UNIVERSITA' DEGLI STUDI DI VERONA

LABORATORIO DI PROBABILITA' E STATISTICA

Docente: Bruno Gobbi

6 – ESERCIZI RIEPILOGATIVI PRIME 3 LEZIONI

REGRESSIONE LINEARE: SPORT - COLESTEROLO

ESERCIZIO 8: La tabella seguente riporta i risultati di uno studio su 8 persone, per le quali si sono misurati il numero dedicate allo sport settimanalmente e il livello di colesterolo.
Analizzare la relazione fra i due fenomeni utilizzando la regressione lineare, disegnando il grafico, calcolando i parametri della retta interpolante, i residui con grafico, il coefficiente di correlazione lineare e giudicandone la bontà di accostamento.

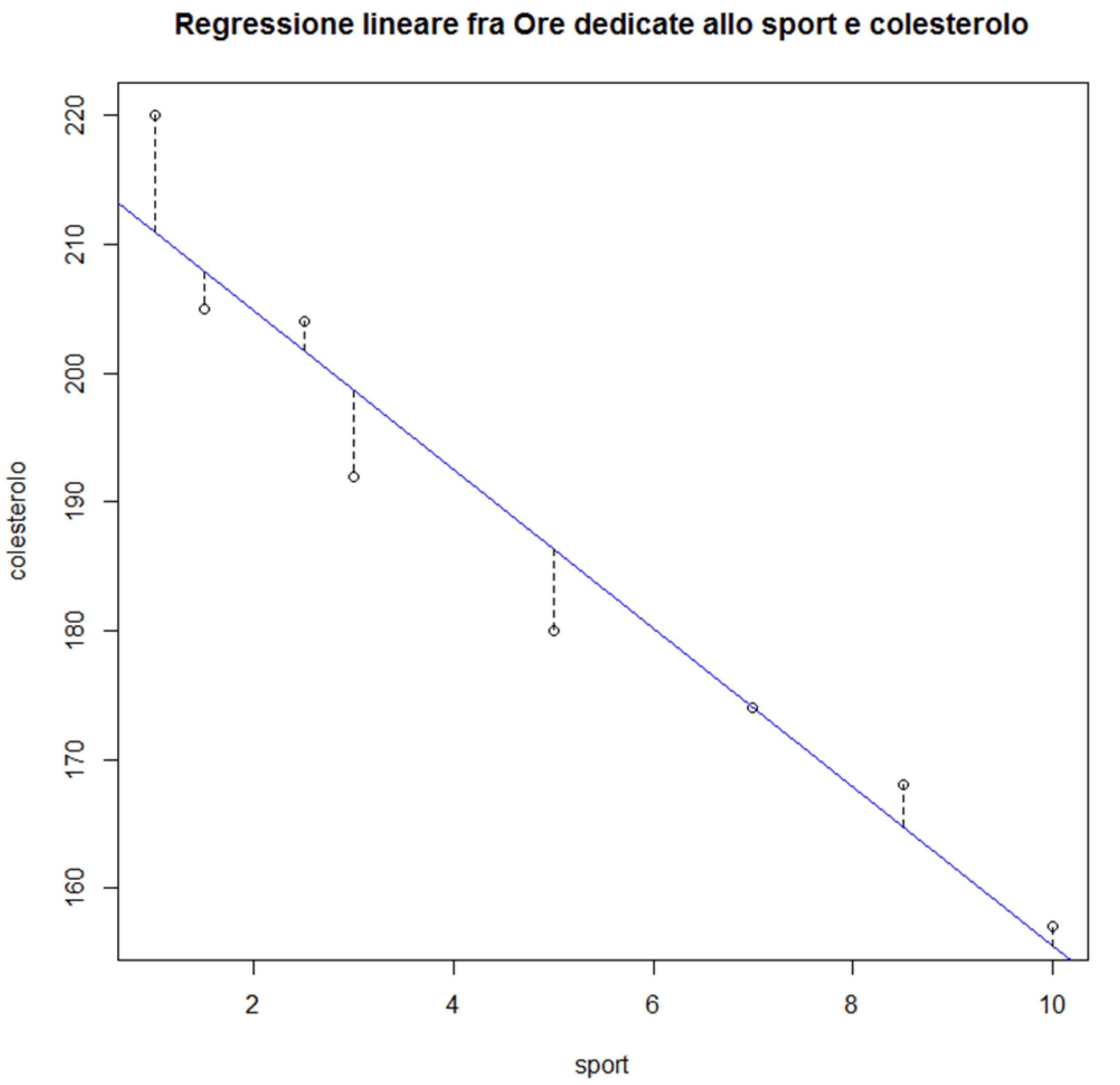
N. ore sport sett.	Livello colesterolo
1,5	205
10	157
8,5	168
7	174
1	220
3	192
5	180
2,5	204

