



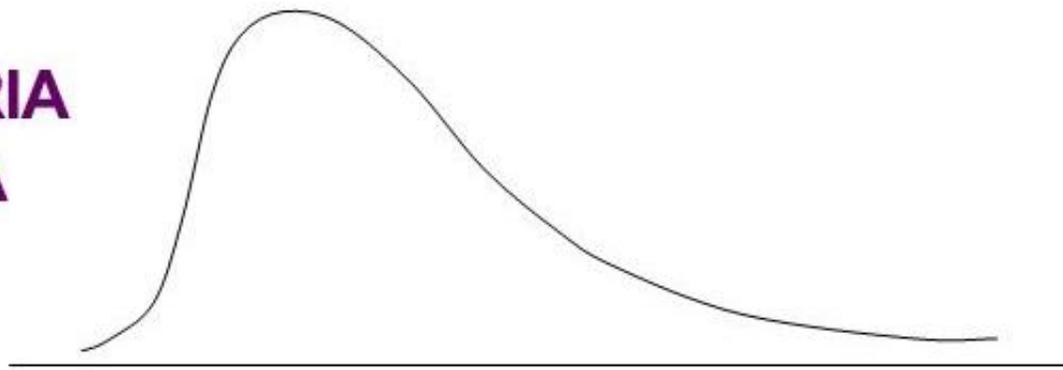
UNIVERSITA' DEGLI STUDI DI VERONA

# LABORATORIO DI PROBABILITA' E STATISTICA

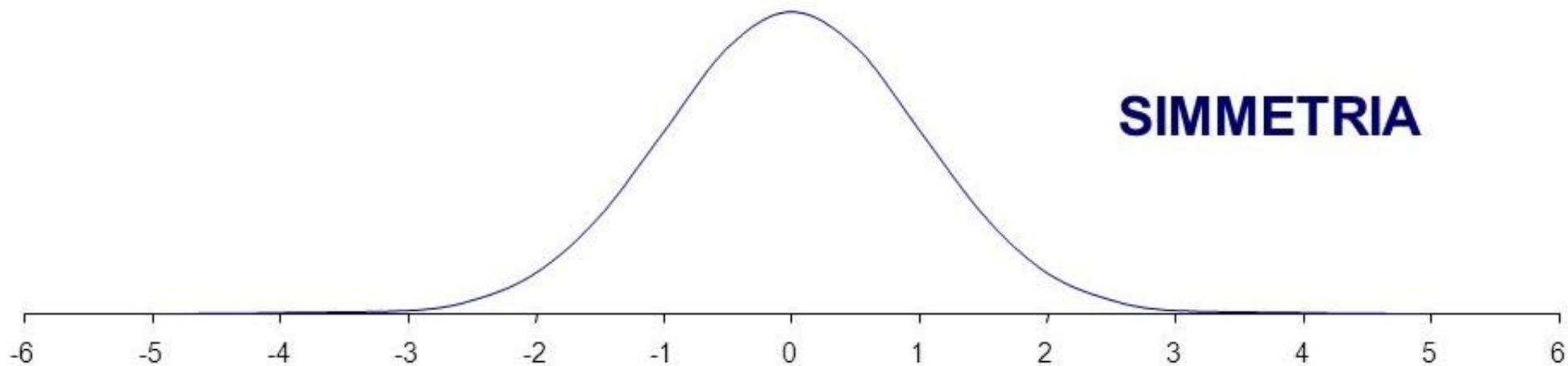
Docente: Bruno Gobbi

## 2 - INDICI DI SIMMETRIA E APPIATTIMENTO

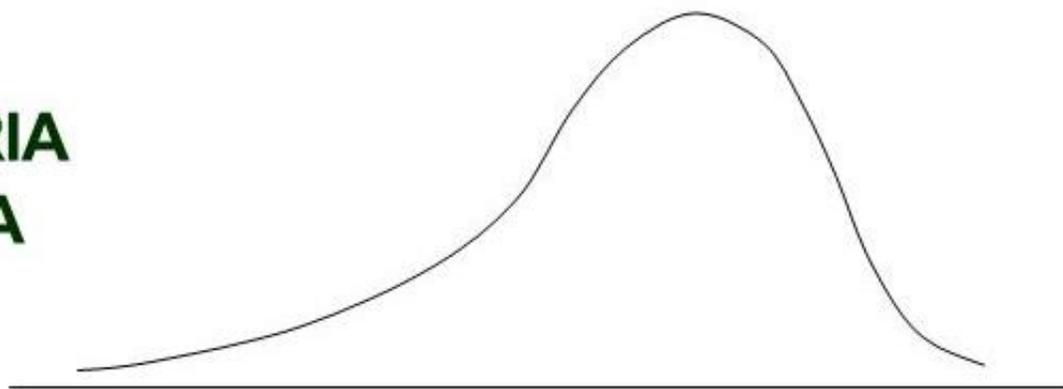
