

Esercitazioni R

Michele Grassi

Dipartimento di Psicologia, Università di Trieste

Istruzioni

Cosa fare:

- Rivediamo step-by-step quello che abbiamo fatto durante questi due incontri.
- Leggete le slide e implementate in R!
- Agli oggetti viene dato un nome: alla fine del ripasso andate sulla finestrella FILE e salvate il WORKSPACE su un floppy.
- Controllerò che abbiate lavorato :-)

Sommario delle funzioni

Sommario delle funzioni

Costruire e
manipolare un
vettore

Costruire e
manipolare una
matrice

Caricare un
data-frame esterno
(.txt)

Manipolare un
data-frame

Statistica
descrittiva

Costruire un
data-frame

Grafico a scatola

Cosa abbiamo visto:

- Costruire e manipolare un vettore
- Costruire e manipolare una matrice
- Caricare un data-frame esterno (.txt)
- Manipolare un data-frame
- Statistica descrittiva
- Costruire un data-frame
- Grafico a scatola
- Le sommatorie in R

Sommario delle funzioni

Costruire e manipolare un vettore
Costruire e manipolare una matrice
Caricare un data-frame esterno (.txt)
Manipolare un data-frame
Statistica descrittiva
Costruire un data-frame
Grafico a scatola

Costruire e manipolare un vettore

- Per creare un vettore di numeri desiderati:

```
vettore<-c(1,2,3,4)
vettore2<-c(seq(1,3,0.5))
```

- Per estrarre l'elemento in base alla sua posizione seriale:

```
###prendiamo il terzo elemento di vettore2:
vettore2[3]
```

```
###prendiamo una serie (dal primo al terzo):
vettore2[1:3]
```

- Una funzione viene applicata ad ogni singolo elemento del vettore:

```
sqrt(vettore)
```

Costruire e manipolare un vettore

- Costruiamo il vettore **X** contenente 25 numeri:

```
x<-c(seq(1,13,0.5))
```

```
length(x)
```

- Costruiamo due nuovi vettori **X1** e **X2** che contengono rispettivamente i primi 10 e gli ultimi 15 numeri di **X**:

```
x1<-x[1:10]
```

```
x2<-x[11:25]
```

```
length(x1)
```

```
length(x2)
```

Sommario delle
funzioni

Costruire e
manipolare un
vettore

Costruire e
manipolare una
matrice

Caricare un
data-frame esterno
(.txt)

Manipolare un
data-frame
Statistica

descrittiva
Costruire un
data-frame

Grafico a scatola

Costruire e manipolare un vettore

- Ampliamo il vettore **X** con l'elemento finale 4000 e assegnamo il nome **Xampl**

```
xampl<-c(x,4000)
```

- Se volessimo inserire il valore 4000 in posizione 10 [... l'unico modo che conosco io:-)]

```
xampl<-c(x[1:9],4000,x[10:25])
```

```
###verifichiamo....
```

```
xampl[10]
```

```
length(xampl)
```

Sommario delle
funzioni

Costruire e
manipolare un
vettore

Costruire e
manipolare una
matrice

Caricare un
data-frame esterno
(.txt)

Manipolare un
data-frame

Statistica
descrittiva

Costruire un
data-frame

Grafico a scatola

Costruire e manipolare una matrice

- Rivediamo la costruzione di una matrice:

```
### riempimento "per righe"  
matrix(x,ncol=2,byrow=T)
```

```
      [,1] [,2]  
[1,]    1  1.5  
[2,]    2  2.5  
[3,]    3  3.5  
[4,]    4  4.5  
[5,]    5  5.5  
[6,]    6  6.5  
[7,]    7  7.5  
[8,]    8  8.5  
[9,]    9  9.5  
[10,]   10 10.5  
[11,]   11 11.5  
[12,]   12 12.5  
[13,]   13  1.0  
Warning message:  
Replacement length not a multiple of the elements to replace in matrix(...)
```

Costruire e manipolare una matrice

```
### riempimento "per colonne" (byrow=F)  
matrix(x,ncol=2)
```

```
      [,1] [,2]  
[1,]  1.0  7.5  
[2,]  1.5  8.0  
[3,]  2.0  8.5  
[4,]  2.5  9.0  
[5,]  3.0  9.5  
[6,]  3.5 10.0  
[7,]  4.0 10.5  
[8,]  4.5 11.0  
[9,]  5.0 11.5  
[10,] 5.5 12.0  
[11,] 6.0 12.5  
[12,] 6.5 13.0  
[13,] 7.0  1.0
```

Warning message:

Replacement length not a multiple of the elements to replace in matrix(...)

Warning message: il vettore **X** di lunghezza 25 non basta per riempire la matrice.

R rimpiazza questo valore mancante con il primo!!!

