UNIVERSITA' DEGLI STUDI DI VERONA

LABORATORIO DI PROBABILITA' E STATISTICA

Docente: Bruno Gobbi

Corso di laurea in Informatica e Bioinformatica

4 – ESERCIZI RIEPILOGATIVI PRIME 3 LEZIONI

1 - STATISTICA DESCRITTIVA – VENDITE PC

ESERCIZIO 1: La seguente tabella riporta i volumi di vendita (in migliaia di pezzi) dei principali produttori di computer nel 2012.

Creare una tabella in R che riporti i volumi di vendita in migliaia di pezzi e in percentuale. Alla fine creare un grafico a istogramma per i volumi di vendita in migliaia e uno a torta per le percentuali.

MARCHIO	VENDITE
Dell	9.000
HP	14.800
Lenovo	14.000
Acer	8.700
Asus	6.500
Apple Mac	4.000

1 - STATISTICA DESCRITTIVA – VENDITE PC

```
> marchio=c("Dell", "HP", "Lenovo", "Acer", "ASUS", "Apple Mac")
> vendite=c(9000, 14800, 14000, 8700, 6500, 4000)
> venditepc=data.frame(marchio, vendite)
> venditepc
    marchio vendite
1      Dell    9000
2        HP   14800
3    Lenovo   14000
4      Acer    8700
5      ASUS    6500
6 Apple Mac    4000
```

1 - STATISTICA DESCRITTIVA – VENDITE PC

 # CREIAMO LA COLONNA DELLE PERCENTUALI DI VENDITA

```
> tot_vendite=sum(vendite)
> tot_vendite
[1] 57000
> perc=vendite/tot_vendite
> perc
[1] 0.15789474 0.25964912 0.24561404 0.15263158 0.11403509
0.07017544
```

 # SE VOLESSIMO LE PERCENTUALI FORMATTATE CON IL %

```
> sprintf("%1.2f%%", 100*perc)
[1] "15.79%" "25.96%" "24.56%" "15.26%" "11.40%" "7.02%"
```

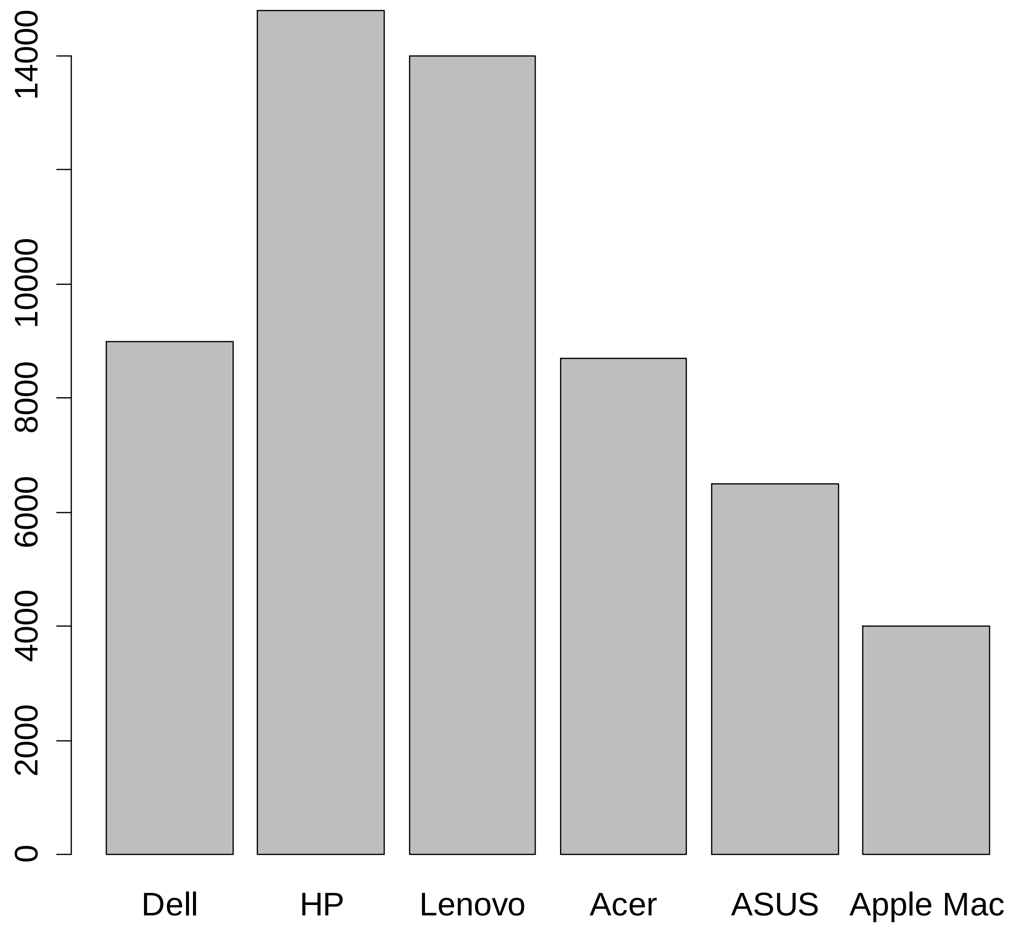
1 - STATISTICA DESCRITTIVA – VENDITE PC

```
# CREIAMO LA COLONNA DELLE PERCENTUALI DI VENDITA
> venditepc=data.frame(venditepc, perc)
> venditepc
    marchio vendite       perc
1      Dell    9000 0.15789474
2        HP   14800 0.25964912
3    Lenovo   14000 0.24561404
4      Acer    8700 0.15263158
5      ASUS    6500 0.11403509
6 Apple Mac    4000 0.07017544
```

