

Introducció a R

“R is really important to the point that it’s hard to overvalue it”

Daryl Pregibon at NY Times

R – Visió general

- Pàgina web: http://cran.r-project.org/
- És un software gratuït i de codi lliure
- És un paquet que permet el tractament estadístic de dades.
- Però R és molt més:
  - És un llenguatge de programació orientat a objectes. R és un llenguatge funcional, un avaluador d'expressions i un intèrpret extensible.
  - Realitza gràfics adaptables i d'alta qualitat.
  - És una eina de càlcul. Permet fer operacions amb matrius y vectors de qualsevol dimensionalitat.
  - Permet un emmagatzematge senzill i manejable d'objectes i funcions. Les comandes proporcionen un output mínim, però els resultats intermedis es poden guardar en objectes.
- És un entorn amb moltes de les tècniques estadístiques clàssiques i actuals implementades:
  - R té una base amb uns paquets estàndards i després tot un seguit d'implementacions afegides (inconvenient: pot haver certa diversitat sintàctica en el seu ús)

R – Instal·lació

1. Obriu la pàgina web del repositori 'CRAN': http://cran.r-project.org/
2. Cliqueu en l'opció 'Mirrors' del menú de l'esquerra i escolliu un dels servidors (normalment el temps de descarrega diferirà molt poc d'un a un altre)
3. Segons el vostre sistema operatiu, feu clic en Linux, MacOS X o Windows.
4. Si useu Windows, feu clic a 'base' i tot seguit, a 'Download R 2.x.y for Windows' (*x* i *y* indiquen la versió actual de R)
5. Executeu l'arxiu des del directori on l'heu emmagatzemat i seguiu les instruccions d'instal·lació.

NOTA: D'aquesta forma s'instal·la la versió bàsica del R que conté els paquets bàsics. Si en algun moment necessiteu uns altres per dur a terme anàlisis més sofisticades....

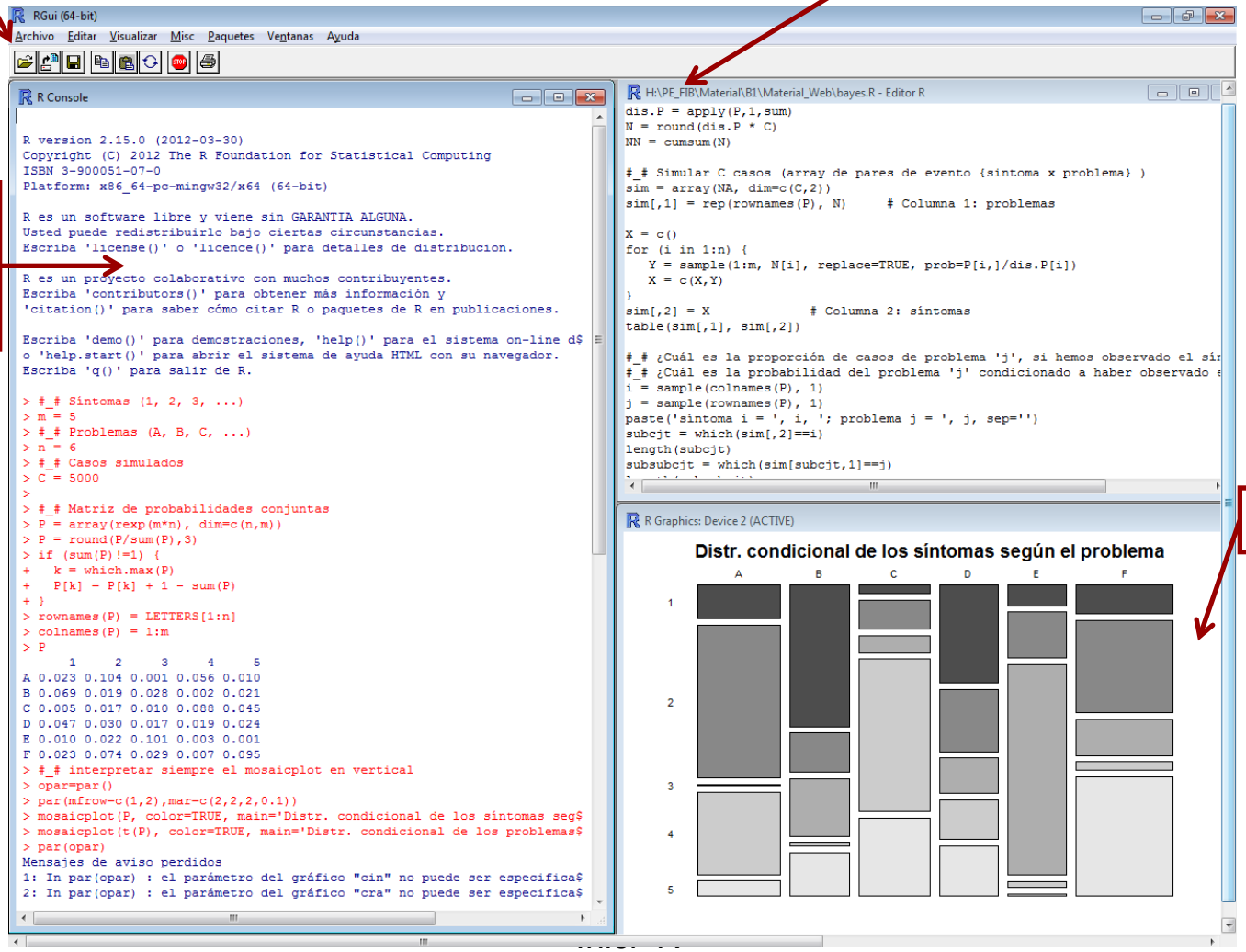
1. Obriu el R.
2. Feu clic en el menú 'Paquetes'
3. Escolliu 'Instalar paquete(s)'
4. Seleccioneu el servidor i el paquet desitjat

R – Interfície

Barra d'eines
Instal·lar paquets,
Guardar scripts, àrees de treball...