Sommario delle
funzioni
Costruire e
manipolare un
vettore

Costruire e
manipolare una
matrice

Caricare un
data-frame esterno
(.txt)

Manipolare un
data-frame
Statistica

descrittiva
Costruire un
data-frame

Grafico a scatola

Costruire e manipolare una matrice

- Abbiamo già costruito il vettore **xampl** contenente 26 numeri. Se lo utilizziamo nel comando `matrix()` non ci sarà bisogno di rimpiazzo:

```
matrix(xampl, ncol=2, byrow=T)
```

```
###oppure
```

```
matrix(xampl, ncol=2, byrow=F)
```

- Con 25 numeri (**X**) e cinque colonne la matrice dovrebbe avere 5 righe, proviamo:

```
matrix(x, ncol=5, byrow=T)
```

```
###oppure
```

```
matrix(x, ncol=5, byrow=F)
```

Sommario delle

funzioni

Costruire e
manipolare un
vettore

Costruire e
manipolare una
matrice

Caricare un
data-frame esterno
(.txt)

Manipolare un
data-frame
Statistica

descrittiva
Costruire un
data-frame

Grafico a scatola

Costruire e manipolare una matrice

```
matrix(x,ncol=5,byrow=T)
```

```
      [,1] [,2] [,3] [,4] [,5]  
[1,]  1.0  1.5  2.0  2.5  3.0  
[2,]  3.5  4.0  4.5  5.0  5.5  
[3,]  6.0  6.5  7.0  7.5  8.0  
[4,]  8.5  9.0  9.5 10.0 10.5  
[5,] 11.0 11.5 12.0 12.5 13.0
```

```
###oppure
```

```
matrix(x,ncol=5,byrow=F)
```

```
      [,1] [,2] [,3] [,4] [,5]  
[1,]  1.0  3.5  6.0  8.5 11.0  
[2,]  1.5  4.0  6.5  9.0 11.5  
[3,]  2.0  4.5  7.0  9.5 12.0  
[4,]  2.5  5.0  7.5 10.0 12.5  
[5,]  3.0  5.5  8.0 10.5 13.0
```

Sommario delle
funzioni

Costruire e
manipolare un
vettore

Costruire e
manipolare una
matrice

Caricare un
data-frame esterno
(.txt)

Manipolare un
data-frame

Statistica
descrittiva

Costruire un
data-frame

Grafico a scatola

Costruire e manipolare una matrice

■ Abbiamo visto come manipolare una matrice:

- ◆ bisogna dare un nome alla matrice

```
m1<-matrix(x,ncol=5,byrow=F)
```

- ◆ utilizziamo poi il nome e le parentesi quadre

```
m1[,]
```

■ alla sinistra della virgola viene specificato il numero (o la serie) di righe, alla destra il numero (o la serie) di colonne.

```
### riga 1 colonna 3
```

```
m1[1,3]
```

```
### riga da 1 a 3 e TUTTE le colonne
```

```
m1[1:3,]
```

```
### tre righe e colonne 2 e 3
```

```
m1[1:3,2:3]
```

Sommario delle

funzioni

Costruire e

manipolare un

vettore

Costruire e

manipolare una

matrice

Caricare un

data-frame esterno

(.txt)

Manipolare un

data-frame

Statistica

descrittiva

Costruire un

data-frame

Grafico a scatola

Caricare un data-frame esterno (.txt)

Carichiamo in R un data-frame presente sul blocco note del computer (file.txt). Il file è "anoressia.txt".

- con il mouse facciamo "*copia*" del nostro file;
- andiamo poi su *risorse del computer* e selezioniamo disco locale *C:*;
- *entriamo* nella cartella *programmi* e nelle sotto-cartelle *R* e *rw1080* (il numero dipende dalla versione!!);
- si "*incolla*" finalmente il file "anoressia.txt"!

Il file è nelle cartelle di R. Lo richiameremo semplicemente come "anoressia.txt" all'interno di read.table().

Caricare un data-frame esterno (.txt)

- Utilizziamo l'argomento `header=T` per sfruttare l'intestazione del data-frame:

```
read.table("anoressia.txt",header=T)
```

- Prima di passare alla manipolazione del data-frame, sarà bene dargli un nome:

```
df<-read.table("anoressia.txt",header=T)
```

- D'ora in poi per richiamare i nostri dati basta digitare **df**

Sommario delle
funzioni

Costruire e
manipolare un
vettore
Costruire e
manipolare una
matrice

Caricare un
data-frame esterno
(.txt)

Manipolare un
data-frame
Statistica
descrittiva
Costruire un
data-frame
Grafico a scatola

Manipolare un data-frame

Abbiamo detto che un data-frame non è altro che una tabella di dati. Tali dati sono numeri che derivano dagli stati che un campione di soggetti assume su una serie di proprietà sotto analisi. Ogni colonna di modalità viene quindi identificata dal nome della variabile a cui si riferisce.

- Possiamo manipolare il data frame in quanto *matrice di numeri*;
- possiamo altrimenti (... e questo è il vantaggio del data-frame) utilizzare il nome delle variabili in gioco.

Negli esempi che seguono utilizzerò il data frame "anoressia.txt" che abbiamo chiamato **df**. Ma sicuramente nulla vi vieta di utilizzare altri data-frame contenuti in R, attraverso i comandi: (es.)