1 - STATISTICA DESCRITTIVA – VENDITE PC

GRAFICO DEI VOLUMI DI VENDITA

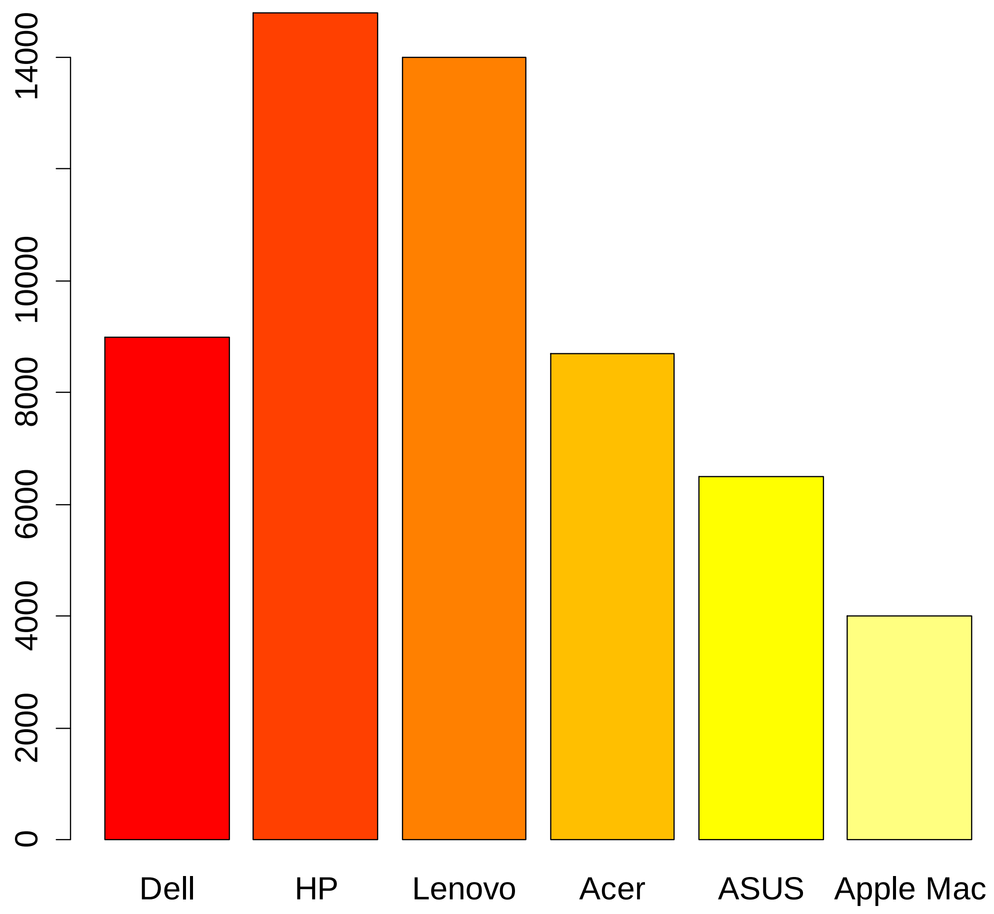
> barplot(vendite, names.arg=marchio)



1 - STATISTICA DESCRITTIVA – VENDITE PC

GRAFICO DEI VOLUMI DI VENDITA

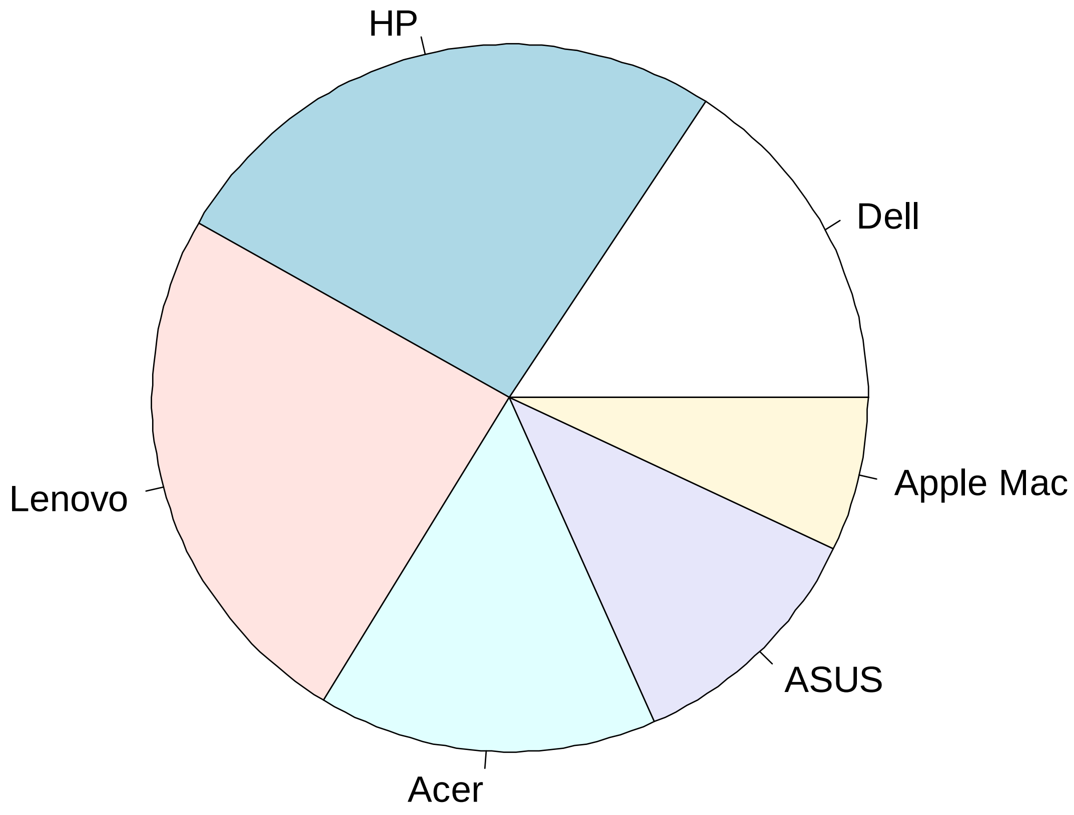
> barplot(vendite, names.arg=marchio, col=heat.colors(6))



1 - STATISTICA DESCRITTIVA – VENDITE PC

 # GRAFICO A TORTA DELLE PERCENTUALI DI VENDITA

```
> pie(perc, labels=marchio)
```



2 – CURTOSI E APPIATTIMENTO – VENDITE PC

ESERCIZIO 2: Sui dati della tabella precedente calcolare la simmetria e l'appiattimento della distribuzione delle vendite in migliaia utilizzando degli opportuni indici.

