

Misure di tendenza centrale

Le misure o indici di tendenza centrale descrivono sinteticamente la distribuzione di una variabile fornendo un ordine di grandezza del fenomeno.

Mediana

La mediana è il valore che divide a metà la distribuzione di una variabile ordinabile in senso crescente o decrescente. La funzione:

```
median(x, na.rm = FALSE)
```

Calcola la mediana solo per vettori numerici. La mediana, così come la media, può essere calcolata per una variabile con valori mancanti, ricorrendo all'argomento `na.rm = TRUE`. Ad esempio:

```
# dati
library(carData)

# la variabile non contiene dati mancanti
median(SLID$age)
```

```
## [1] 41
```

```
# la variabile contiene dati mancanti
median(SLID$education, na.rm = TRUE)
```

```
## [1] 12.1
```

Nel pacchetto [LabRS](#) è disponibile la funzione `med()`, che può essere utilizzata anche per le variabili ordinali.

Media aritmetica

«La media aritmetica è il tipo di media impiegato più comunemente e quello al quale, con il termine “media”, si fa in genere riferimento nel parlare comune. Viene usata per riassumere con un solo numero un insieme di dati su un fenomeno misurabile (per esempio, l'altezza media di una popolazione).

Viene calcolata sommando i diversi valori a disposizione, i quali vengono divisi per il loro numero complessivo» [Media_\(statistica\)#Media_aritmetica](#).

$$\overline{X} = \frac{\sum_{i=1}^N x_i}{N}$$

```
sum(SLID$age) / length(SLID$age)
```

```
# oppure  
mean(SLID$age)
```

```
## [1] 43.98276
```

Anche per la media aritmetica, in caso di valori mancanti, utilizziamo l'argomento `na.rm = TRUE`:

```
mean(SLID$education, na.rm = TRUE)
```

```
## [1] 12.49608
```

Trimmed mean

Dal momento che la media aritmetica risente dei valori estremi, per ridurre tale effetto possiamo calcolare la media solo sui valori centrali di una distribuzione asimmetrica.

```
x <- c(2, 3, 3, 5, 5, 6, 6, 7, 8, 60)  
  
mean(x)
```

```
## [1] 10.5
```

A tale scopo, useremo l'argomento `trim`, che può assumere valori compresi fra 0 e 0,5, e che elimina le corrispondenti proporzioni di casi (code) dall'inizio e dalla fine della distribuzione:

```
mean(x, trim=0.1)
```

```
## [1] 5.375
```

Altre medie

Le altre medie non sono disponibili nella distribuzione standard di R, e si trovano in altri pacchetti. Esse sono però facilmente calcolabili con le funzioni disponibili.

Media quadratica

La media quadratica è la radice quadrata (`sqrt()`) della media aritmetica dei valori al quadrato (`SLID$age^2`):

$$\overline{X_2} = \sqrt{\frac{\sum_{i=1}^N x_i^2}{N}}$$

Quindi:

```
sqrt(mean(SLID$age^2))
```

```
## [1] 47.40821
```

Media geometrica

La media geometrica è la radice n-esima del prodotto degli n valori:

$$\overline{X}_{\text{geom}} = \left(\prod_{i=1}^n x_i \right)^{1/n}$$

```
n = length(SLID$age)
prod(SLID$age)^(1/n)
```

```
# oppure
prod(SLID$age)^(1/length(SLID$age))
```

```
## [1] Inf
```

Se la variabile ha valori mancanti:

```
n = sum(!is.na(SLID$education))
```

```
# oppure
n = nval(SLID$education) # pacchetto LabRS

prod(SLID$education, na.rm = TRUE)^(1/n)
```

```
## 0
```

Funzione per la media geometrica

Questa è una funzione per il calcolo della media geometrica:

[g_mean.R](#)

```
g_mean<- function(x) {
  n = sum(!is.na(x))
  res <- prod(x, na.rm = T)^(1/n)
  res
}
```

oppure

g_mean.R

```
g_mean <- function(x) exp(mean(log(x), na.rm = T))
```

Media armonica

La media armonica è il reciproco della media aritmetica ($1/\text{mean}()$) dei reciproci dei valori ($1/x$):

$$H = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}, \text{ per } x_i > 0$$

Quindi:

```
1/mean(1/SLID$age)
```

```
## [1] 36.74617
```

Script di esempio

E' possibile scaricare ed eseguire lo script dell'esempio:

nome_file.R

```
# dati
library(carData)

# MEDIANA
median(SLID$age)

# mediana per variabile con dati mancanti
median(SLID$education, na.rm = TRUE)

# MEDIA ARITMETICA
mean(SLID$age)

# media per variabile con dati mancanti
mean(SLID$education, na.rm = TRUE)

# trimmed mean
x <- c(2, 3, 3, 5, 5, 6, 6, 7, 8, 60)
mean(x)
```

```
mean(x, trim = 0.1)

# MEDIA QUADRATICA
sqrt(mean(SLID$age^2))

# MEDIA GEOMETRICA

n = length(SLID$age)
prod(SLID$age)^(1/n)

# oppure
prod(SLID$age)^(1/length(SLID$age))

# media geometrica con valori mancanti

# numero di casi validi
n = sum(!is.na(SLID$education))

# oppure
n = nval(SLID$education) # pacchetto LabRS

# calcolo della media
prod(SLID$education, na.rm = TRUE)^(1/n)

# MEDIA ARMONICA
1/mean(1/SLID$age)
```

Analisi monovariata

From:

<https://www.agnesevardanega.eu/wiki/> - Ricerca Sociale con R

Permanent link:

https://www.agnesevardanega.eu/wiki/r/analisi_descrittiva/misure_di_tendenza_centrale

Last update: 11/08/2025 14:35