- **data()** per visualizzare i data-frame esistenti:
- **library(base)** per caricare la libreria (package) in cui si trovano i dati;
- **data(USArrests)** per caricare i dati desiderati che si trovano all'interno della libreria selezionata.

Sommario delle
funzioni
Costruire e
manipolare un
vettore
Costruire e
manipolare una
matrice
Caricare un
data-frame esterno
(.txt)

Manipolare un
data-frame

Statistica
descrittiva
Costruire un
data-frame
Grafico a scatola

MANIPOLARE UN DATA-FRAME COME MATRICE DI NUMERI:

- selezionare riga e colonna del data-frame "anoressia.txt":

```
###riga 1 e tutte le colonne:  
df[1,]
```

```
###riga 1 e colonna 2:  
df[1,2]
```

- selezionare una sequenza di righe e di colonne:

```
###dalla riga 1 alla riga 10. Tutte le  
###colonne:  
df[1:10,]
```

```
###dalla riga 1 alla riga 10 per le colonne  
### dalla 2 alla 3:  
df[1:10,2:3]
```

Sommario delle
funzioni
Costruire e
manipolare un
vettore
Costruire e
manipolare una
matrice
Caricare un
data-frame esterno
(.txt)

Manipolare un
data-frame

Statistica
descrittiva
Costruire un
data-frame
Grafico a scatola

MANIPOLARE UN DATA-FRAME UTILIZZANDO LE VARIABILI:

- selezionare la colonna *therapy*:

```
df$therapy
```

- selezionare le righe che si riferiscono ad uno specifico livello di therapy (es. *livello "b"*):

```
df[df$therapy=="b" , ]
```

Sommario delle
funzioni
Costruire e
manipolare un
vettore
Costruire e
manipolare una
matrice
Caricare un
data-frame esterno
(.txt)

Manipolare un
data-frame

Statistica
descrittiva
Costruire un
data-frame
Grafico a scatola

Con il comando precedente possiamo separare le «prestazioni» dei tre gruppi di ragazze anoressiche sottoposte a tre diversi trattamenti terapeutici;

- gruppo trattamento «behaviorista»:

```
grb<-df[df$therapy=="b" , ]
```

- gruppo trattamento «famiglia»:

```
grf<-df[df$therapy=="f" , ]
```

- gruppo trattamento «controllo»:

```
grc<-df[df$therapy=="c" , ]
```

Sommario delle

funzioni

Costruire e
manipolare un
vettore

Costruire e
manipolare una
matrice

Caricare un
data-frame esterno
(.txt)

Manipolare un
data-frame

Statistica
descrittiva

Costruire un
data-frame

Grafico a scatola

grb, **grf** e **grc** sono tre data-frame che derivano dal data-frame complessivo **df**. In ciascuno di essi troviamo i valori della variabili *after* e *before* relativi ad uno specifico gruppo di ragazze anoressiche. Vi ricordo che potevano separare le ragazze in base al trattamento semplicemente utilizzando i numeri di riga, dal momento che i livelli di *therapy* erano messi in ordine:

```
grb<-df[1:29, ]  
grf<-df[30:46, ]  
grc<-df[47:72, ]
```

Se i livelli b,c,f sono disposti in modo disordinato, per selezionare (e «compattare») le righe di interesse si utilizza la sintassi precedente!!!

Sommario delle
funzioni
Costruire e
manipolare un
vettore
Costruire e
manipolare una
matrice
Caricare un
data-frame esterno
(.txt)

Manipolare un
data-frame
Statistica
descrittiva
Costruire un
data-frame
Grafico a scatola

COME SONO ANDATE LE COSE IN CIASCUN GRUPPO DI RAGAZZE ANORESSICHE?

QUALE TRATTAMENTO HA DATO I RISULTATI MIGLIORI NEL CAMPIONE ANALIZZATO?

Queste sono le domande fondamentali alle quali si cerca di rispondere utilizzando la statistica descrittiva. Prima di tutto dobbiamo separare i risultati relativi ai tre gruppi di ragazze. Questo l'abbiamo già fatto creando i tre data-frame **grb**, **grc** e **grf**. L'operazione successiva consiste nello stabilire quale sia la **variabile dipendente** dell'esperimento che ci permette di confrontare i tre gruppi. Questa è l'**incremento di peso** per ciascuna ragazza, definito come la differenza tra il peso dopo (*after*) il trattamento terapeutico e il peso prima (*before*).

CREIAMO IL VETTORE **incremento di peso** PER CIASCUN GRUPPO DI RAGAZZE:

```
incpb<-c(grb$after-grb$before)  
incpc<-c(grc$after-grc$before)  
incpf<-c(grf$after-grf$before)
```

Abbiamo ora gli incrementi di peso di ciascuna ragazza, per ciascuno dei tre gruppi di terapia.

- Sommario delle funzioni
- Costruire e manipolare un vettore
- Costruire e manipolare una matrice
- Caricare un data-frame esterno (.txt)
- Manipolare un data-frame
- Statistica descrittiva**
- Costruire un data-frame
- Grafico a scatola

Come riassumere le tre variabili in **pochi numeri** informativi sulla *tendenza centrale* e sulla *dispersione* all'interno di ciascun campione di ragazze anoressiche?