MARCHIO	VENDITE
Dell	9.000
HP	14.800
Lenovo	14.000
Acer	8.700
Asus	6.500
Apple Mac	4.000

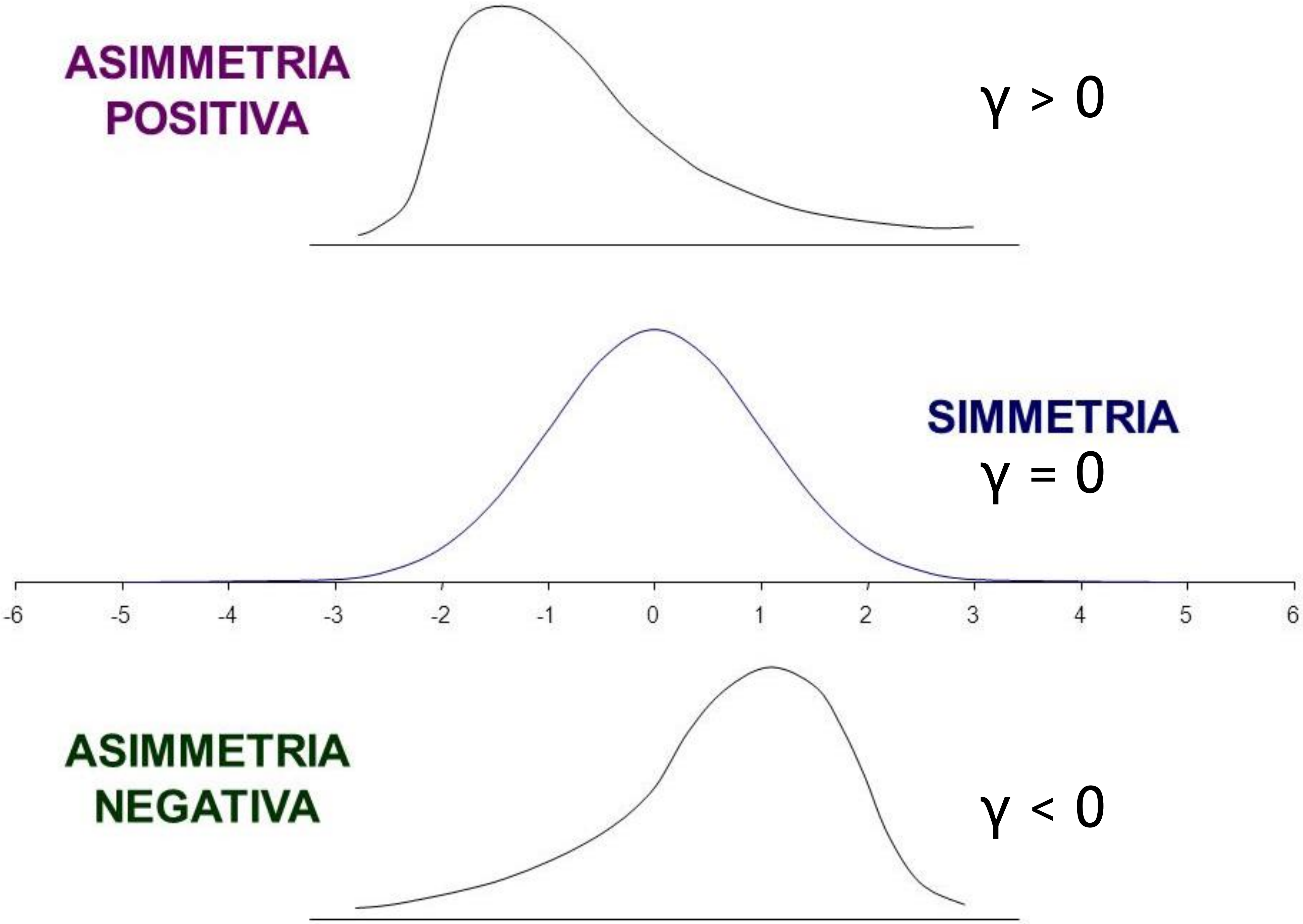
INDICE DI SIMMETRIA γ (gamma) DI FISHER

$$\gamma = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \mu}{\sigma}\right)^3$$

Se $\gamma = 0$ → allora la distribuzione è simmetrica

Se $\gamma < 0$ → allora la distribuzione è asimmetrica negativa

Se $\gamma > 0$ → allora la distribuzione è asimmetrica positiva


ASIMMETRIA POSITIVA
γ > 0
SIMMETRIA
γ = 0
ASIMMETRIA NEGATIVA
γ < 0
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6

CREAZIONE DI UNA FUNZIONE PER GAMMA

$$\gamma = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \mu}{\sigma}\right)^3$$

```
gamma =  function(x) {
m3 = mean((x-mean(x))^3)
skew = m3/(sd(x)^3)
skew
}
```

{ = AltGr + 7
} = AltGr + 0
NO tastiera numerica

2 – CURTOSI E APPIATTIMENTO – VENDITE PC

> gamma(x) = 0.1029673

C’È UN’ASIMMETRIA POSITIVA, LA DISTRIBUZIONE PRESENTA UNA CODA PIÙ LUNGA A DESTRA.

INDICE DI CURTOSI β (beta) DI PEARSON

$$\beta = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \mu}{\sigma}\right)^4$$

Se $\beta = 3$ → allora la distribuzione è MESOCURTICA

Se $\beta < 3$ → allora la distribuzione è PLATICURTICA

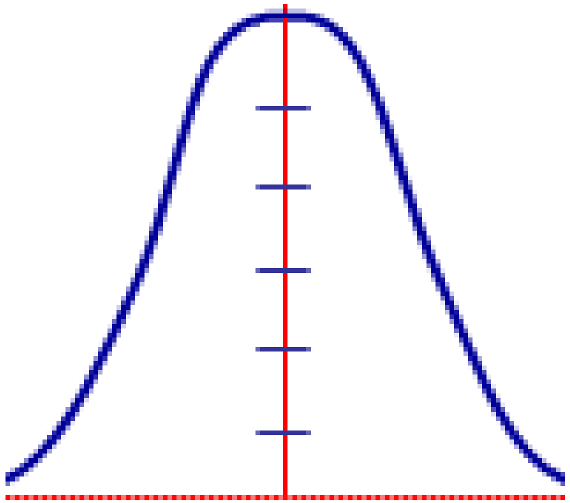
Se $\beta > 3$ → allora la distribuzione è LEPTOCURTICA

INDICE DI CURTOSI γ_2 (gamma2) DI FISHER

$$\gamma_2 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \mu}{\sigma}\right)^4 - 3$$

Se $\gamma_2 = 0$ → allora la distribuzione è MESOCURTICA
Se $\gamma_2 < 0$ → allora la distribuzione è PLATICURTICA
Se $\gamma_2 > 0$ → allora la distribuzione è LEPTOCURTICA

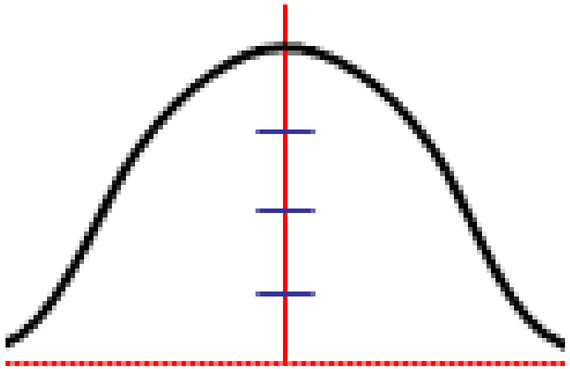
INDICI DI APPIATTIMENTO (CURTOSI)