**ASIMMETRIA  
POSITIVA**



**SIMMETRIA**



**ASIMMETRIA  
NEGATIVA**



## INDICE DI SIMMETRIA $\gamma$ (gamma) DI FISHER

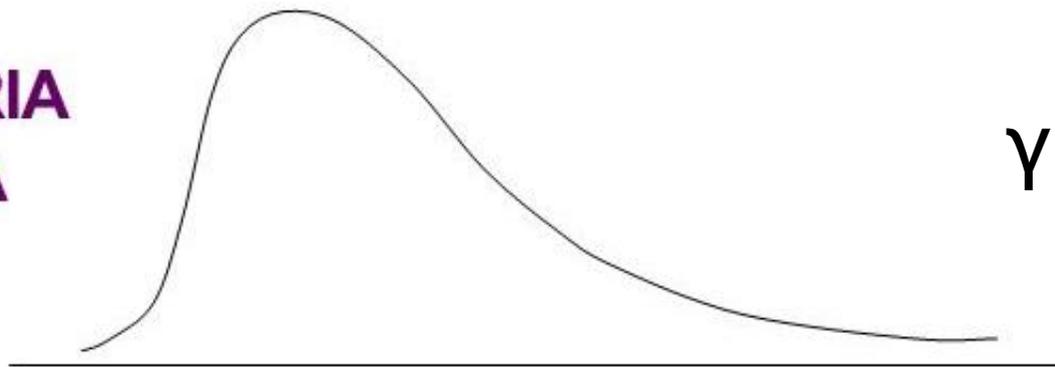
$$\gamma = \frac{1}{N} \sum_{i=1}^N \left( \frac{x_i - \mu}{\sigma} \right)^3$$