- Possiamo calcolare media e varianza per ciascuna variabile:

```
###(es.) per la variabile incpb:  
mean(incpb)  
var(incpb)
```

- Possiamo calcolare il primo quartile, la mediana e il terzo quartile:

```
###(es.) per la variabile incpf:  
summary(incpf)
```

- Sommario delle funzioni
- Costruire e manipolare un vettore
- Costruire e manipolare una matrice
- Caricare un data-frame esterno (.txt)
- Manipolare un data-frame
- Statistica descrittiva**
- Costruire un data-frame
- Grafico a scatola

Il data-frame può essere costruito direttamente in R, attraverso il comando **data.frame()**.

- Prima di tutto costruiamo le variabili da inserire nel data-frame:

```
x<-c(1,2,3,4,5)
```

```
y<-c(11,12,13,14,15)
```

```
f<-factor(c("bello","brutto","bello","bello","brutto"))
```

- Mostriamo ora le variabili sotto forma di data-frame:

```
data.frame(f,x,y)
```

```
      f x  y
1 bello 1 11
2 brutto 2 12
3 bello 3 13
4 bello 4 14
5 brutto 5 15
```

- Dando un nome al data-frame appena costruito possiamo applicare le manipolazioni discusse:

```
dfnew<-data.frame(f,x,y)
dfnew$f
dfnew$x
dfnew[1:3,]
dfnew[dfnew$f=="bello",]
dfnew[dfnew$f=="brutto",]
```

Sommario delle
funzioni

- Costruire e
manipolare un
vettore
- Costruire e
manipolare una
matrice
- Caricare un
data-frame esterno
(.txt)
- Manipolare un
data-frame
- Statistica
descrittiva
- Costruire un
data-frame**
- Grafico a scatola

Possiamo dare una descrizione grafica delle nostre variabili, utilizzando il GRAFICO A SCATOLA.

- La funzione da applicare alla variabile di interesse è:

```
boxplot(incpb)
```

- La descrizione dei valori numerici disegnati si ottiene con la classica funzione di estrazione \$:

```
boxplot(incpb)$stats
```

```
[ ,1]
```

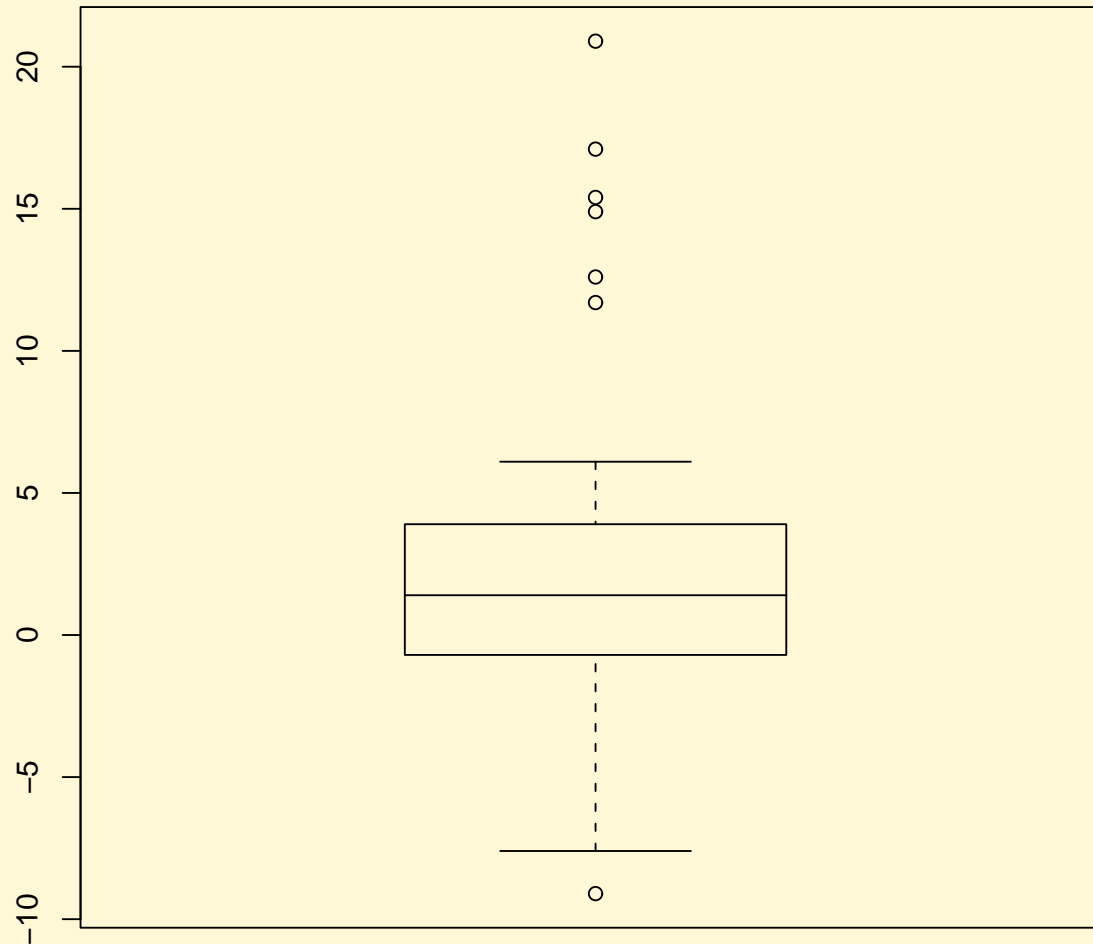
```
[1,] -7.6 valore adiacente inferiore
```

```
[2,] -0.7 primo quartile
```

```
[3,] 1.4 mediana
```

```
[4,] 3.9 terzo quartile
```

```
[5,] 6.1 valore adiacente superiore
```