Leptocurtica

$\beta > 3$

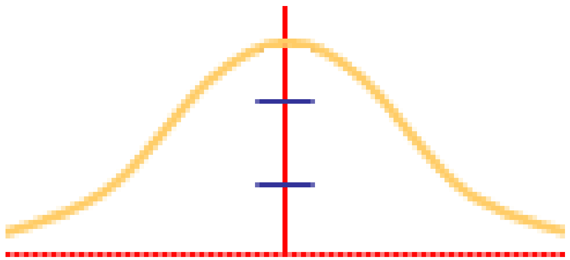
$\gamma_2 > 0$



Mesocurtica

$\beta = 3$

$\gamma_2 = 0$



Platicurtica

$\beta < 3$

$\gamma_2 < 0$

CREAZIONE DI UNA FUNZIONE PER BETA

$$\beta = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \mu}{\sigma}\right)^4$$

```
beta =  function(x) {
m4 = mean((x-mean(x))^4)
curt = m4/(sd(x)^4)
curt
}
```

2 – CURTOSI E APPIATTIMENTO – VENDITE PC

```
> beta(vendite)
[1] 1.168586
```

LA DISTRIBUZIONE APPARE SCHIACCIATA, PLATICURTICA

```
> beta(vendite)-3
[1] -1.831414
```

3 – STATISTICHE E BOXPLOT - LAGO HURON

ESERCIZIO 3: Utilizzando la base dati già presente in R relativamente ai livelli del Lago Huron fra il 1875 e il 1972 (nome del database: ‘’LakeHuron’’), calcolare:

- Media
- Mediana
- Primo e terzo quartile
- Minimo e Massimo
- Varianza campionaria
- Numero di elementi del database

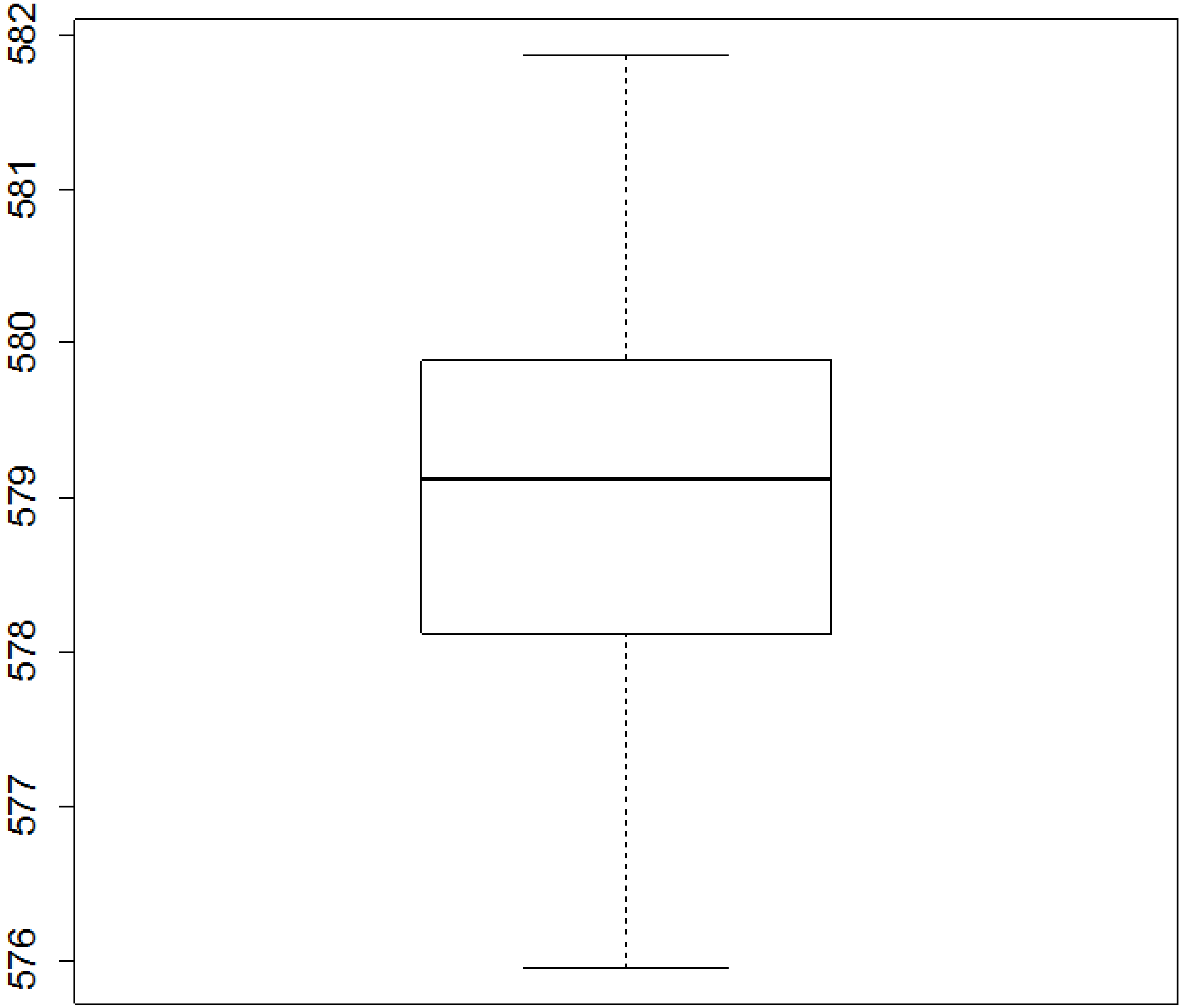
Infine disegnare il grafico boxplot della serie storica.

3 – STATISTICHE E BOXPLOT - LAGO HURON

```
> summary(LakeHuron)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  576.0   578.1   579.1   579.0   579.9   581.9 

> var(LakeHuron)
[1] 1.737911

> length(LakeHuron)
[1] 98
```

3 – STATISTICHE E BOXPLOT - LAGO HURON

> boxplot(LakeHuron)



ES. PRECIPITAZIONI INVERNALI E TEMPERATURE ESTIVE

ESERCIZIO 4: La tabella riporta la distribuzione delle precipitazioni medie nei mesi invernali dal 1950 in 10 città italiane e le temperature medie nelle estati seguenti. Giudicare se esiste una connessione fra la quantità di pioggia caduta d'inverno e le temperature delle estati seguenti.