ES. STUDIO RELAZIONE ORE DI SPORT - COLESTEROLO

```
> sport=c(1.5, 10, 8.5, 7, 1, 3, 5, 2.5)
> colesterolo=c(205, 157, 168, 174, 220, 192, 180, 204)
> plot(sport, colesterolo)
> rettasport=lm(colesterolo~sport)
> abline(rettasport, col="blue")
> segments(sport, fitted(rettasport), sport, colesterolo, lty=2)
> title(main="Regressione lineare fra Ore dedicate allo sport e
colesterolo")
```

Per scrivere la tilde ~ in Ubuntu premere:
ALT GR + ì

ES. STUDIO RELAZIONE ORE DI SPORT - COLESTEROLO



ES. STUDIO RELAZIONE ORE DI SPORT - COLESTEROLO

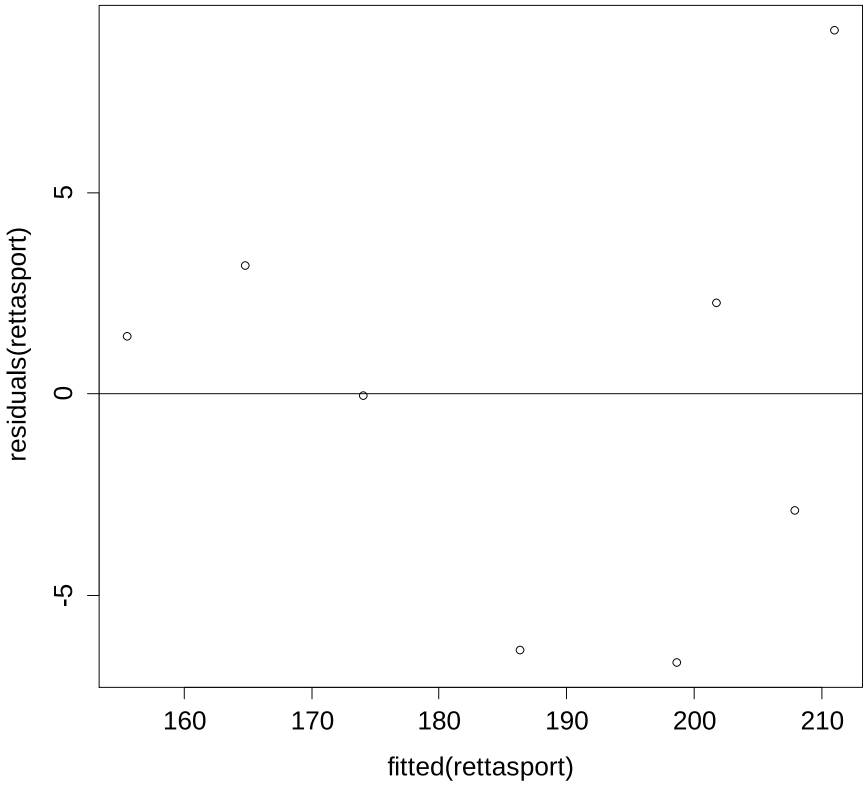
```
> summary (rettasport)

Call:
lm(formula = colesterolo ~ sport)

Residuals:
    Min      1Q  Median      3Q     Max
-6.6587 -3.7565  0.7021  2.4978  9.0283

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 217.1282     3.6498  59.490 1.52e-09 ***
sport        -6.1565     0.6344  -9.704 6.87e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.656 on 6 degrees of freedom
Multiple R-squared:  0.9401,	Adjusted R-squared:  0.9301
F-statistic: 94.16 on 1 and 6 DF,  p-value: 6.874e-05
```