Sommario delle
funzioni
Costruire e
manipolare un
vettore
Costruire e
manipolare una
matrice
Caricare un
data-frame esterno
(.txt)
Manipolare un
data-frame
Statistica
descrittiva
Costruire un
data-frame

Grafico a scatola

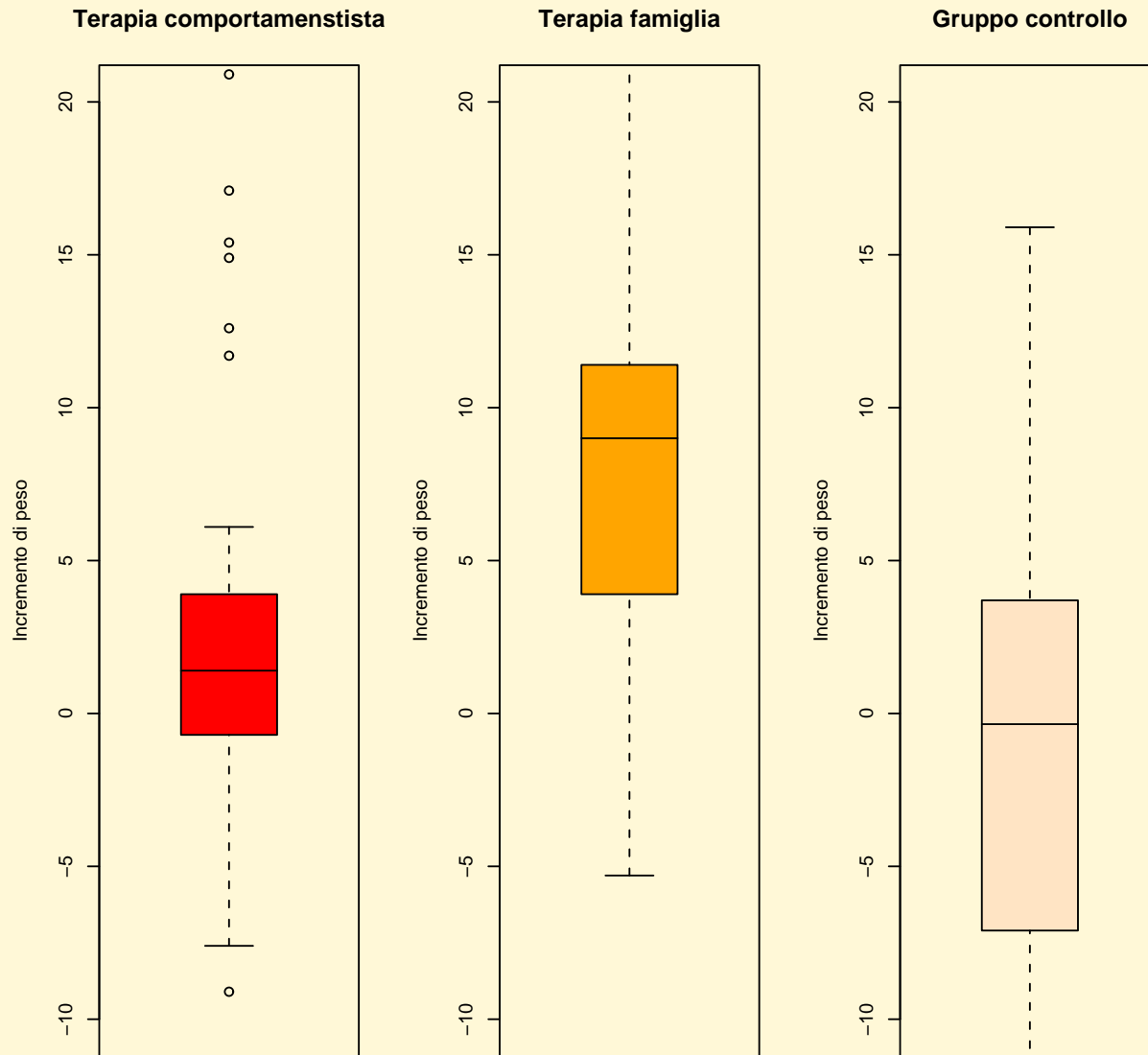
A noi interessa sfruttare il **boxplot** per confrontare direttamente (con solo 5 numeri) come sono andate le cose nei diversi gruppi di terapia di ragazze anoressiche:

