

Funciones en



Francisco Balao
23/06/2015

Funciones en R

- ¿Qué es una función?
- ¿Por qué necesitamos crear una función?
- ¿Cómo hacer una función?
- Bucles y otros elementos
- Recomendaciones finales

¿Qué es una función?

Una función es un grupo de **instrucciones** con un objetivo en particular

Las funciones reciben **datos** (entrada) y deben entregar un **resultado** (salida). Las funciones pueden contener argumentos que modifican las instrucciones.



¿Por qué necesitamos crear una función en R?

- 1) No encuentras tu función
- 2) Tareas repetitivas

Handwritten mathematical notes and formulas:

- $|b(T, a, b)| \leq 2$
- $\varphi(\sigma_1 t) \varphi(\sigma_2 t) = \varphi(\sqrt{\sigma_1^2 + \sigma_2^2} t)$
- $\ell(\omega) = \frac{\sum_{k=1}^n p_k^* \log_2 \frac{1}{p_k}}{\sum_{k=1}^n p_k^*}$
- $\eta_1 = \sum_{k=1}^n a_k \xi_k$
- $y = \phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$
- $S(\alpha, \tau) = \frac{2}{\pi} \int_0^{\pi} \frac{\sin \alpha t}{t} dt$
- $P(\eta < y | \xi = x) = \sup_{y < y', y' < x} P(\eta < y' | \xi = x)$
- $W_k = \binom{k}{n} p^k (1-p)^{n-k}$
- $g^{-1} \cdot g = e$
- $f(x) = \frac{1}{2} \left| \int_{|x|>A} f(x) \log_2 \frac{1}{f(x)} dx \right| < \epsilon$
- $\int_{-\infty}^{\infty} dG_k(x) \geq \frac{1}{2} \sum_{k=0}^{\infty} e^{-\frac{k^2}{2n}} = H(k)$
- $f_{n-1}(u) = \int_{-\infty}^{\infty} f_n(u) f_n(t-u) du = \frac{2^{n+1} e^{-2t}}{n!} \lim_{t \rightarrow \infty} (e^{2t}) = 0$
- $\log \varphi(t) = i j^t t - c |t|^a [1 + i b \frac{t}{|t|} \omega(t, a)]$
- $\int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du = \sqrt{2\pi}$
- $\prod_{m=1}^r \Gamma(\frac{1}{m}) = r$
- $|X \cup Y| = |X| + |Y| - |X \cap Y|$
- $g: X \rightarrow X \cap W$
- $Q(A) = \int_A f(x) dx$
- $Q(A) = \int_A f(x) dx$
- $Q(A) = \int_A f(x) dx$
- $\lim_{N \rightarrow \infty} \int_{-N}^N f(x) dx \geq \int_{-\infty}^{\infty} f(x) dx$
- $D^2(\hat{y}_n) \leq \frac{k}{n} + 2k \left(\frac{1}{k} \sum_{k=1}^n R(k) \right)$
- $\det(M^*) = \det(M) + \det(M^*) = \det(M)$
- $h(x, y) = \frac{1}{2\pi} [\sqrt{2} e^{-\frac{x^2}{2}} - e^{-x^2}]$
- $1/M(\epsilon, \delta, n) \leq C \frac{n}{m-n}$

¿Cómo hacer una función?

Definiendo funciones:

```
> mifuncion1 <- function(arg1, arg2, ...) {  
  cuerpo_de_la_funcion  
}
```

Llamando mi función:

```
> mifuncion1(arg1=..., arg2=...)
```

1er Ejemplo: la media

Media aritmética

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

```
> media <-function(x) { sum(x)/length(x) }
```

x: vector de números

```
> media(c(1,2,3,10))
```

```
[1] 4
```

1er Ejemplo: la media

Resumen

```
> resumen <-function(x, tr=0.2) {  
  cat(" Parámetros de localización\n")  
  cat("media:",mean (x), "mediana_muestral:",  
median(x), "media_recortada:", mean (x, tr), "\n")}  
  
> resumen(c(1,2,3,10,5,3,2,1,3,4,6))  
Parámetros de localización  
media: 3.636364 mediana_muestral: 3 media_recortada:  
3.142857
```

2º Ejemplo: utilizando valores por defecto

$$CV = \frac{s}{|\bar{x}|}$$

Coeficiente de variación

```
>var.coeff<- function (x, square = FALSE)
{
x2<-na.omit(x)
n <- length(x2)
V <- sqrt((n - 1) * var(x2)/n)/mean(x2)
if (square = TRUE)
V <- V^2
V
}
```


2º Ejemplo: utilizando valores por defecto

$$CV = \frac{s}{|\bar{x}|}$$

Coeficiente de variación

```
> var.coeff(c(3,4,5,2,4,5,6))  
[1] 0.3006137
```

```
> var.coeff(c(3,4,5,2,4,5,6), square=TRUE)  
[1] 0.0903686
```

Bucles y otros elementos

- Un **bucle** es la ejecución repetidas veces de un mismo conjunto de sentencias

- Bucles **for**

Se le indica el número de iteraciones

```
for (i in x:y){
```

```
  z[i]<- cuerpo del bucle }
```

Bucles y otros elementos

- Un **bucle** es la ejecución repetidas veces de un mismo conjunto de sentencias

- Bucles **for**

Se le indica el número de iteraciones

```
for (i in x:y){
```

```
  z[i]<- cuerpo del bucle }
```

3^{er} Ejemplo

Elevar al cuadrado un vector

```
> z <- NULL
> for (i in 1:10) {z[i]<-i^2}
> z
 [1]  1  4  9 16 25 36 49 64 81
100
```

