

©I diritti d'autore sono riservati. Ogni sfruttamento commerciale non autorizzato sarà perseguito.

Cognome, Nome e Numero di matricola:

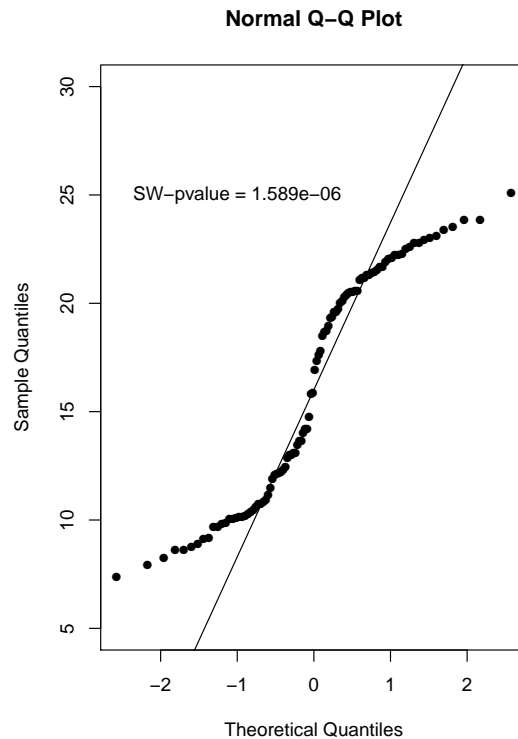
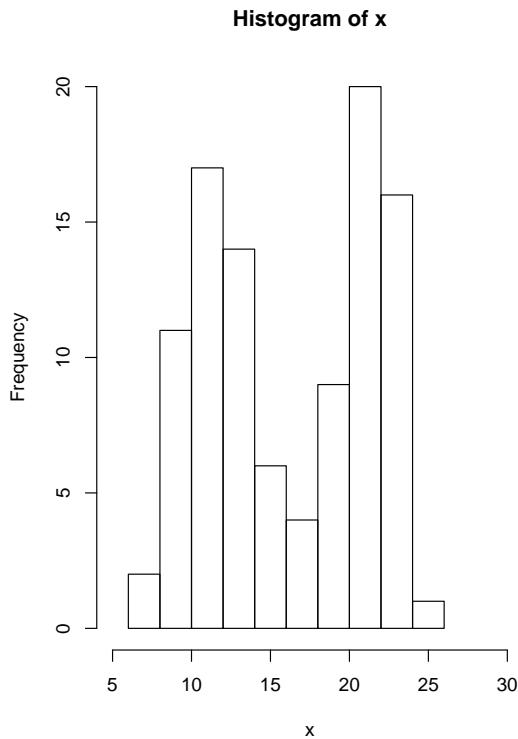
Problema 1. I primi Re di Gondor sapevano bene come è distribuita l'altezza X dei servitori di Sauron. Il cuore degli uomini però si corrompe facilmente e alcune cose che non avrebbero dovuto essere dimenticate andarono perdute. Di quell'antico sapere al Sovrintendente Denethor II rimane solo la varianza: misurando l'altezza in magli, si ha $\sigma^2 = 29$.

Denethor incarica quindi il figlio Faramir di trovare almeno una buona stima della altezza media μ dei servitori di Sauron. Faramir deve quindi catturare un campione casuale di servitori di Sauron, misurarne l'altezza, e tornare dal padre con una stima intervallare al 99% di μ , bilaterale, dove l'errore deve essere di 1 maglia al massimo.

(a) Supponendo normale la distribuzione di X , quanti servitori di Sauron deve catturare Faramir?

Alla fine Faramir cattura solo 100 servitori di Sauron, trovando:

$$\bar{x}_{100} = 16.1557, \quad s_{100}^2 = 27.78307,$$



(b) Fornire un stima intervallare bilaterale al 99% di μ .

(c) Siete sicuri che il livello di confidenza dell'intervallo calcolato sia sotto controllo? Perché?

Al suo ritorno Faramir apprende che, secondo un'antica pergamena appena ritrovata nella biblioteca di Minas Tirith, l'altezza media dei servitori di Sauron sarebbe di 15 magli. O almeno lo era ai tempi descritti dalla pergamena.