- Creiamo tre boxplot per le tre variabili *incremento di peso* nella stessa sessione grafica:

```
par(mfrow=c(1,3))
boxplot(incpb,col="red",ylim=c(-10,20),
        ylab="Incremento di peso",
        main="Terapia comportamenstista")
boxplot(incpf,col="orange",ylim=c(-10,20),
        ylab="Incremento di peso",
        main="Terapia famiglia")
boxplot(incpc,col="bisque",ylim=c(-10,20),
        ylab="Incremento di peso",
        main="Gruppo controllo")
```

Sommario delle
funzioni
Costruire e
manipolare un
vettore
Costruire e
manipolare una
matrice
Caricare un
data-frame esterno
(.txt)
Manipolare un
data-frame
Statistica
descrittiva
Costruire un
data-frame

Grafico a scatola



Sommario delle
funzioni
Costruire e
manipolare un
vettore
Costruire e
manipolare una
matrice
Caricare un
data-frame esterno
(.txt)
Manipolare un
data-frame
Statistica
descrittiva
Costruire un
data-frame
Grafico a scatola

Per semplificare la sintassi precedente è necessario inserire la variabile *incremento di peso complessiva* nel data frame **df**.

- costruiamo un vettore **incp** che contenga gli incrementi di peso relativi a ciascun gruppo di terapia:

```
incp<-c(incpb,incpf,incpc)
```

- ampliamo il data-frame **df** con tale vettore:

```
df2<-data.frame(df,incp)
```

```
df2
```

- La sintassi per ottenere i tre box plot precedenti si semplifica alquanto:

```
boxplot(df2$incp~df2$therapy,df2,col="red",  
        xlab="tipo di terapia",  
        ylab="incremento di peso")
```

```
###opp.
```

```
plot(df2$therapy,df2$incp,df2,col="cyan",  
     xlab="tipo di terapia",  
     ylab="incremento di peso")
```

```
###opp.
```

```
plot(df2$incp~df2$therapy,df2,col="yellow",  
     xlab="tipo di terapia",  
     ylab="incremento di peso")
```

Sommario delle
funzioni
Costruire e
manipolare un
vettore
Costruire e
manipolare una
matrice
Caricare un
data-frame esterno
(.txt)
Manipolare un
data-frame
Statistica
descrittiva
Costruire un
data-frame

Grafico a scatola

Sommario delle

funzioni

Costruire e

manipolare un

vettore

Costruire e

manipolare una

matrice

Caricare un

data-frame esterno

(.txt)

Manipolare un

data-frame

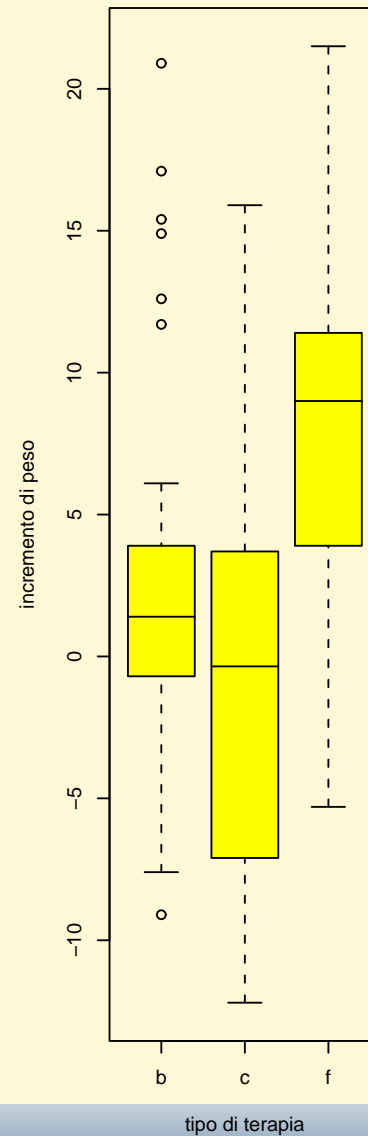
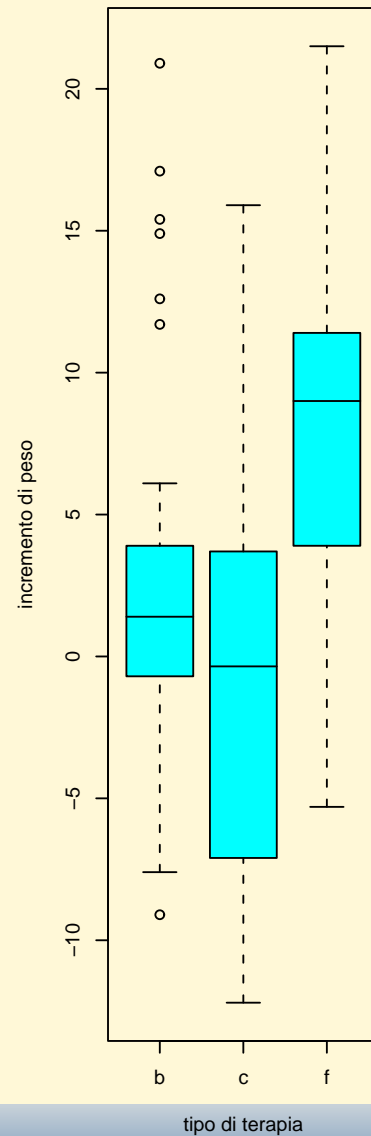
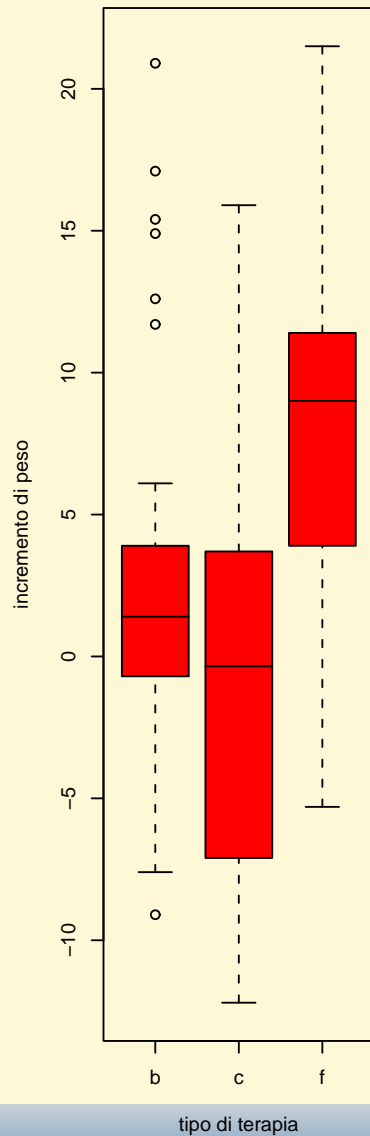
Statistica