	TEMPERATURE MEDIE ESTIVE		
PRECIPITAZIONI INVERNALI (IN MM)	Da 26 a 27	Da 27 a 28	Oltre 28
Da 40 a 50	50	53	49
Da 50 a 60	35	65	60
Da 60 a 70	40	56	50
Oltre 70	32	60	50

	alpha (significatività)	
g.d.l.	1%	5%
1	6,64	3,84
2	9,21	5,99
3	11,35	7,82
4	13,28	9,49
5	15,09	11,07
6	16,81	12,59
7	18,48	14,07
8	20,09	15,51
9	21,67	16,92
10	23,21	18,31

ES. PRECIPITAZIONI INVERNALI E TEMPERATURE ESTIVE

```
> meteo=matrix(c(50, 53, 49, 35, 65, 60, 40, 56, 50, 32, 60,
50), nrow=4, byrow=TRUE)
> pioggia=c("Da 40 a 50", "Da 50 a 60", "Da 60 a 70", "Oltre 70")
> temp=c("Da 26 a 27", "Da 27 a 28", "Oltre 28")
> dimnames(meteo)=list(pioggia, temp)
> meteo
           Da 26 a 27 Da 27 a 28 Oltre 28
Da 40 a 50         50         53       49
Da 50 a 60         35         65       60
Da 60 a 70         40         56       50
Oltre 70           32         60       50
> mosaicplot(meteo)
```

ES. PRECIPITAZIONI INVERNALI E TEMPERATURE ESTIVE

```
> testchiq=chisq.test(meteo)
> testchiq
        Pearson's Chi-squared test
 data:  meteo
X-squared = 6.3715, df = 6, p-value = 0.3829
```

 # I GRADI DI LIBERTA' SONO 6 PERCHE' DATI DA (r-1)*(c*1)=(4-1)*(3-1)
POICHE' IL VALORE CALCOLATO DEL CHI-QUADRATO E' 6.3715, INFERIORE ALLA SOGLIA CRITICA DI 16,81 VALIDO ALL'1% PER 6 G.D.L., SI ACCETTA L'IPOTESI NULLA DI INDIPENDENZA A LIVELLO DELL'1%. LA STESSA COSA VALE PER LA SOGLIA PER IL LIVELLO DI SIGNIFICATIVITA' DEL 5% E 6 G.D.L., IN QUANTO IL CHI-QUADRATO CALCOLATO E' SUPERIORE A 12,59
PROVIAMO COMUNQUE A CALCOLARE IL V DI CRAMER

	alpha (significatività)	
g.d.l.	**1%**	**5%**
1	6,64	3,84
2	9,21	5,99
3	11,35	7,82
4	13,28	9,49
5	15,09	11,07
6	16,81	12,59
7	18,48	14,07
8	20,09	15,51
9	21,67	16,92
10	23,21	18,31

ES. PRECIPITAZIONI INVERNALI E TEMPERATURE ESTIVE

 # CALCOLIAMO IL VALORE DELLA STATISTICA V DI CRAMER

```
> chiquadrato= testchiq$statistic
> chiquadrato
X-squared
6.371519
```

 # IL TOTALE DI ELEMENTI PRESENTI SI OTTIENE IN QUESTO MODO:

```
> N = sum(meteo)
> N
[1] 600
```

 # SI SCEGLIE IL MINORE FRA IL NUMERO DI RIGHE E DI COLONNE E SI SOTTRAE 1

```
> V=sqrt( chiquadrato / (N*(3-1)) )
> V
X-squared
0.07286699
```

 # IL RISULTATO PORTA AD AFFERMARE CHE C'È UNA BASSISSIMA CONNESSIONE FRA I DUE FENOMENI. IN ALTRE PAROLE NON SEMBRA ESSERCI UN LEGAME FRA LA QUANTITA' DI PIOGGIA CHE CADE IN INVERNO E LE TEMPERATURE MEDIE DELLE ESTATI SUCCESSIVE.

ES. LAUREATI PER FACOLTA' E LAVORO

ESERCIZIO 5: La tabella presenta la distribuzione di alcuni laureati in diverse facoltà e la loro condizione occupazionale dopo un anno dalla laurea.

	FACOLTA'		
CONDIZIONE OCCUPAZIONALE	**Informatica**	**Lettere**	**Giurisprudenza**
Occupati	70	40	30
In cerca di occupazione	20	15	30

	alpha (significatività)	
g.d.l.	**1%**	**5%**
1	6,64	3,84
2	9,21	5,99
3	11,35	7,82
4	13,28	9,49
5	15,09	11,07
6	16,81	12,59
7	18,48	14,07
8	20,09	15,51
9	21,67	16,92
10	23,21	18,31

ES. LAUREATI PER FACOLTA’ E LAVORO

```
> laureati=matrix(c(70, 40, 30, 20, 15, 30), nrow=2, byrow=TRUE)
> condizione=c("Occupati", "In cerca di occupazione")
> facoltà=c("Informatica", "Lettere", "Giurisprudenza")
> dimnames(laureati)=list(condizione, facoltà)
> laureati
                        Informatica Lettere Giurisprudenza
Occupati                         70      40             30
In cerca di occupazione          20      15             30

> mosaicplot(laureati)
```

ES. LAUREATI PER FACOLTA' E LAVORO

```
> testchiq=chisq.test(laureati)
> testchiq

        Pearson's Chi-squared test

data:  laureati
X-squared = 13.5108, df = 2, p-value = 0.001165
```

 # POICHE' IL VALORE CALCOLATO DEL CHI-QUADRATO E' 13.5108, SUPERIORE ALLA SOGLIA CRITICA DI 9,21 VALIDO ALL'1% PER 2 G.D.L., SI RIFIUTA L'IPOTESI NULLA DI INDIPENDENZA E SI CONFERMA LA CONNESSIONE FRA I FENOMENI

I GRADI DI LIBERTA' SONO 2 PERCHE' DATI DA (r-1)*(c*1)=(2-1)*(3-1)

	alpha (significatività)	
g.d.l.	**1%**	**5%**
1	6,64	3,84
2	9,21	5,99
3	11,35	7,82
4	13,28	9,49
5	15,09	11,07
6	16,81	12,59
7	18,48	14,07
8	20,09	15,51
9	21,67	16,92
10	23,21	18,31

ES. LAUREATI PER FACOLTA' E LAVORO

 # CALCOLIAMO IL VALORE DELLA STATISTICA V DI CRAMER

```
> chiquadrato= testchiq$statistic
> chiquadrato
X-squared
13.51079
```

 # IL TOTALE DI ELEMENTI PRESENTI SI OTTIENE IN QUESTO MODO:

```
> N = sum(laureati)
> N
[1] 205
```

 # SI SCEGLIE IL MINORE FRA IL NUMERO DI RIGHE E DI COLONNE E SI SOTTRAE 1

```
> V=sqrt( chiquadrato / (N*(2-1)) )
> V
X-squared
0.2567223
```

 # IL RISULTATO PORTA AD AFFERMARE CHE C'È UNA DISCRETA CONNESSIONE FRA I DUE FENOMENI

REGRESSIONE LINEARE: carotene - eritema

ESERCIZIO 8: 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.