ES. STUDIO RELAZIONE ORE DI SPORT - COLESTEROLO

 # I PARAMETRI TROVATI SONO a=217.1282 E b=-6.1565
QUINDI IL MODELLO TEORICO SARA':
Y'=217.1282-6.1565 *sport

 # EFFETTUIAMO L'ANALISI DEI RESIDUI

```
> plot(fitted(rettasport), residuals(rettasport))
> abline(0, 0)
```



L'analisi dei residui conferma che questi si distribuiscono in maniera uniforme e apparentemente casuale attorno all'asse zero, quindi si può confermare l'ipotesi di distribuzione casuale degli stessi, con media nulla e incorrelazione.

ES. STUDIO RELAZIONE ORE DI SPORT - COLESTEROLO

 # CALCOLIAMO IL COEFFICIENTE DI CORRELAZIONE LINEARE:

```
> R=cor(sport, colesterolo)
> R
[1] -0.9695861
```

 # POICHE' R E' MOLTO VICINO A -1 POSSIAMO AFFERMARE CHE C'E' UNA FORTE RELAZIONE LINEARE INDIRETTA FRA LE DUE VARIABILI

 # CALCOLIAMO IL COEFFICIENTE DI DETERMINAZIONE FACENDO IL QUADRATO DI R PER GIUDICARE LA BONTA' DI ACCOSTAMENTO:

```
> R2=R^2
> R2
[1] 0.9400972
```

 # DATO CHE R2 E' QUASI UGUALE A 1, DICIAMO CHE IL MODELLO TEORICO USATO SI ADATTA MOLTO BENE AI VALORI OSSERVATI

REGRESSIONE LINEARE: carotene - eritema

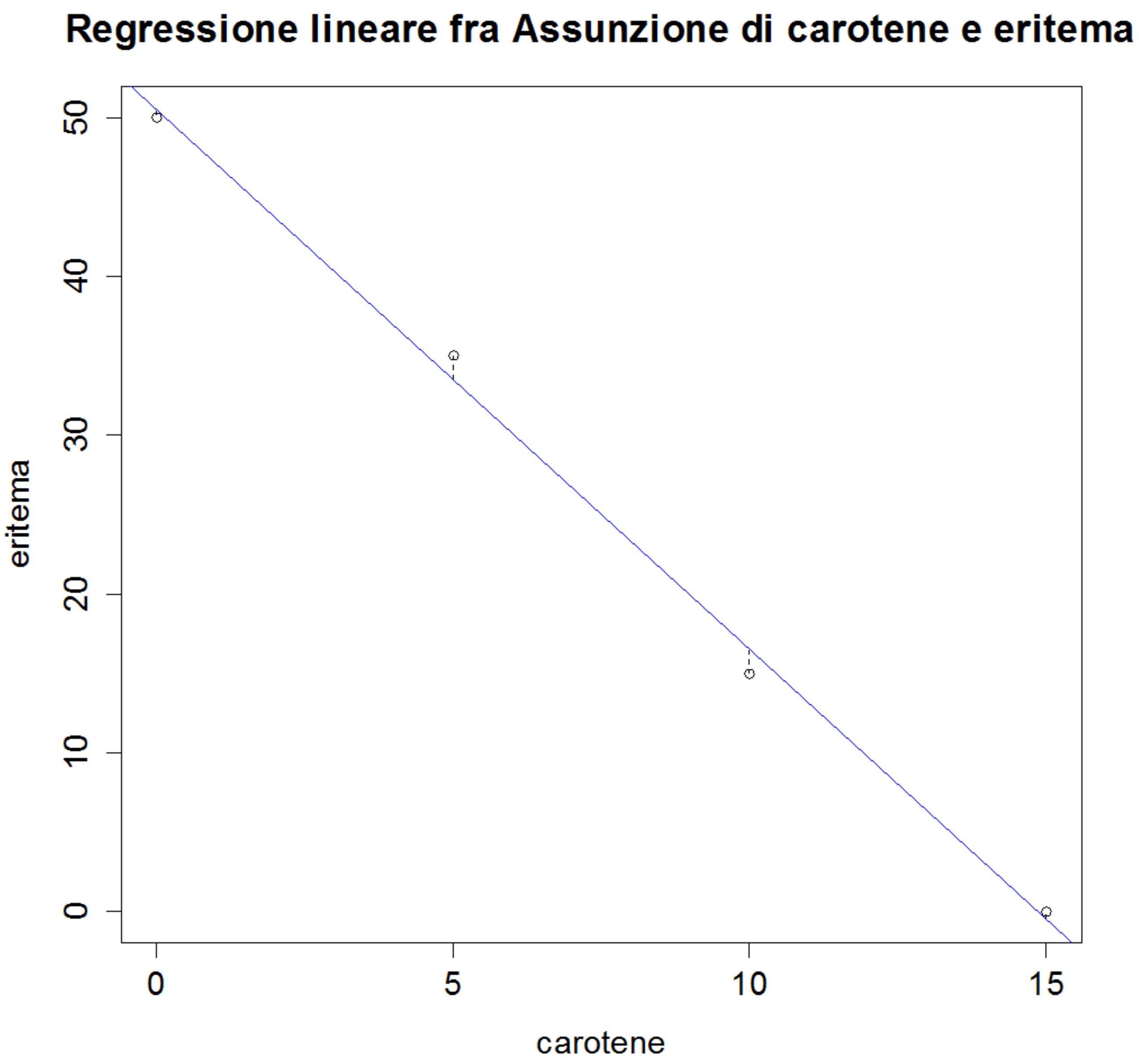
ESERCIZIO 9: Una ricerca sulla relazione fra quantità assunta di un integratore a base di beta carotene e il rischio di subire un eritema solare ha dato i risultati presenti in tabella.
Analizzare la relazione fra i due fenomeni utilizzando la regressione lineare, disegnando il grafico, calcolando i parametri della retta interpolante, i residui con grafico, il coefficiente di correlazione lineare e giudicandone la bontà di accostamento.