Script
Fitxer de text on guardem les instruccions que executem en la sessió

Consola
On s'executen les instruccions
(surt a l'obrir el R)



Finestra de gràfics

R - Funcions

- Les funcions (ó instruccions ó expressions) permeten realitzar diferents tipus d'operacions, anàlisis, gràfics, etc...
- *Sintaxi: nom de la funció (paràmetres o arguments):*

Com a funció:	Com a expressió matemàtica
> sum(2,3)	> 2+3
[1] 5	[1] 5

- Paràmetres: poden ser posicionals o per nom:

Posicionals (ordre per defecte):	Per nom (alterem l'ordre per defecte)
> plot(5,5,"h")	> plot(type="h", x=5, y=5)

- Output: els resultats poden ser:
  - a pantalla

```
> sum(2,3)
[1] 5
```

  - assignats ('=' ó '<-' ó '->') a un objecte.

```
> nota=sum(2,3)        > nota <- sum(2,3)        > nota -> sum(2,3)
```

  - efectes colaterals sobre fitxers, gràfics,…
- # indica comentari fins al final de la línia (no s'executa la línia de comanda)

R - Objectes

- Llenguatge orientat a objectes: les variables, dades, funcions, resultats, etc., es guarden en la memòria activa del computador en forma de objectes amb un nom específic.
- Nom:
  - Qualsevol cadena alfanumèrica (distingint minúscules de majúscules) llevat de les reservades (function, for, if, in, next, while, return, repeat, …)
  - Es poden usar noms compostos separats per punt, ja que el punt no és cap caràcter especial (Ex: nota.sum)
- Manipulació: L'usuari pot modificar o manipular aquests objectes amb operadors (aritmètics, lògics i comparatius) i amb funcions (que alhora són objectes).
- Tipus bàsics: numèrics (3) ; alfanumèrics("b") ; lògics (TRUE/FALSE) ...
- Classes:
  - Cada objecte té una classe concreta que permet emmagatzemar un determinat tipus de valors i fer unes operacions determinades.
  - Exemples de classes: vectors, factors, matrius, arrays, data.frames, llistes ....

R – Exemples (I)

```
# Demanar ajuda d'una funció o d'un tema general
help(mean)
?median
help.search ("histogram")

# Creació d'un vector numèric i calcular la seva mitjana
edat <- c(20,21,20,22,23,20,25,26,20,21)
edat.mitjana <- sum(edat)/length(edat)
edat.mean <- mean(edat)

# Creació d'un vector alfanumèric i fer la taula d'efectius
genere <- c("h","h","h","h","h","h","h","d","d","d")
table(genere)
summary(genere)

# Altres...
ls()            # Veure els objectes guardats en memòria
genere          # Visualitzar el contingut de l'objecte genere
length(edat)    # Longitud del vector edat
mode(genere)    # Tipus del vector genere
class(genere)   # Classe del vector genere
```

R – Exemples (II)

```
### LECTURA DE FITXERS

#Després de copiar les columnes del fitxer Notes.xls (disponible a la web)
notes <- read.delim("clipboard") # Guardem les dades a l'objecte notes
mean(notes$nota,na.rm=TRUE)      # Fem la mitjana de les notes
table(notes$sexe)                # Fem la taula del genere

#També podem passar les dades a un fitxer Notes.txt amb les variables
#separades per tabuladors o espais
setwd("E:/carpeta1/carpeta2")    # Defineixo el directori on tinc el fitxer
notes2 <- read.table("Notes.txt",header=TRUE) # Guardem a notes2
summary(notes2)                  # Fem descriptiva de les variables
```

R - Miscel·lània

- R distingeix majúscules i minúscules. Així, no és el mateix l'objecte *Edat* que l'objecte *edat*.
- Els valors mancants (*missings*) els codifica com *NA (Not Available).*
- La funció *summary* és una funció genèrica que és pot aplicar a un gran ventall de classes d'objectes i ens dóna un resum de la informació continguda en el mateix.
- Podem crear les pròpies funcions amb *function*

```
# Exemple de funció que escriu la mediana i el rang intercuartílic
med.IQR <- function(var) {
   med <- median(var)
   iq <- IQR(var)
   cat("Mediana:",med," IQR:",iq,"\n")
}
edat <- c(20,21,20,22,23,20,25,26,20,21)
med.IQR(edat)     # Executa la funció definida
```

- Per a més informació consultar el tutorial:

  http://www-eio.upc.edu/teaching/pe/B1/TutorialR.pdf