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

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

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

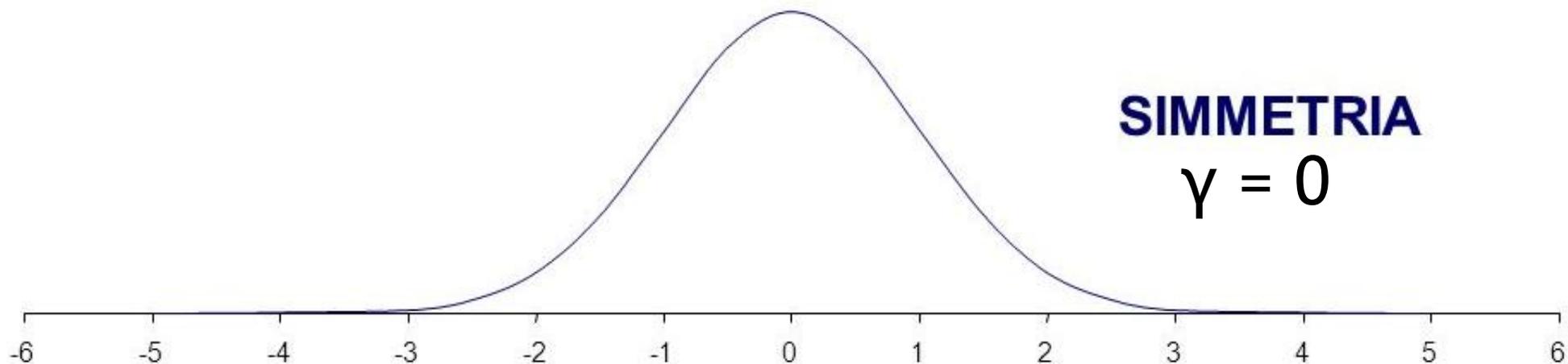
**ASIMMETRIA  
POSITIVA**



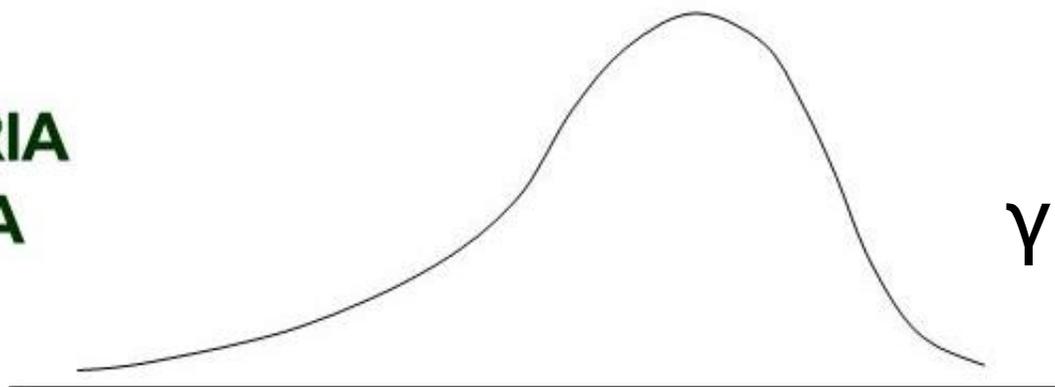
$$\gamma > 0$$

**SIMMETRIA**

$$\gamma = 0$$



**ASIMMETRIA  
NEGATIVA**



$$\gamma < 0$$

## SIMMETRIA (O SKEWNESS) IN R

In R esistono diversi pacchetti aggiuntivi che aiutano a calcolare la simmetria di una distribuzione.

ES.

- moment
- e1071
- fUtilities

# 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

# SIMMETRIA (O SKEWNESS) IN R

ES.

$x = c(0, 1, 1, 2, 2, 3, 4, 5)$

Valutare la simmetria di tale distribuzione.

# SIMMETRIA (O SKEWNESS) IN R

ES.