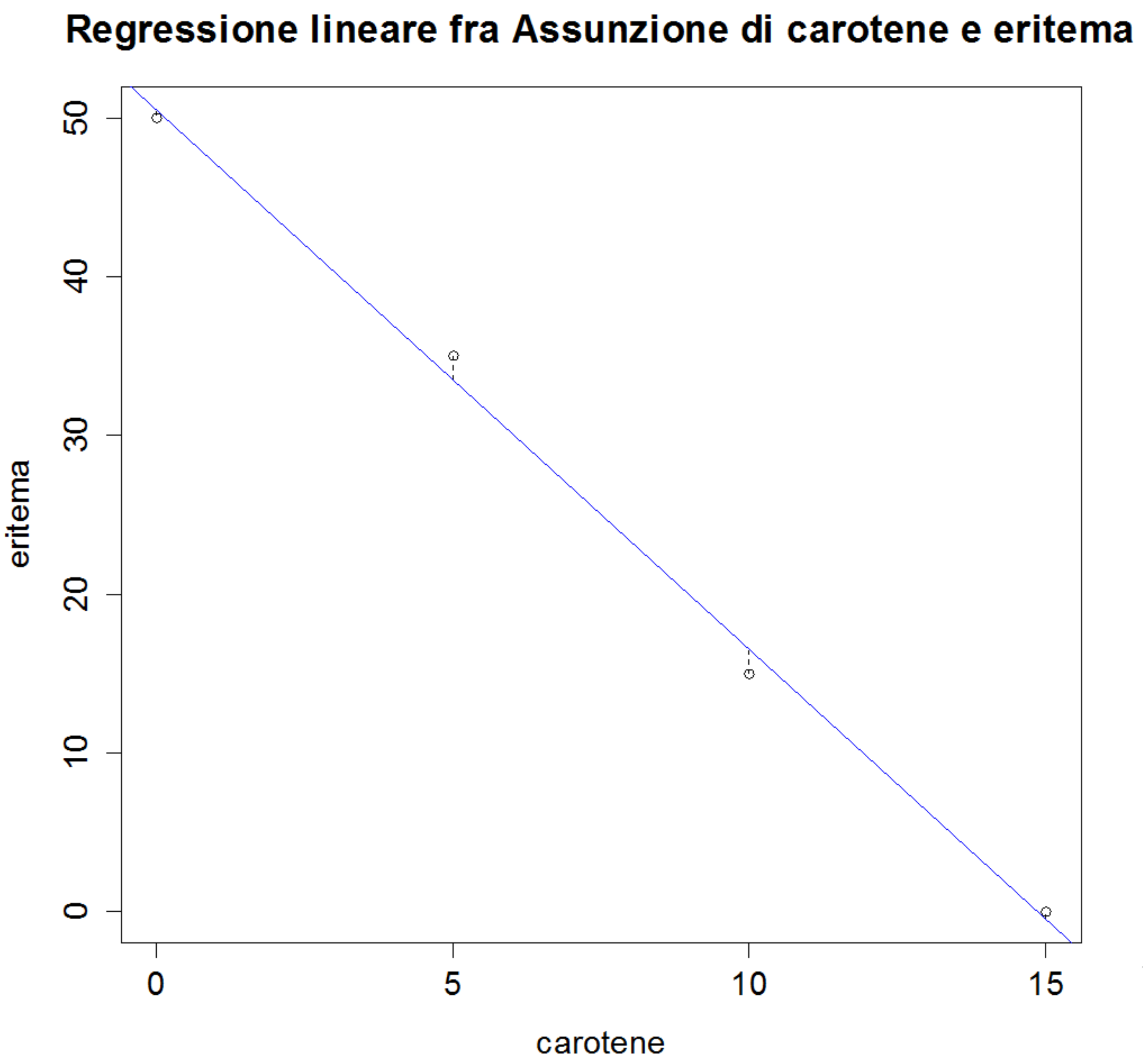
Quantità beta carotene	Rischio eritema
0	50
10	15
5	35
15	0

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")
```

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

ES. STUDIO RELAZIONE carotene - eritema



ES. STUDIO RELAZIONE carotene - eritema

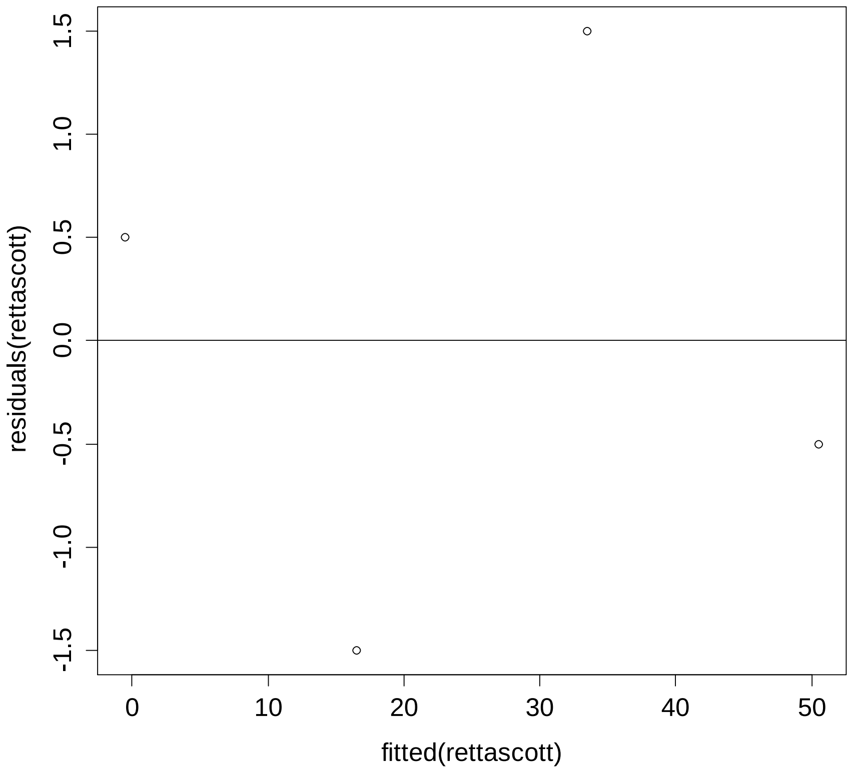
```
> summary (rettascott)