descrittiva

Costruire un

data-frame

Grafico a scatola



Sommatorie in R

- Consideriamo la seguente distribuzione della variabile X : $\{3, 7, 1, 9, 2\}$.
- Usiamo l'indice i per fare riferimento alle diverse u.s. che costituiscono la distribuzione.
- Il valore assunto dalla variabile X nel caso di una generica unità statistica viene indicato con X_i .
- Quando l'indice i assume il valore 2, 3, o 5 questo significa che facciamo riferimento al valore assunto da X nel caso della seconda, terza o quinta unità statistica.
- Esempio: $X_1 = 3$; $X_4 = 9$; $X_2 = 7$.

- Un simbolo di cui viene fatto un grande uso nella statistica è quello di sommatoria:

$$\sum_{i=1}^n X_i = X_1 + X_2 + \cdots + X_n$$

- Nel caso della distribuzione precedente avremo:

$$\sum_{i=1}^5 X_i = 3 + 7 + 1 + 9 + 2 = 22$$

- Quando non ci sono ambiguità, la notazione precedente può essere semplificata con

$$\sum_i X_i \quad \text{oppure} \quad \sum X_i$$

- Un caso più complicato si ha quando le singole osservazioni fanno parte di più gruppi.
- Per esempio, assumiamo che vi siano 10 studenti in 3 aule. Per distinguere tutte queste osservazioni abbiamo bisogno di 2 indici, uno per gli studenti e uno per le aule. Sia i l'indice per gli studenti e j l'indice per le aule. Quindi l'indice i assume valori da 1 a 10 e l'indice j da 1 a 3.
- Ciascuno studente sarà dunque identificato da due indici: X_{ij} . Se usiamo la convenzione per cui l'indice i precede l'indice j , allora $X_{7,3}$ indicherà il settimo studente nella terza aula.

- Spesso accade che vogliamo sommare i punteggi di tutti gli individui in tutti i gruppi. Per fare questo usiamo la seguente notazione:

$$\sum_{j=1}^J \sum_{i=1}^I X_{ij}$$

- Nel caso dell'esempio precedente, avremo

$$\sum_{j=1}^3 \sum_{i=1}^{10} X_{ij} = X_{1,1} + X_{1,2} + \cdots + X_{3,10}$$

Regole per la manipolazione delle sommatorie

Regole per la manipolazione delle sommatorie

Regola 1
Regola 2
Regola 3
Regola 4
Regola 5
Regola 6
Regola 7
Sommatoria parziale

- Se a è una costante, allora

$$\sum_{i=1}^n a = n \times a$$

- Esempio: se $a = 10$ e $n = 3$ allora

$$\sum_{i=1}^3 10 = 10 + 10 + 10 = 3 \times 10$$

$a < -10$

$a + a + a$

[1] 30

$3 * a$

[1] 30

- Se ciascuna delle osservazioni che entrano nella sommatoria viene moltiplicata per la costante a , allora

$$\sum_{i=1}^n aX_i = a \sum_{i=1}^n X_i$$

```
x<-c(3,5,1)
```

```
a<-2
```

```
sum(x*a)
```

```
[1] 18
```

```
a*sum(x)
```

```
[1] 18
```