$x = c(0, 1, 1, 2, 2, 3, 4, 5)$

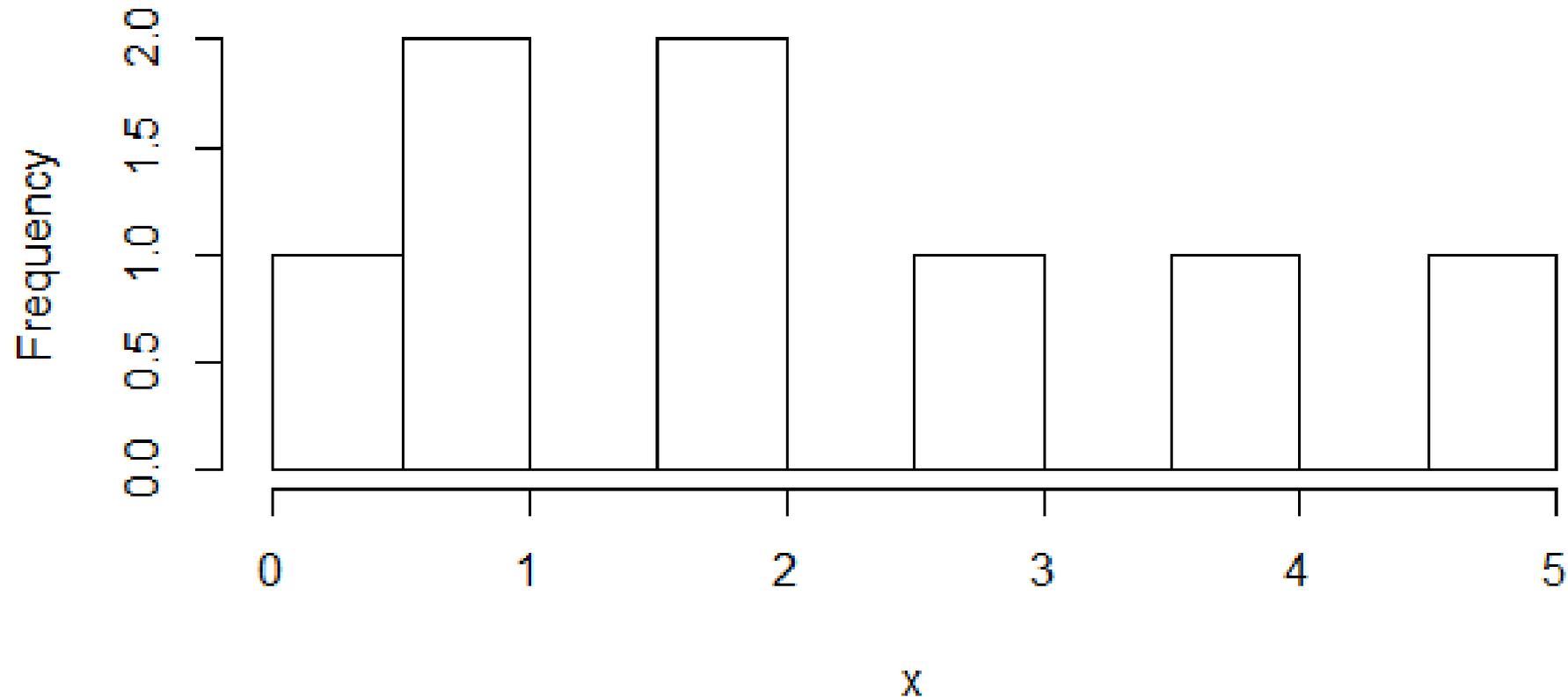
$\text{gamma}(x) = 0.3024528$

C'è asimmetria positiva, la distribuzione presenta una coda più lunga a destra.

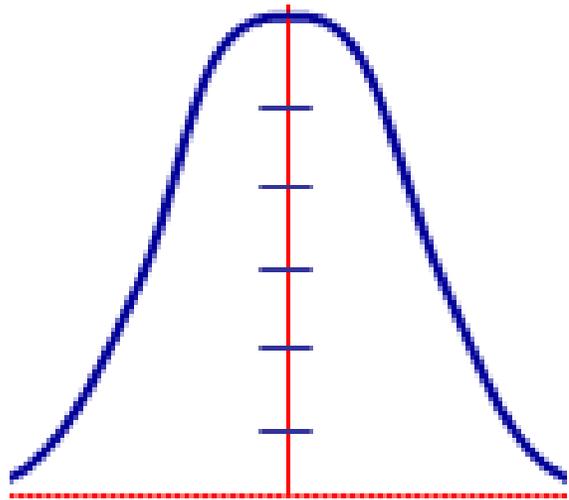
# VERIFICA GRAFICO SIMMETRIA

```
hist(x, freq=TRUE, breaks=length(x))
```

