)=$$

$$95\cdot\left(1-\frac{42,25}{53,4375}\right)=19,\overline{8}$$

$$\hat{\sigma}^2=\frac{19,\overline{8}}{32-2}=0,663$$

$$\left[Var\left(\hat{\beta}_1\right)\right]=\frac{\hat{\sigma}^2}{n}\cdot\frac{1}{Var(x)}=\frac{0,663}{32\cdot 18}=1,095$$

c) Esso ha per estremi i valori

$$\hat{\beta}_1\mp t_{0,99}(30)\sqrt{\frac{\hat{\sigma}^2}{nVar(x)}}=-0,36\overline{1}\mp 2,457\sqrt{1,095}$$

$$[-2,9321;2,2101]$$

d) $\hat{\mu}(30)=23,2\overline{7}-0,361\cdot 30=12,44\overline{7}$

L'intervallo di confidenza ha per estremi i valori

$$\hat{\mu}(30)\mp t_{0,95}(30)\sqrt{\frac{\hat{\sigma}^2}{n}\left[1+\frac{(x-\bar{x})^2}{Var(x)}\right]}=$$

$$12{,}44\overline{7} \mp 1{,}697\sqrt{\frac{0{,}663}{32}\left[1+\frac{(30-34)^2}{18}\right]} =$$

$[12{,}1121;12{,}7835]$

e) Rifiuto H_0 se

$$\frac{\left|\hat{\beta}_0-\beta_0\right|}{\sqrt{\frac{\hat{\sigma}^2}{n}\left(1+\frac{\bar{x}^2}{Var(x)}\right)}} > t_{1-\frac{\alpha}{2}}(n-2)$$

Poichè

$$\frac{\left|23{,}2\overline{7}-22\right|}{\sqrt{\frac{0{,}663}{32}\left(1+\frac{34^2}{18}\right)}} = 1{,}0992 < 1{,}697,$$

accetto l'ipotesi nulla.

f) Rifiuto H_0 se

$$\frac{\hat{\beta}_1-\beta_1}{\sqrt{\frac{\hat{\sigma}^2}{n}\frac{1}{Var(x)}}} < -t_{1-\alpha}(n-2)$$

Poichè

$\frac{-0{,}361+0{,}3}{\sqrt{1{,}095}} = -0{,}0584 > -2{,}457,$ accetto l'ipotesi nulla.

g) Esso ha per estremi i valori

$$\frac{\sum\left(y_i-\hat{\beta}_0-\hat{\beta}_1 x_i\right)^2}{\chi^2_{1-\frac{\alpha}{2}}(n-2)};\frac{\sum\left(y_i-\hat{\beta}_0-\hat{\beta}_1 x_i\right)^2}{\chi^2_{\frac{\alpha}{2}}(n-2)}$$

$$\left[\frac{19,\overline{8}}{40,3}=0,4935;\frac{19,\overline{8}}{15}=1,3259\right].$$