- Se dobbiamo eseguire un'operazione algebrica (quadrato, radice, logaritmo, ecc.) sulle singole osservazioni che devono essere sommate, questa operazione deve essere eseguita prima di sommare le n osservazioni.
- Esempio:

$$\sum_{i=1}^n X_i^2 = X_1^2 + X_2^2 + \cdots + X_n^2$$

```
sum(x^2)
```

```
[1] 35
```

```
x[1]^2+x[2]^2+x[3]^2
```

```
[1] 35
```

- Se la sola operazione che deve essere eseguita sulle n osservazioni prima della sommatoria è essa stessa una somma (o sottrazione), allora **la sommatoria può essere distribuita.**

- Esempio:

$$\sum_{i=1}^n (X_i - Y_i) = \sum_{i=1}^n X_i - \sum_{i=1}^n Y_i$$

```
x<-c(3,5,1)
```

```
y<-c(4,7,6)
```

```
sum(x-y)
```

```
[1] -8
```

```
sum(x)-sum(y)
```

```
[1] -8
```

- Se ciascuna osservazione viene misurata su due variabili, X_i e Y_i , allora

$$\sum_{i=1}^n X_i Y_i = X_1 Y_1 + X_2 Y_2 + \dots + X_n Y_n$$

dobbiamo eseguire prima il prodotto tra i punteggi appaiati e poi la sommatoria.

```
sum(x*y)
```

```
[1] 53
```

```
x[1]*y[1]+ x[2]*y[2]+ x[3]*y[3]
```

```
[1] 53
```

```
###non si può distribuire la sommatoria
```

```
###all'interno di un prodotto tra vettori!!!
```

```
sum(x)*sum(y)
```

```
[1] 153 sbagliato!!!
```

$$\sum_{i=1}^n aX_iY_i = a \sum_{i=1}^n X_iY_i$$

```
sum(a*x*y)
[1] 106
a*sum(x*y)
[1] 106
```

Regole per la
manipolazione
delle sommatorie

Regola 1

Regola 2

Regola 3

Regola 4

Regola 5

Regola 6

Regola 7

Sommatoria
parziale

$$\sum_{i=1}^n (aX_i + bY_i) = a \sum_{i=1}^n X_i + b \sum_{i=1}^n Y_i$$

a < -2

b < -3

x

[1] 3 5 1

y

[1] 4 7 6

sum(a*x+b*y)

[1] 69

sum(a*x)+sum(b*y)

[1] 69

a*sum(x)+b*sum(y)

[1] 69

Regole per la
manipolazione
delle sommatorie

Regola 1

Regola 2

Regola 3

Regola 4

Regola 5

Regola 6

Regola 7

Sommatoria
parziale

La sommatoria generica

$$\sum_{i=1}^n$$

indica la sommatoria di **TUTTI** i numeri di un **vettore**: questo è ciò che fa il comando **sum()**.

```
x<-c(10,20,30,40,50,60,70,80,90,100)  
sum(x)
```

$$\sum_{i=1}^n x_i = \sum_{i=1}^{10} x_i = x_1 + x_2 + \dots + x_{10}$$

La sommatoria parziale

$$\sum_{i=1}^5 x_i$$

si ottiene specificando la **serie di numeri del vettore x** su cui applicare la sommatoria.

```
sum(x[1:5])
```

$$\sum_{i=1}^5 x_i = x_1 + x_2 + x_3 + x_4 + x_5$$

Esempio

Considerando i vettori

$x \leftarrow c(10, 20, 30, 40, 50, 60, 70, 80, 90, 100)$

$y \leftarrow c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10)$

semplifichiamo la seguente sommatoria con le regole prima viste:

$$\sum_{i=1}^5 \left(x_i \left(\sum_{i=1}^{10} y_i \right) / 10 - 12 \right) =$$

$$\sum_{i=1}^5 x_i \left(\sum_{i=1}^{10} y_i \right) / 10 - 5 \times 12 =$$

$$\bar{y} \sum_{i=1}^5 x_i - 5 \times 12$$

Verifichiamo in R la correttezza della nostra semplificazione:

```
x
```

```
[1] 10 20 30 40 50 60 70 80 90 100
```

```
y
```

```
[1] 1 2 3 4 5 6 7 8 9 10
```

```
sum(x[1:5]*(sum(y[1:10])/10)-12)
```

```
[1] 765
```

```
mean(y)*sum(x[1:5])-12*5
```

```
[1] 765
```

```
###opp.
```

```
sum(y)/length(y)*sum(x[1:5])-12*5
```

```
[1] 765
```

Esercizio

Utilizzando i vettori X e Y di prima, si semplifichi la seguente espressione:

$$\sum_{i=1}^7 \left(X_i - \left(\sum_{i=1}^n y_i / n \right) \right)$$

Verificate in R che i risultati della forma intera e semplificata coincidano!!