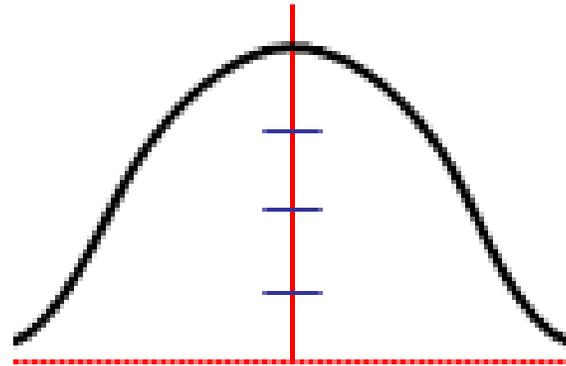
**Histogram of x**



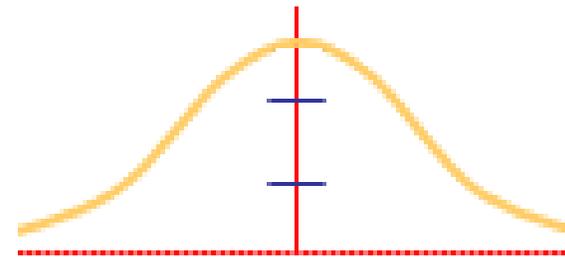
# INDICI DI APPIATTIMENTO (CURTOSI)



Leptocurtica



Mesocurtica



Platicurtica

## INDICE DI CURTOSI $\beta$ (beta) DI PEARSON

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

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

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

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

## INDICE DI CURTOSI $\gamma_2$ (gamma2) DI FISHER

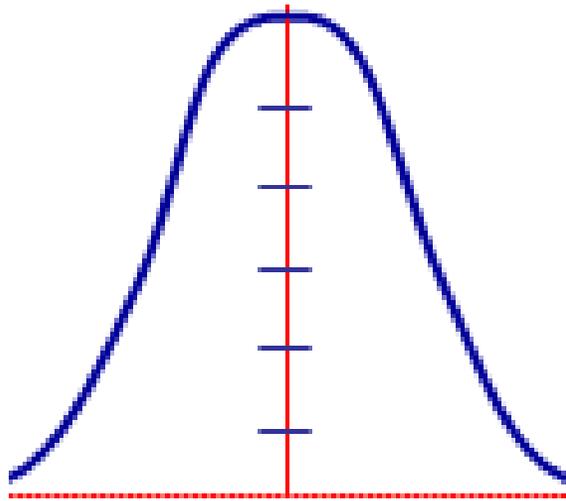
$$\gamma_2 = \frac{1}{N} \sum_{i=1}^N \left( \frac{x_i - \mu}{\sigma} \right)^4 - 3$$

Se  $\gamma_2 = 0 \rightarrow$  allora la distribuzione è MESOCURTICA

Se  $\gamma_2 < 0 \rightarrow$  allora la distribuzione è PLATICURTICA

Se  $\gamma_2 > 0 \rightarrow$  allora la distribuzione è LEPTOCURTICA

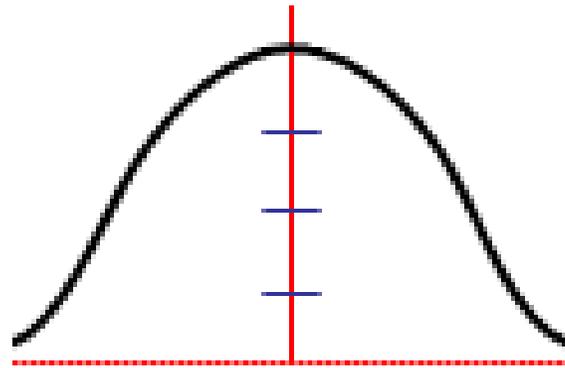
# INDICI DI APPIATTIMENTO (CURTOSI)



Leptocurtica

$$\beta > 3$$

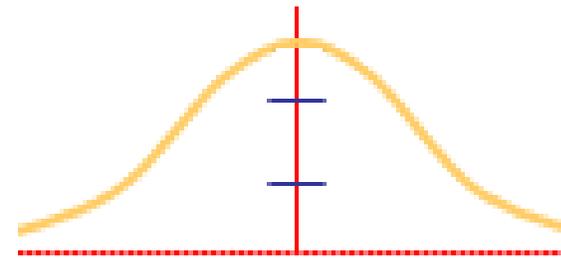
$$\gamma_2 > 0$$



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  
}
```

# APPIATTIMENTO (O CURTOSI) IN R

ES.

$x = c(0, 1, 1, 2, 2, 3, 4, 5)$

Misurare la curtosi di  $x$ .

# APPIATTIMENTO (O CURTOSI) IN R

ES.

$x = c(0, 1, 1, 2, 2, 3, 4, 5)$

$\text{beta}(x) = 1.569003$

La distribuzione presenta un andamento  
“schiacciato” ovvero platicurtico.

# CREAZIONE DI UNA FUNZIONE PER GAMMA2

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