(d) Sulla base dei servitori di Sauron catturati da Faramir, potete confermare o smentire con una significatività dell'1% quanto affermato dalla pergamena?

Formulate un opportuno test statistico specificando ipotesi nulla, ipotesi alternativa, regione critica, conclusione (H_0 o H_1), forza della conclusione (debole o forte).

Risultati.

- (a) $n \geq \sigma^2 z_{0.005}^2 = 29 \cdot 2.576^2 = 192.4 \iff n \geq 193$.
- (b) $\bar{x}_{100} \pm \frac{\sigma}{\sqrt{100}} z_{0.005} = 16.16 \pm 1.39$, ovvero $(14.77, 17.55)$.
- (c) Il livello di confidenza dell'intervallo calcolato è **sotto controllo**, cioè è pari almeno approssimativamente al 99%. Infatti, anche se il test di Shapiro-Wilk, il normal probability plot e l'istogramma del campione catturato smentiscono la normalità dell'altezza dei servitori di Sauron (l'istogramma evidenzia addirittura una bimodalità), tuttavia **il campione è sufficientemente numeroso** ($n = 100$) per ritenere valida l'approssimazione normale per la media campionaria (grazie al Teorema Centrale del Limite).
- (d)

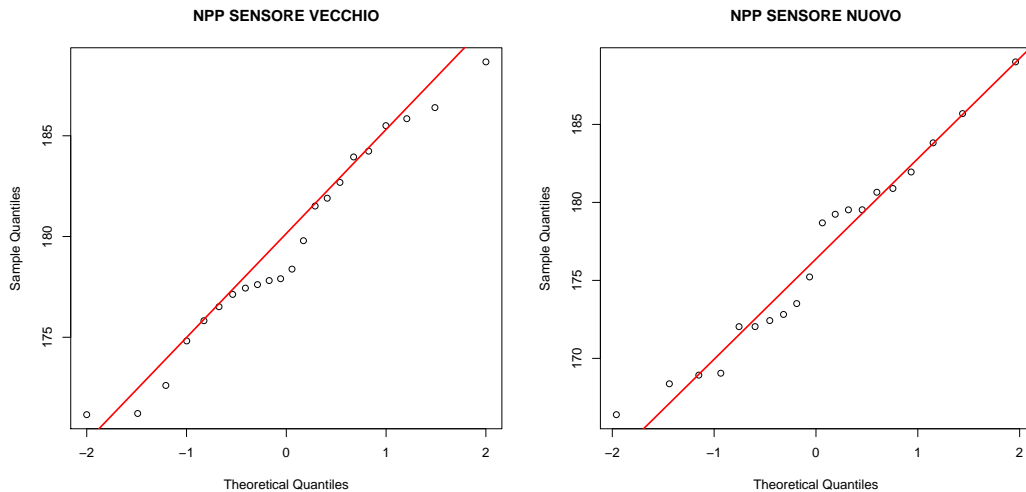
$$H_0 : \mu = 15, \quad H_1 : \mu \neq 15, \quad R : \left| \bar{x}_{100} - 15 \right| > \frac{\sigma}{\sqrt{100}} z_{0.005} \iff 15 \notin \bar{x}_{100} \pm \frac{\sigma}{\sqrt{100}} z_{0.005}$$

Poichè $15 \in 16.16 \pm 1.39$, i dati raccolti non ci consentono di rifiutare H_0 con una significatività dell'1%, i dati cioè non smentiscono la pergamena: i servitori di Sauron hanno una altezza media di 15 magli. La conclusione è debole.

Problema 2. Bill Right, consulente riduzione costi dell'azienda Popolauto, ha proposto alla direzione produzione di introdurre un nuovo sensore per misurare le emissioni di ossidi di azoto (NO_x) più economico del precedente. Bill è riuscito a convincere tutti i membri della direzione produzione, ma Richard Bigriver, un ingegnere del controllo qualità, sospetta che Bill voglia falsificare i dati sulle emissioni. Richard vuole quindi confrontare le rilevazioni fornite dai due sensori, ma questi non possono essere montati contemporaneamente su una medesima automobile. Pertanto decide di eseguire 42 prove indipendenti, tutte su una stessa automobile, sottoposta sempre alle stesse condizioni, rilevando però le emissioni 22 volte col vecchio sensore e 20 volte con quello nuovo. Richard ottiene così le seguenti medie e varianze campionarie:

$$\begin{aligned}\bar{x}_{\text{old}} &= 179.5 \text{ mg/km} & s_{\text{old}}^2 &= 24.8 \text{ (mg/km)}^2 \\ \bar{x}_{\text{new}} &= 176.5 \text{ mg/km} & s_{\text{new}}^2 &= 39.6 \text{ (mg/km)}^2\end{aligned}$$