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] -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: SPORT - COLESTEROLO

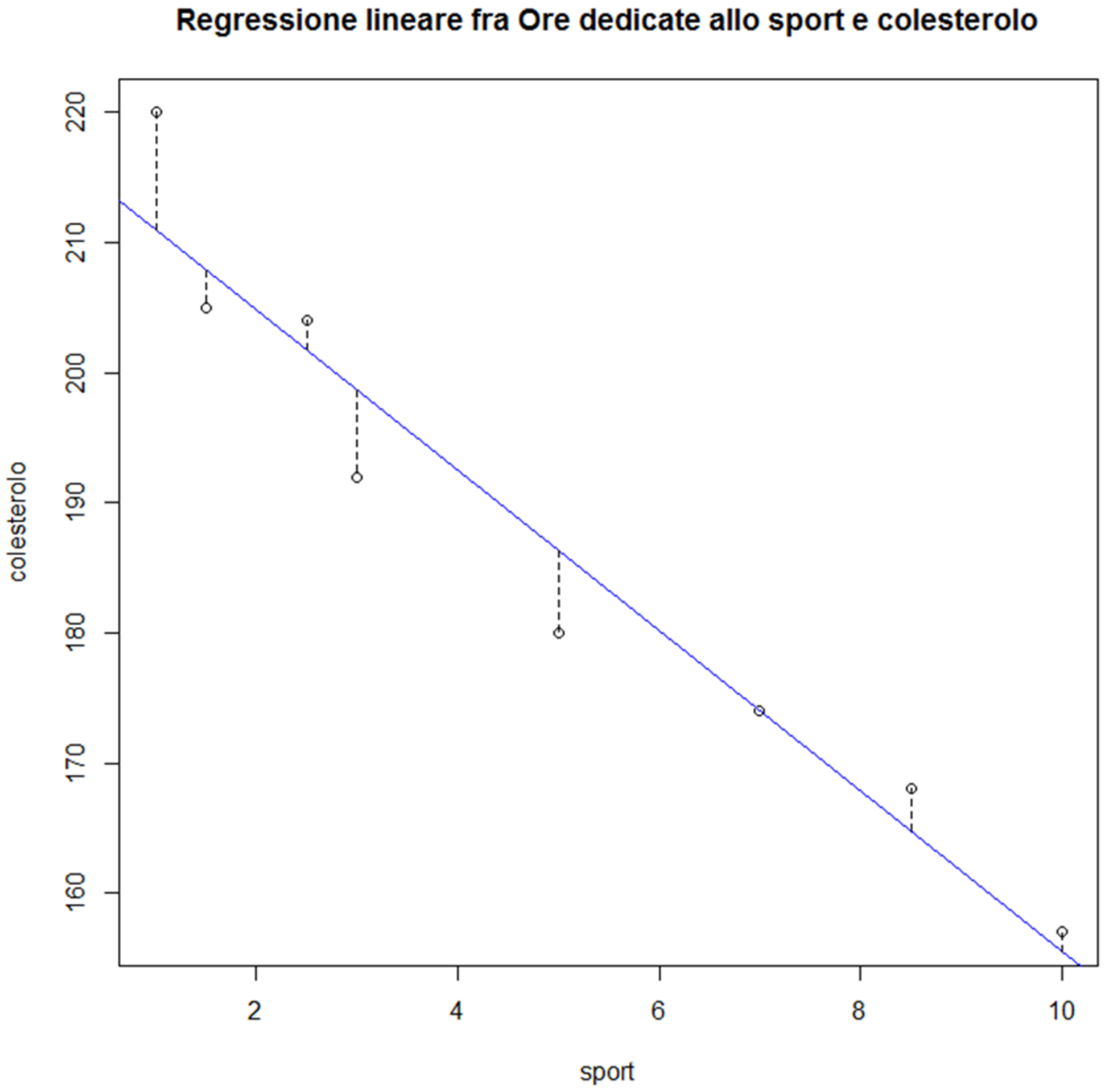
ESERCIZIO 9: La tabella seguente riporta i risultati di uno studio su 8 persone, per le quali si sono misurati il numero dedicate allo sport settimanalmente e il livello di colesterolo.
Analizzare la relazione fra i due fenomeni utilizzando la regressione lineare, disegnando il grafico, calcolando i 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")
```

ES. STUDIO RELAZIONE ORE DI SPORT - COLESTEROLO



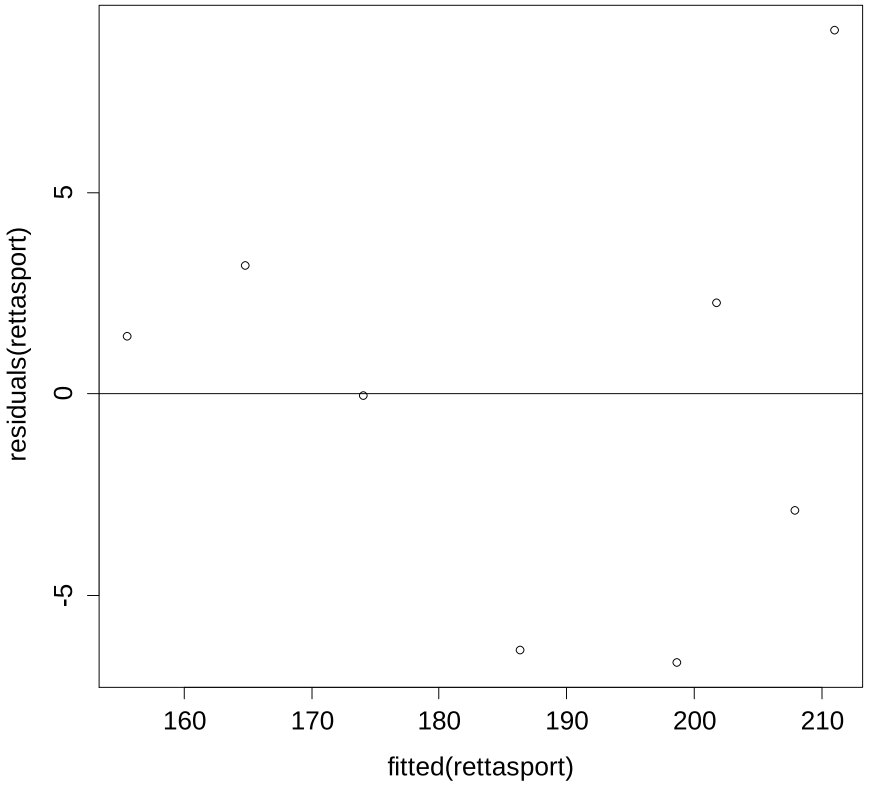
ES. STUDIO RELAZIONE ORE DI SPORT - COLESTEROLO