```
gamma2 = function(x) {  
  m4 = mean((x-mean(x))^4)  
  curt = m4 / (sd(x)^4)  
  curt - 3  
}
```

# APPIATTIMENTO (O CURTOSI) IN R

ES.

$x = c(0, 1, 1, 2, 2, 3, 4, 5)$

Misurare l'appiattimento di  $x$  con  $\gamma_2$

# APPIATTIMENTO (O CURTOSI) IN R

ES.

$x = c(0, 1, 1, 2, 2, 3, 4, 5)$

$\text{gamma2}(x) = -1.430997$

Essendo negativo,  $\gamma_2$  conferma la forma platicurtica della distribuzione.

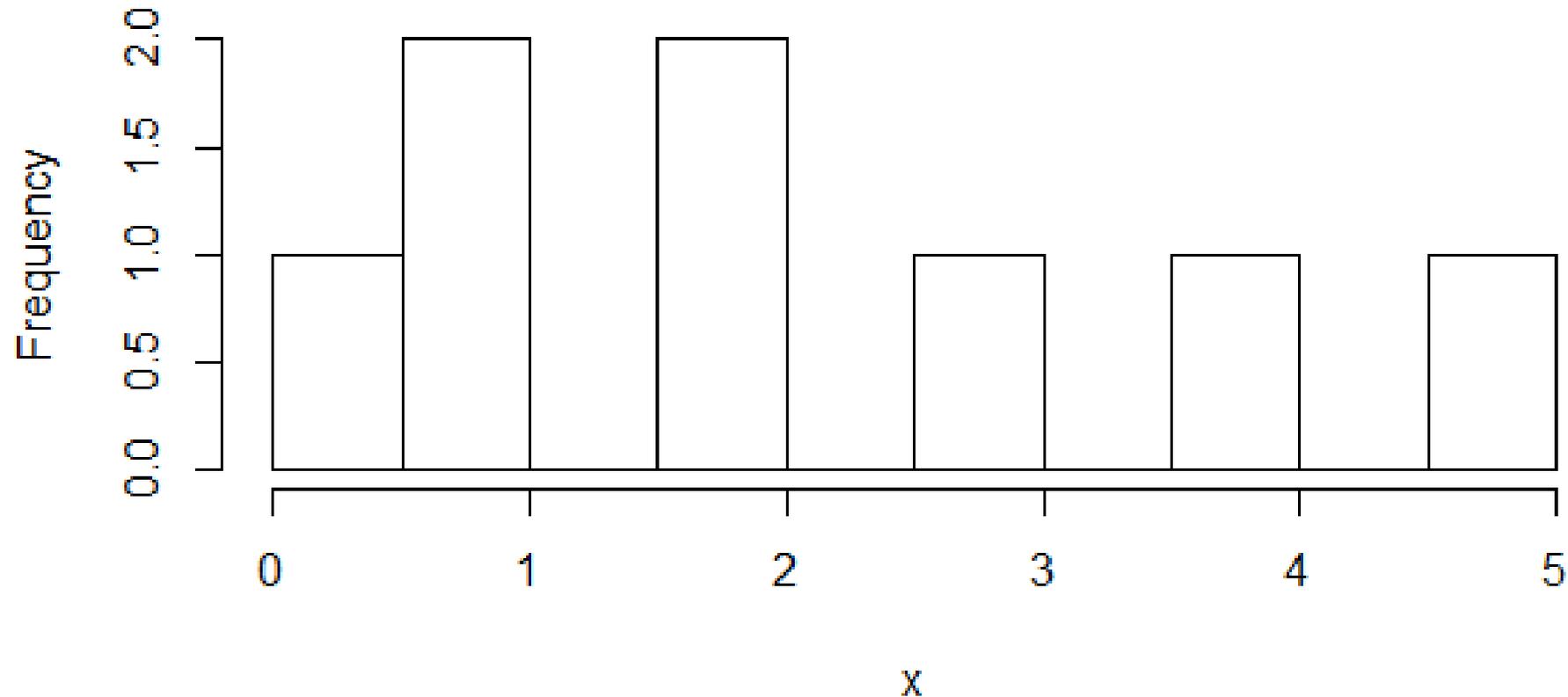
Altro metodo per calcolare gamma2:

$> \text{beta}(x) - 3$

# VERIFICA GRAFICO APPIATTIMENTO