Inoltre i p-value del test di Shapiro-Wilk sono 0.6001 per i dati sul vecchio sensore e 0.5732 per quelli del nuovo, mentre i loro Normal Probability Plot sono:



- (a) Si può assumere la normalità delle rilevazioni per entrambi i sensori?
- (b) Verificare se i dati raccolti consentono di ritenere uguali le varianze delle rilevazioni dei due sensori.
Specificare: popolazioni (o variabili aleatorie) oggetto della inferenza, ipotesi nulla, ipotesi alternativa, regione critica, condizioni di applicabilità del test, p-value dei dati, conclusione (H_0 o H_1), forza della conclusione (forte o debole).
- (c) Verificare se i dati raccolti consentono di sostenere, con conclusione forte ad un livello di significatività del 5%, che il nuovo sensore rilevi mediamente una quantità di NO_x minore rispetto al vecchio sensore.
Specificare: popolazioni (o variabili aleatorie) oggetto della inferenza, ipotesi nulla, ipotesi alternativa, regione critica, condizioni di applicabilità del test, conclusione.

Risultati.

- (a) Dal qq-plot vediamo che per entrambi i sensori i punti si distribuiscono omogeneamente vicino alla retta. Inoltre i p-value dello Shapiro-Wilk test sono molto alti e ci permettono di assumere la normalità delle rilevazioni per entrambi i sensori.
- (b)
- X_{old} = rilevazione col **vecchio** sensore dell'ossidio di azoto emesso dall'automobile utilizzata da Richard durante una prova, effettuata nelle condizioni utilizzate da Richard.
 - X_{new} = rilevazione col **nuovo** sensore dell'ossidio di azoto emesso dall'automobile utilizzata da Richard durante una prova, effettuata nelle condizioni utilizzate da Richard.

Test:

$$H_0 : \sigma_{\text{old}}^2 = \sigma_{\text{new}}^2, \quad H_1 : \sigma_{\text{old}}^2 \neq \sigma_{\text{new}}^2, \quad R_\alpha : \left\{ \frac{s_{\text{old}}^2}{s_{\text{new}}^2} < f_{1-\frac{\alpha}{2}}(21, 19) \right\} \cup \left\{ \frac{s_{\text{old}}^2}{s_{\text{new}}^2} > f_{\frac{\alpha}{2}}(21, 19) \right\}$$

Possiamo usare questa regione critica perché per il punto (a) abbiamo

$$X_{\text{old}} \sim N(\mu_{\text{old}}, \sigma_{\text{old}}^2), \quad X_{\text{new}} \sim N(\mu_{\text{new}}, \sigma_{\text{new}}^2).$$

Per i dati raccolti $\frac{s_{\text{old}}^2}{s_{\text{new}}^2} = 0.63 < 1$, per cui trovo il p-value risolvendo

$$\begin{aligned} \frac{s_{\text{old}}^2}{s_{\text{new}}^2} = 0.63 &= f_{1-\frac{\alpha}{2}}(21, 19) = \frac{1}{f_{\frac{\alpha}{2}}(19, 21)} \\ 0.56 &= \frac{1}{1.78} = \frac{1}{f_{0.1}(20, 21)} \simeq \frac{1}{f_{0.1}(19, 21)} < 0.63 = \frac{1}{f_{\frac{\alpha}{2}}(19, 21)} < \frac{1}{f_{0.25}(19, 21)} \simeq \frac{1}{f_{0.25}(20, 21)} = \frac{1}{1.35} = 0.74 \\ 0.2 < \alpha &= \text{p-value} < 0.5 \end{aligned}$$

Concludiamo dunque che le varianze possono essere supposte uguali agli usuali livelli di significatività. Conclusione debole (ma almeno p-value superiore al 20%).