Quantità beta carotene	Rischio eritema
0	50
10	15
5	35
15	0

ES. STUDIO RELAZIONE carotene - eritema

```
> carotene=c(0, 10, 5, 15)
> eritema=c(50, 15, 35, 0)
> plot(carotene, eritema)
> rettascott=lm(eritema~carotene)
> abline(rettascott, col="blue")
> segments(carotene, fitted(rettascott), carotene, eritema, lty=2)
> title(main="Regressione lineare fra Assunzione di carotene e
eritema")
```

ES. STUDIO RELAZIONE carotene - eritema



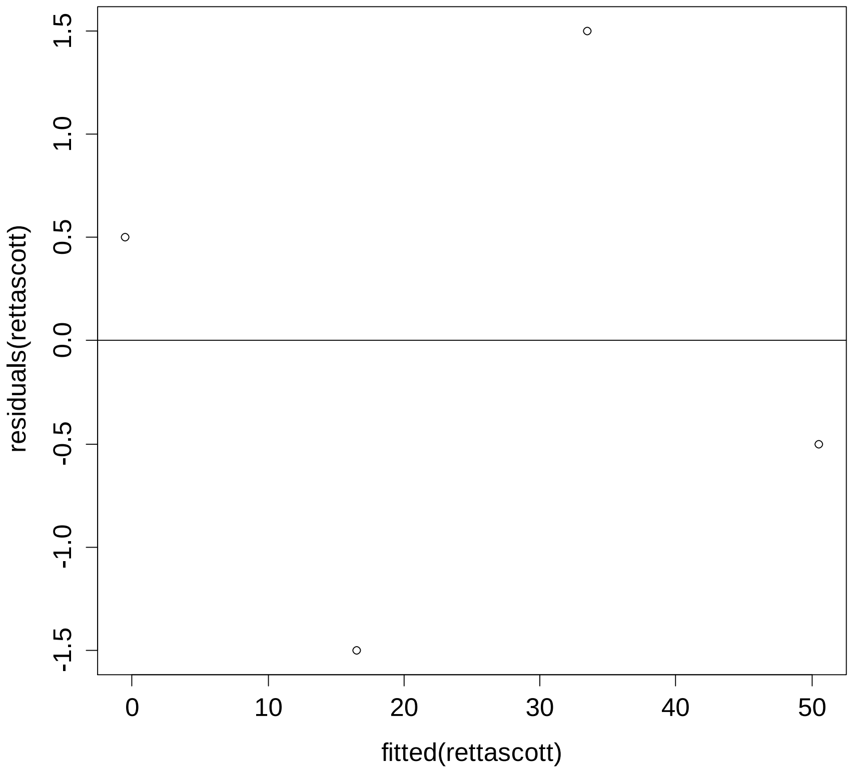
ES. STUDIO RELAZIONE carotene - eritema

```
> summary (rettascott)

Call:
lm(formula = eritema ~ carotene)

Residuals:
   1    2    3    4 
-0.5 -1.5  1.5  0.5 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  50.5000     1.3229   38.17 0.000686 ***
carotene     -3.4000     0.1414  -24.04 0.001726 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.581 on 2 degrees of freedom
Multiple R-squared:  0.9966,	Adjusted R-squared:  0.9948 
F-statistic:   578 on 1 and 2 DF,  p-value: 0.001726
```