```
hist(x, freq=TRUE, breaks=length(x))
```

**Histogram of x**



## ES. CON DATI PONDERATI

<b>Num. esami</b>	<b><math>n_i</math></b>
<b>0</b>	<b>14</b>
<b>1</b>	<b>41</b>
<b>2</b>	<b>83</b>
<b>3</b>	<b>116</b>
<b>4</b>	<b>56</b>
<b>5</b>	<b>5</b>
<b>Totale</b>	<b>315</b>

## ES. CON DATI PONDERATI

Nel caso di dati ponderati è opportuno utilizzare la funzione "rep" per esprimere il numero di volte in cui si ripete ogni elemento.

```
> esami=c(rep(0, 14), rep(1, 41), rep(2, 83),  
rep(3, 116), rep(4, 56), rep(5, 5))
```

Valutare la simmetria e l'appiattimento di questa distribuzione, disegnandone anche un grafico.

## ES. CON DATI PONDERATI

```
> gamma(esami)
```

```
[1] -0.3415149
```

```
# asimmetria negativa, coda a sinistra
```

```
> beta(esami))
```

```
[1] 2.669616
```

```
> gamma2(esami))
```

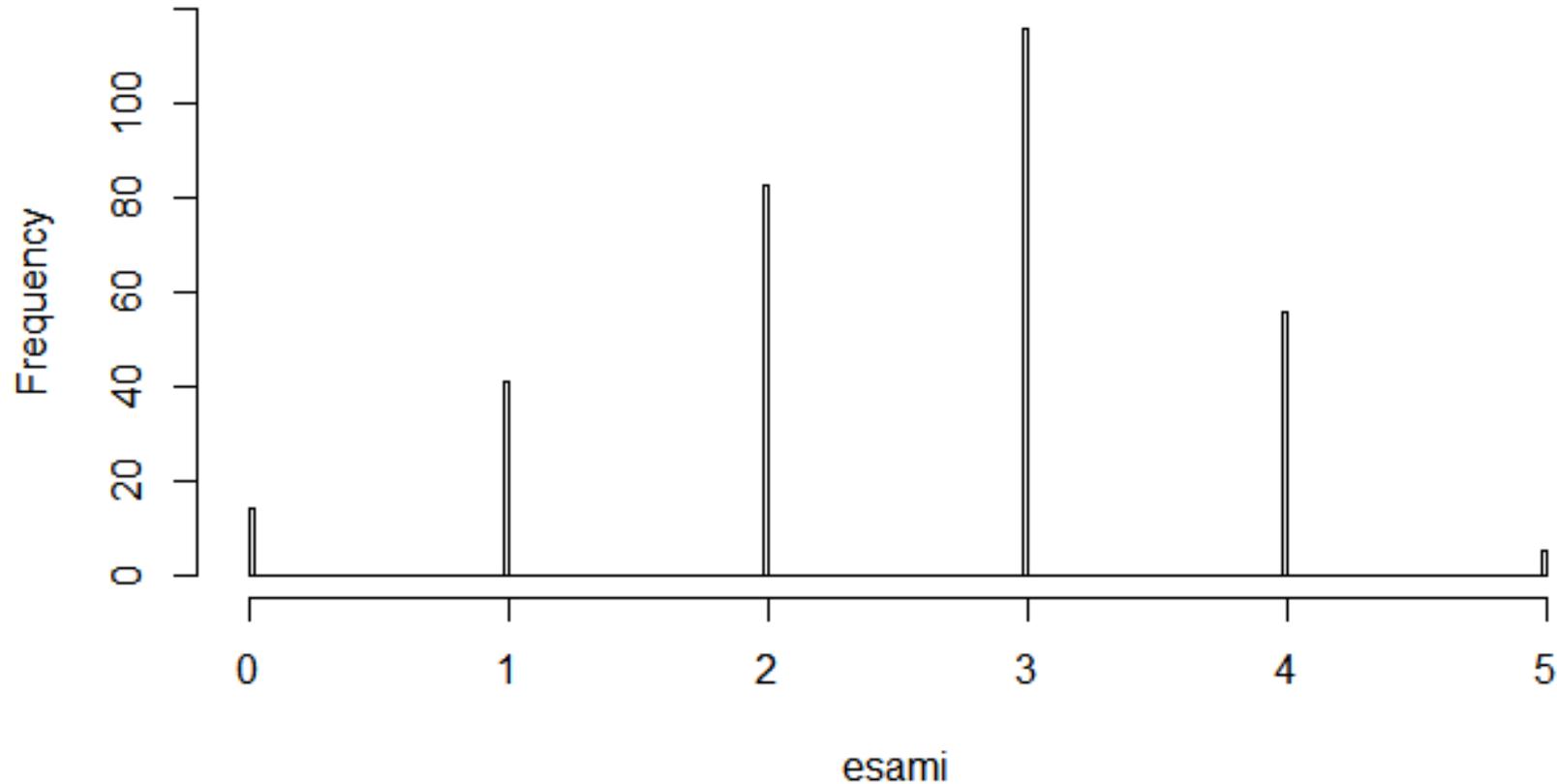
```
[1] -0.330384
```