- (c) Vogliamo avere evidenza che il nuovo sensore misuri mediamente una quantità più bassa di NO_x , quindi per le popolazioni X_{old} e X_{new} già introdotte al punto (b) eseguiamo il seguente test:

$$H_0 : \mu_{\text{old}} \leq \mu_{\text{new}}, \quad H_1 : \mu_{\text{old}} > \mu_{\text{new}}, \quad R_\alpha : t_0 > t_{0.05}(40),$$

dove la statistica test è

$$t_0 = \frac{\bar{x}_{\text{old}} - \bar{x}_{\text{new}}}{s_P \sqrt{1/22 + 1/20}}$$

con

$$s_P^2 = \frac{21 s_{\text{old}}^2 + 19 s_{\text{new}}^2}{40}.$$

Possiamo usare questa regione critica perché, per i punti (a) e (b), confrontiamo due variabili normali con varianze uguali,

$$X_{\text{old}} \sim N(\mu_{\text{old}}, \sigma^2), \quad X_{\text{new}} \sim N(\mu_{\text{new}}, \sigma^2),$$

Per il livello scelto e i dati raccolti otteniamo

$$t_0 = \frac{\bar{x}_{\text{old}} - \bar{x}_{\text{new}}}{s_P \sqrt{1/22 + 1/20}} = 1.72 > t_{0.05}(40) = 1.684,$$

quindi il campione appartiene alla regione critica e rifiuto H_0 : al 5% c'è evidenza che il nuovo sensore fornisca misurazioni mediamente più basse.

Problema 3. Il noto birrifico Biondetti è interessato a prevedere la quantità di alcool Y nelle proprie birre sulla base delle caratteristiche del processo di fermentazione. A tale scopo, oltre alle misurazioni della gradazione alcolica di 185 birre, registra la temperatura durante il processo di fermentazione T e la durata di tale processo D . In Figura 1 sono mostrati gli scatterplot relativi alle coppie di variabili.

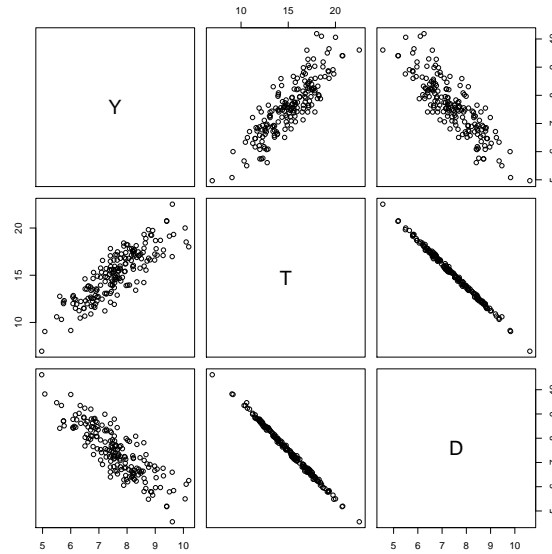


Figura 1: Scatterplot delle variabili

Per prevedere Y vengono elaborati due modelli empirici gaussiani di regressione lineare:

- Y su T e D (modello completo, Figure 2 e 3),
- Y su T (modello ridotto, Figure 4 e 5).

- (a) Scrivere la relazione tra le variabili ipotizzata dai due modelli.
- (b) Scrivere, per i entrambi i modelli, quanto vale la variazione della percentuale alcolica media delle bottiglie di birra se la temperatura viene aumentata di 1°C e la durata viene tenuta costante.
- (c) Valutare la bontà di entrambi i modelli e spiegare quale si preferisce.
- (d) Quale relazione c'è fra gli scatterplot di Figura 1 e le limitazioni del modello scartato al punto (c)?
- (e) Fornire una stima puntuale della percentuale alcolica media delle bottiglie di birra prodotte tramite un processo di fermentazione a temperatura 10°C di durata di 9 giorni.
- (f) I dati raccolti permettono di affermare che, aumentando la temperatura di 1° , si ha un aumento della percentuale alcolica media delle bottiglie di birra superiore a 0.32?