ES. STUDIO RELAZIONE carotene - eritema

 # I PARAMETRI TROVATI SONO a=50.5 E b=-3.4
QUINDI IL MODELLO TEORICO SARA':
Y'=50.5-3.4 *carotene

 # EFFETTUIAMO L'ANALISI DEI RESIDUI
> plot(fitted(rettascott), residuals(rettascott))
> abline(0, 0)



L'analisi dei residui conferma che questi si distribuiscono in maniera uniforme e apparentemente casuale attorno all'asse zero, quindi si può confermare l'ipotesi di distribuzione casuale degli stessi, con media nulla e incorrelazione.

ES. STUDIO RELAZIONE carotene - eritema

 # CALCOLIAMO IL COEFFICIENTE DI CORRELAZIONE LINEARE:

```
> R=cor(carotene, eritema)
> R
[1] -1.1872454
```

 # POICHE' R E' NEGATIVO, POSSIAMO AFFERMARE CHE C'E' UNA FORTE RELAZIONE LINEARE INDIRETTA FRA LE DUE VARIABILI

ES. STUDIO RELAZIONE carotene - eritema

 # CALCOLIAMO IL COEFFICIENTE DI CORRELAZIONE LINEARE:

```
> R=cor(carotene, eritema)
> R
[1] -0.9982744
```

 # POICHE' R E' MOLTO VICINO A -1 POSSIAMO AFFERMARE CHE C'E' UNA FORTE RELAZIONE LINEARE INDIRETTA FRA LE DUE VARIABILI

ES. STUDIO RELAZIONE carotene - eritema

 # CALCOLIAMO IL COEFFICIENTE DI DETERMINAZIONE FACENDO IL QUADRATO DI R PER GIUDICARE LA BONTA' DI ACCOSTAMENTO:

```
> R2=R^2
> R2
[1] 0.9965517
```