```
# leggermente platicurtica
```

# GRAFICO DATI PONDERATI

```
> hist(esami, freq=TRUE, breaks=length(esami))
```

Histogram of esami



# ALCUNE UTILI FUNZIONI

- ▶ `sum()` calcola la somma degli elementi di un vettore di dati;
- ▶ `length()` restituisce la numerosità di un vettore;
- ▶ `range()` per trovare il minimo e il massimo di un vettore;
- ▶ `mean()` calcola la media;
- ▶ `weighted.mean(x, pesi)` calcola la media ponderata;
- ▶ `median()` calcola la mediana;
- ▶ `sd()` calcola lo scarto quadratico medio campionario di un vettore di dati;
- ▶ `var()` calcola la varianza campionaria di un vettore di dati o la covarianza tra due vettori;
- ▶ `cor()` calcola la correlazione tra due vettori;
- ▶ `summary()` riporta le principali statistiche descrittive di un vettore o di una matrice di dati.

# ALCUNI PACCHETTI UTILI

- ▶ **ggplot2**: pacchetto per migliorare e facilitare la creazione di grafici
- ▶ **plyr**: per analizzare i dati raggruppandoli in sotto insiemi o combinandoli fra di loro (analisi group-by)
- ▶ **rcpp**: per scrivere funzioni di R che richiamano codice C++
- ▶ **XML**: per creare documenti XML
- ▶ **zoo**: per l'analisi delle serie storiche
- ▶ **quantmod**: fornisce degli strumenti per il download di dati finanziari, grafici e la loro analisi
- ▶ **shiny**: per trasformare le analisi di R in applicazioni interattive per il web

<https://support.rstudio.com/hc/en-us/articles/201057987-Quick-list-of-useful-R-packages>

# ALCUNI SITI WEB UTILI

- ▶ <http://www.statmethods.net/>
- ▶ <http://www.r-bloggers.com/>
- ▶ <http://www.rdocumentation.org/>
- ▶ <http://rseek.org/>
- ▶ <http://www.inside-r.org/>
- ▶ <http://www.ats.ucla.edu/stat/r/>