Rispondere tramite un opportuno test di ipotesi, specificando: ipotesi nulla, ipotesi alternativa, regione critica, condizioni di applicabilità del test, p-value dei dati, conclusione.

```
Call:
lm(formula = Y ~ T + D)

Residuals:
    Min       1Q   Median       3Q      Max
-1.22601 -0.37101 -0.03651  0.31084  1.64710

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.49482    9.07014   0.165   0.869
T            0.36114    0.26923   1.341   0.181
D            0.08681    0.67150   0.129   0.897

Residual standard error: 0.5278 on 182 degrees of freedom
Multiple R-squared:  0.7169,    Adjusted R-squared:  0.7138
F-statistic: 230.4 on 2 and 182 DF,  p-value: < 2.2e-16
```

Figura 2: Summary modello completo

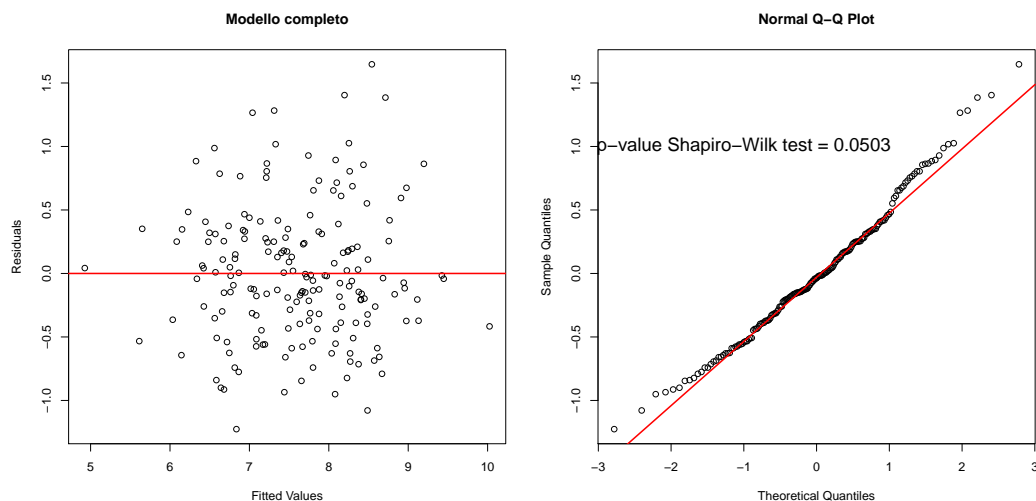


Figura 3: Grafici residui modello completo

```

Call:
lm(formula = Y ~ T)

Residuals:
    Min       1Q   Median       3Q      Max
-1.22441 -0.37099 -0.03221  0.30971  1.64376

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.66694    0.23265   11.46  <2e-16 ***
T            0.32639    0.01516   21.52  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5264 on 183 degrees of freedom
Multiple R-squared:  0.7168,    Adjusted R-squared:  0.7153
F-statistic: 463.3 on 1 and 183 DF,  p-value: < 2.2e-16

```

Figura 4: Summary modello ridotto

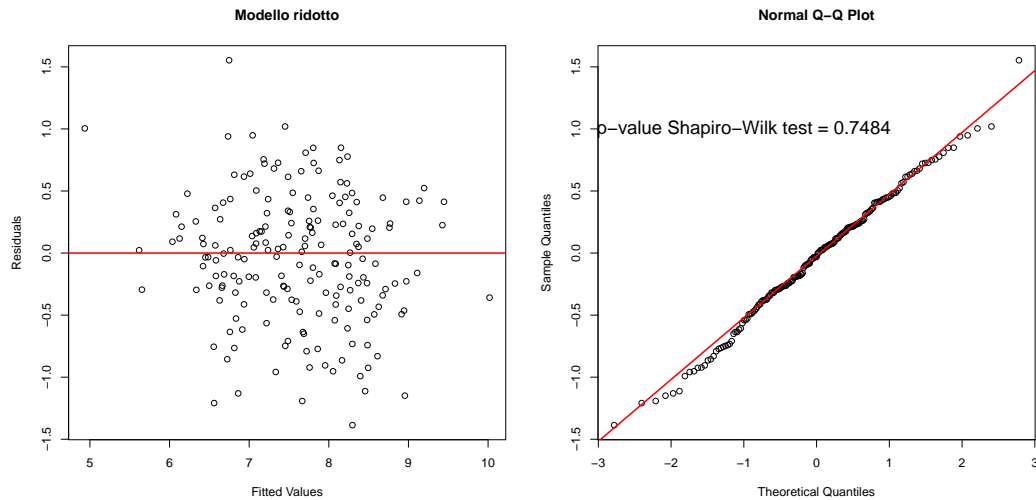


Figura 5: Grafici residui modello ridotto

Risultati.