3^{er} Ejemplo

Calculando la media una vez más

```
> medialoop <- function(x) {  
  total <- 0  
  n <- length(x)  
  for (i in seq_along(x)) {  
    total <- total + (x[i] / n)  
  }  
  total  
}  
> medialoop(rnorm(100, 5, 2))  
[1] 5.188459
```

3^{er} Ejemplo

A veces los bucles son ineficientes en R

```
> system.time(media_loop(rnorm(10000,5,2)))
user  system elapsed
0.004   0.000   0.005

> system.time(mean(rnorm(10000,5,2)))
user  system elapsed
0.001   0.000   0.001
```

El comando **if**

- El comando ***if*** permite implementar condiciones en la función.
- Si se cumple una condición entonces se realiza una acción
- Puede ir acompañado de ***else***. Si no se cumple la condición entonces se realiza otra acción

```
if (condición){ acción } else {acción_alternativa}
```

4º Ejemplo

```
> barplotfran<-function(x,barra=c("var","sd","se")){
  #Definimos la función del error estándar
  std <- function(x) sd(x)/sqrt(length(x))

  # calculamos el valor que queramos, varianza, desviación estándar o error estándar

  if (barra=="var") { barra2<-tapply(x[,1],x[,2],var)}
  if (barra=="sd") { barra2<-tapply(x[,1],x[,2],sd)}
  if (barra=="se") { barra2<-tapply(x[,1],x[,2],std)}

#Calculamos la media
medias<-tapply(x[,1],x[,2],mean)

#Definimos el límite del eje y
mmax<-max(medias)*2

#Dibujamos las barras
bp<-barplot(medias, ylim=c(0,mmax))

#Dibujamos las barras
barra2<-medias+barra2
segments(bp,medias,bp,barra2)
segments(bp-0.05,barra2,bp+0.05,barra2)
abline(h=0)
}
```


4º Ejemplo

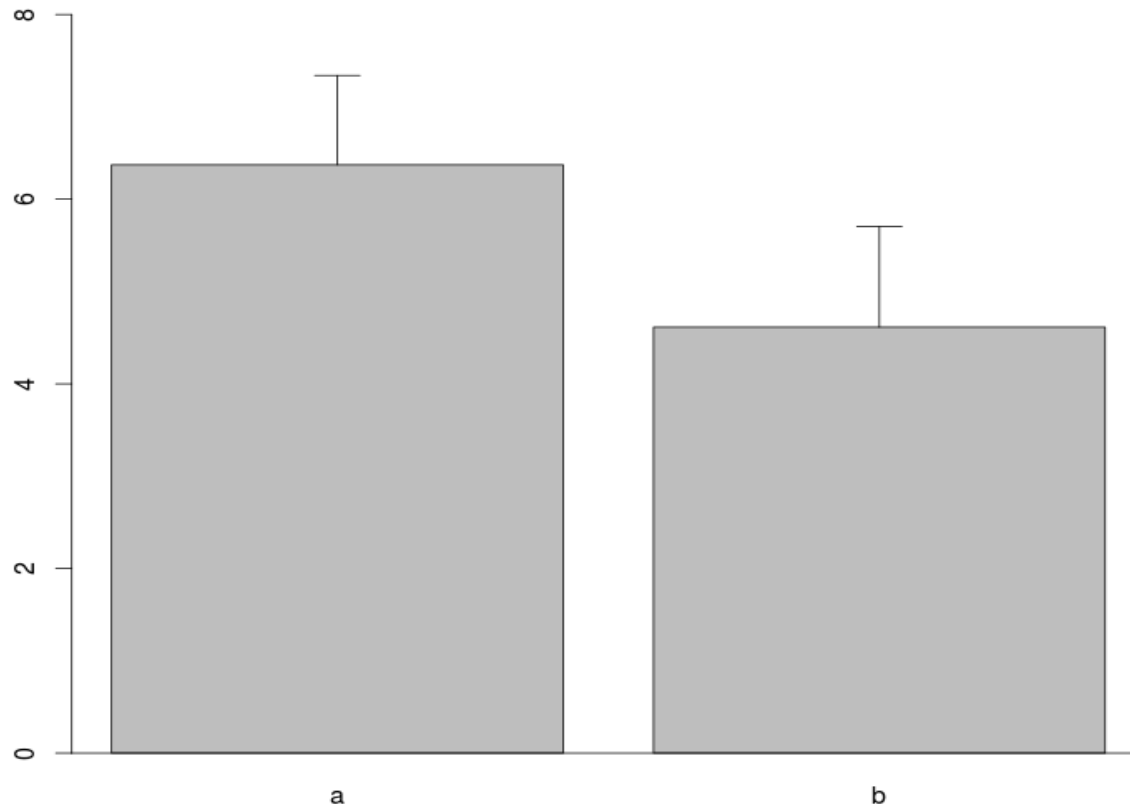
```
> datos<-data.frame (x=rnorm(20, mean=5, sd=2), y=rep(c("a","b"),10))
```

```
> datos
```

	x	y
1	48.52913	a
2	52.47802	b
3	47.97604	a
4	47.11545	b
5	51.75183	a
6	44.71220	b
7	50.48086	a
8	50.09812	b
9	49.87890	a
10	51.79377	b
11	47.24488	a
12	49.12151	b
13	49.66861	a
14	49.33583	b
15	51.16097	a
16	49.89615	b
17	53.29374	a
18	49.52362	b
19	48.96675	a
20	48.15767	b

4º Ejemplo

```
> barplotfran(datos, barra="se")
```



Para finalizar, recuerda...

- Para asignar funciones se utiliza la palabra ***function***
- Evitar asignar nombres de funciones existentes
- Es útil utilizar valores por defecto en los argumentos