 # DATO CHE R2 E' QUASI UGUALE A 1, IL MODELLO TEORICO USATO SI ADATTA MOLTO BENE AI VALORI OSSERVATI

REGRESSIONE LINEARE: vitamina c - radicali

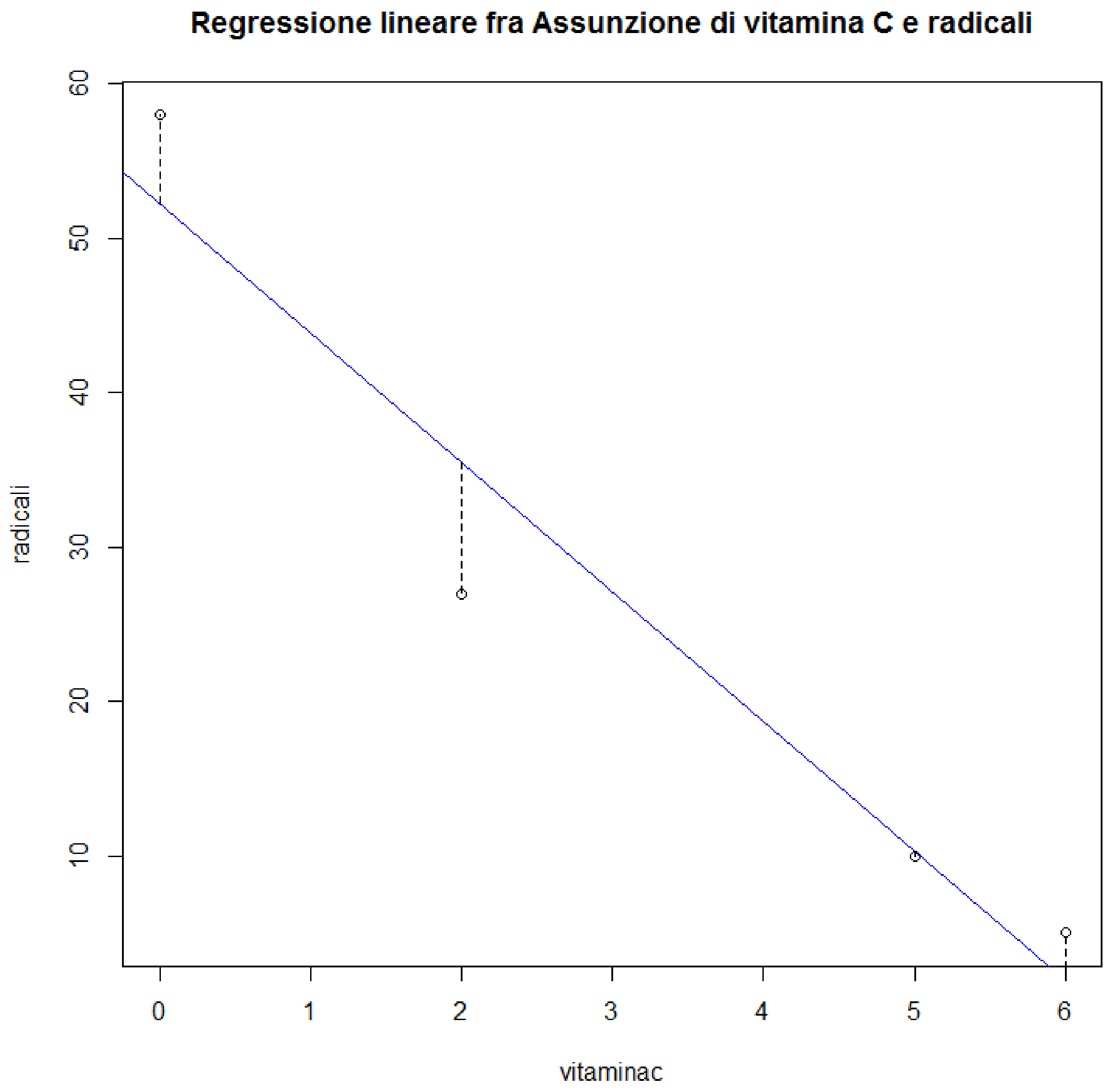
ESERCIZIO 10: Una ricerca sulla relazione fra quantità assunta di un integratore a base di vitamina C e il livello di radicali liberi sulle pareti cellulari vascolari ha dato i risultati presenti in tabella. Analizzare la relazione fra i due fenomeni utilizzando la regressione lineare, disegnando il grafico, calcolando i parametri della retta interpolante, i residui con grafico, il coefficiente di correlazione lineare e giudicandone la bontà di accostamento.

Quantità di vitamina C	Livello radicali liberi
0	58
2	27
5	10
6	5

ES. STUDIO RELAZIONE vitamina c - radicali

```
> vitaminac=c(0, 2, 5, 6)
> radicali=c(58, 27, 10, 5)
> plot(vitaminac, radicali)
> rettavit=lm(radicali~vitaminac)
> abline(rettavit, col="blue")
> segments(vitaminac, fitted(rettavit), vitaminac, radicali, lty=2)
> title(main="Regressione lineare fra Assunzione di vitamina C e
radicali")
```