```
> 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: vitamina c - radicali

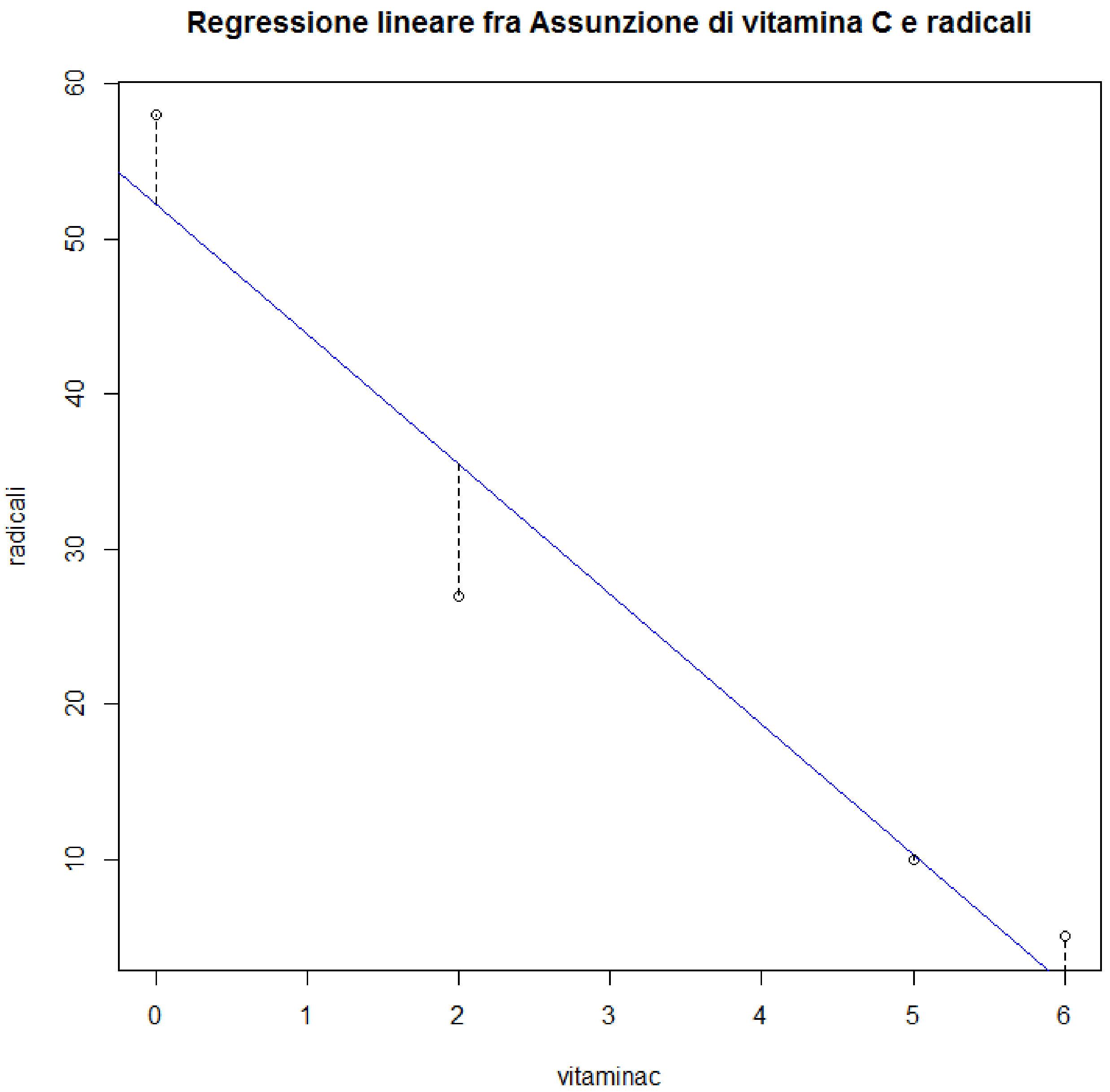
ESERCIZIO 10: Una ricerca sulla relazione fra quantità assunta di un integratore a base di vitamina C e il livello di radicali liberi sulle pareti cellulari vascolari ha dato i risultati presenti in tabella. Analizzare la relazione fra i due fenomeni utilizzando la regressione lineare, disegnando il 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



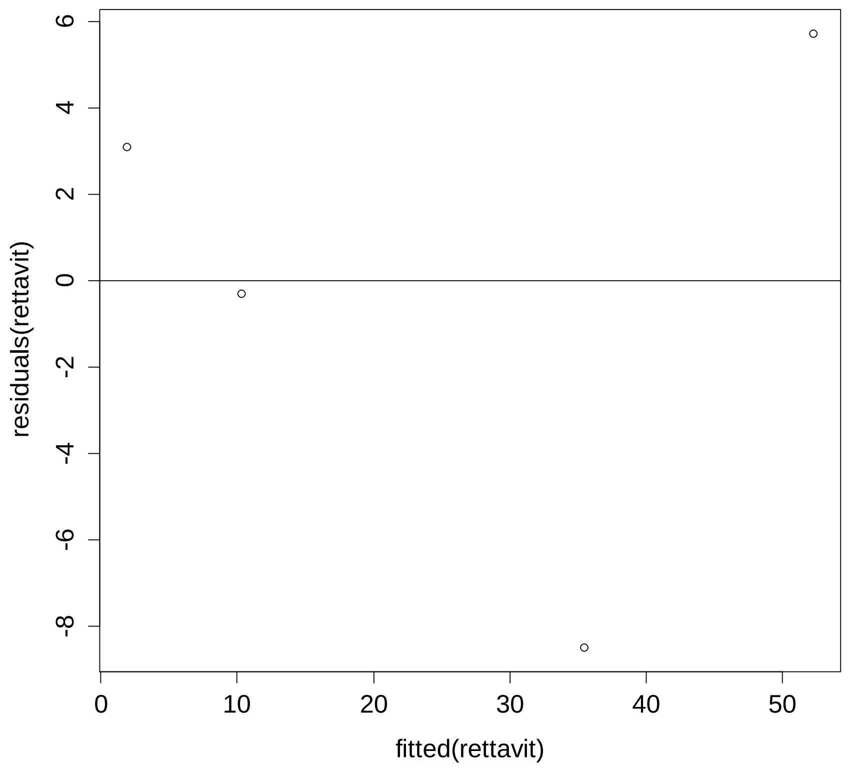
ES. STUDIO RELAZIONE vitamina c - radicali

```
> 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
```