- (a) • modello completo: $Y = \beta_0 + \beta_1 T + \beta_2 D + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma^2),$
 • modello ridotto: $Y = \beta_0 + \beta_1 T + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma^2).$
- (b) • modello completo: $\mathbb{E}[Y|T+1, D] - \mathbb{E}[Y|T, D] = \beta_1,$
 • modello ridotto: $\mathbb{E}[Y|T+1, D] - \mathbb{E}[Y|T, D] = \beta_1.$
- (c) Modello completo. L'indice $R_{\text{aggiustato}}^2 = 0.7138$ è abbastanza alto. Ipotesi gaussiana dubbia: anche se l'omoschedasticità dei residui può essere assunta, dato che i residui si distribuiscono in maniera omogenea attorno allo zero come si vede nel grafico a sinistra della Figura 3, tuttavia l'ipotesi viene messa in dubbio dal NPP e soprattutto dal p-value basso (0.0525) del test di Shapiro Wilks. Ad ogni modo, se anche con test al 5% non rinunciassimo a tale ipotesi, ci troveremmo con un modello globalmente significativo (p-value molto basso, zero numerico), ma con intercetta (p-value = 0.869) e il predittore D (p-value = 0.897) poco significativi. Pertanto dovremmo valutare l'eliminazione prima di D , poi eventualmente dell'intercetta.

Modello ridotto. L'indice $R_{\text{aggiustato}}^2 = 0.7153$ è leggermente più alto. Ipotesi gaussiana confermata: l'omoschedasticità dei residui può essere assunta, dato che i residui si distribuiscono in maniera omogenea attorno allo zero come si vede nel grafico a sinistra della Figura 5, così come il test di Shapiro Wilks fornisce un p-value elevato (0.7484) e non notiamo particolari problemi nel qq-plot. Infine intercetta e predittore risultano entrambi molto significativi (p-value molto bassi, zero numerico). Pertanto, eliminata D dal modello completo, non si deve più procedere nella selezione.

Il modello ridotto risulta quindi preferibile sotto tutti i punti di vista.

- (d) Lo scatterplot di Figura 1 evidenzia una forte correlazione tra i predittori T e D . Si capisce quindi che, in presenza del predittore T , il secondo predittore D possa risultare poco utile nella previsione del responso Y .
- (e) Usiamo la stima del modello ridotto:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 T = 2.667 + 0.326 \cdot T$$

e sostituendo $T = 10$ otteniamo

$$\hat{Y} = 2.667 + 0.326 \cdot 10 \approx 5.93$$

- (f)

$$H_0 : \beta_1 \leq 0.32, \quad H_1 : \beta_1 > 0.32, \quad R_\alpha : \hat{\beta}_1 > 0.32 + \text{se}(\hat{\beta}_1) t_\alpha(183)$$

Possiamo usare questa regione critica perché è confermata l'ipotesi gaussiana.

Calcolo p-value dei dati:

$$\begin{aligned} \hat{\beta}_1 &= 0.32 + \text{se}(\hat{\beta}_1) t_\alpha(183) \\ t_\alpha(183) &= \frac{\hat{\beta}_1 - 0.32}{\text{se}(\hat{\beta}_1)} = \frac{0.00639}{0.01516} = 0.4215 \\ z_\alpha &\simeq 0.42 \\ \alpha &= 1 - \Phi(0.42) = 1 - 0.6628 = 0.3372 \end{aligned}$$

Non posso quindi rifiutare H_0 agli usuali livelli di significatività: anche se $\hat{\beta}_1 = 0.32639 > 0.32$, non ottengo una conclusione forte a favore di $\beta_1 > 0.32$.