ES. STUDIO RELAZIONE vitamina c - radicali



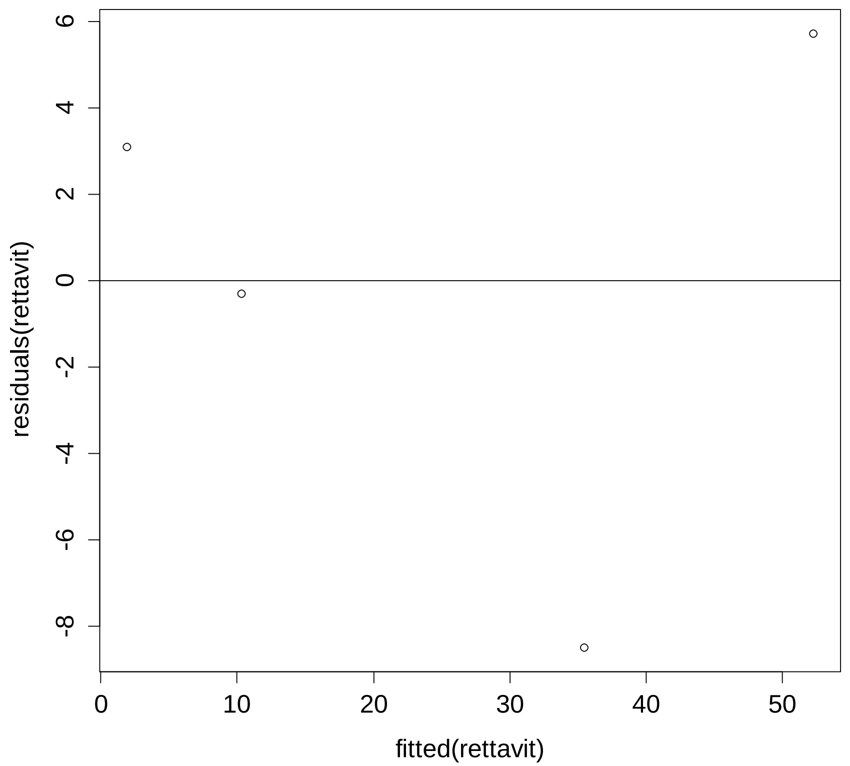
ES. STUDIO RELAZIONE vitamina c - radicali

```
> summary (rettavit)

Call:
lm(formula = radicali ~ vitaminac)

Residuals:
      1       2       3       4
 5.7143 -8.4945 -0.3077  3.0879

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   52.286      6.393   8.179   0.0146 *
vitaminac     -8.396      1.586  -5.294   0.0339 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 7.564 on 2 degrees of freedom
Multiple R-squared:  0.9334,	Adjusted R-squared:  0.9001
F-statistic: 28.02 on 1 and 2 DF,  p-value: 0.03388
```

ES. STUDIO RELAZIONE vitamina c - radicali

 # I PARAMETRI TROVATI SONO a=52.286 E b=-8.396
QUINDI IL MODELLO TEORICO SARA':
Y'=52.286-8.396 *vitaminac

 # EFFETTUIAMO L'ANALISI DEI RESIDUI
```
> plot(fitted(rettavit), residuals(rettavit))
> abline(0, 0)
```



L'analisi dei residui conferma che questi si distribuiscono in maniera uniforme e apparentemente casuale attorno all'asse zero, quindi si può confermare l'ipotesi di distribuzione casuale degli stessi, con media nulla e incorrelazione.

ES. STUDIO RELAZIONE vitamina c - radicali

 # CALCOLIAMO IL COEFFICIENTE DI CORRELAZIONE LINEARE:

```
> R=cor(vitaminac, radicali)
> R
[1] -0.96612
```

 # POICHE' R E' MOLTO VICINO A -1 POSSIAMO AFFERMARE CHE C'E' UNA FORTE RELAZIONE LINEARE INDIRETTA FRA LE DUE VARIABILI

 # CALCOLIAMO IL COEFFICIENTE DI DETERMINAZIONE FACENDO IL QUADRATO DI R PER GIUDICARE LA BONTA' DI ACCOSTAMENTO:

```
> R2=R^2
> R2
[1] 0.9333879
```

 # DATO CHE R2 E' QUASI UGUALE A 1, DICIAMO CHE IL MODELLO TEORICO USATO SI ADATTA MOLTO BENE AI VALORI OSSERVATI