

Introducció

Els seminaris de l'assignatura d'Estadística es realitzaran utilitzant el llenguatge de programació R. R és un llenguatge de programació orientat a objectes, interpretat i amb una sintaxis semblant al C/C++. Aquest llenguatge forma part del sistema GNU i neix originalment de dos altres llenguatges, S i Scheme, als Bell Laboratories d' AT&T. A grans trets, R proporciona una gran varietat de tècniques estadístiques i gràfiques que estan incloses en nombroses llibreries que la comunitat actualitza contínuament. A més, R conté estructures de control del flux i la capacitat de definir instruccions addicionals a partir de la definició de noves funcions.

A continuació donarem uns primers detalls sobre el procés d'instal·lació i una primera familiarització amb el llenguatge i la seva línia de comandes.

Instal·lació

En general, la instal·lació depèn del sistema operatiu, aquí trobareu la gran majoria de fitxers d'instal·lació per qualsevol OS:

<http://cran.es.r-project.org/bin>

Windows: en el següent link trobareu l'executable necessari que a l'obrir-lo instal·larà el sistema base i un conjunt de paquets recomanats.

<http://cran.es.r-project.org/bin/windows/base>

Per comprovar que tot ha anat bé, feu doble click sobre la icona del R i se us hauria d'obrir una nova finestra amb accés a una línia de comandes de R preparada per treballar en una nova sessió (workspace).

macOS: trobareu els fitxers binaris i informació addicional a:

<https://cran.r-project.org/bin/macosx/>

Linux: en funció de la distribució en la que estigui basat el vostre sistema s'instal·larà d'una manera o una altra. En general, podeu baixar-vos les fonts de qualsevol versió (`R-x.y.z.tar.gz`) i compil·lar-les. Sinó, es pot instal·lar des dels fitxers binaris (`.deb` per Debian). Una altra opció, en el cas de les distribucions basades en Debian (Ubuntu), és utilitzar el software `apt-get`. Per exemple:

```
apt-get update  
apt-get install r-base r-base-dev
```

Podeu comprovar que tot ha anat bé obrint una terminal i executant la comanda R que us hauria d'obrir una línia de comandes en una nova sessió de R.

La primera sessió

En qualsevol dels sistemes operatius que utilitzeu, si teniu R instal·lat disposareu d'una interfície on fer servir la línia de comandes de R. Això és una consola que està preparada per rebre inputs (`>`) i bolcar els resultats corresponents (`[1]`). Alguns comentaris sobre les instruccions de R:

- S'executen amb la tecla **Intro**.
- Consisteixen en expressions i assignacions (definir funcions, variables, etc.).
- Dues instruccions en una mateixa línia s'han de separar fent servir `;`.
- Poden agrupar-se utilitzant els delimitadors `{ }`.
- Poden escriure's scripts per executar diferents comandes fent servir una sola instrucció.
- Els comentaris s'indiquen utilitzant `#`.

Com a primera comanda proveu que s'executa:

```
> 2+2
[1] 4
```

Alguns exemples més d'instruccions on s'inclouen algunes definicions de variables i algunes expressions matemàtiques:

```
> 2+2
[1] 4
> exp(3)
[1] 20.08554
> pi
[1] 3.141593
> x<-pi
> print(x)
[1] 3.141593
> 10+3^4+sqrt(3)
[1] 92.73205
> log(2)
[1] 0.6931472
> assign('a',10)
> a
[1] 10
```

A l'exemple que hem vist hem declarat dues variables amb els noms `a` i `x` que contenen 10 i π , respectivament. Quan declareu variables, tinguen en compte que R és *case-sensitive*, distingeix entre majúscules i minúscules i per tant, dues variables `a` i `A` no són la mateixa. Existeixen altres restriccions pels noms de les variables (veure documentació) que haureu de tenir en compte.

Per últim, en aquest exemple concret hem assignat a dues variables dos valors numèrics. Tanmateix, R permet assignar a variables qualsevol tipus d'objecte. Aquests poden ser caràcters, valors numèrics, valors lògics, factors, matrius, llistes, data frames, funcions, etc. Alguns d'ells els anirem veient al llarg de les següents sessions.

Per veure quines són les variables que teniu assignades en una sessió podeu fer servir la instrucció

```
> ls()
[1] "a" "x"
```

que retorna el nom de les variables declarades. Per eliminar-les de la sessió podeu fer servir `rm()`

```
> rm(a,x)
> rm (list=ls())
> ls()
```

Per últim, per executar un script que està guardat en un fitxer de text que s'anomena `script.R`, podeu utilitzar `source("script.R")`.

Ajuda dins de R

La quantitat de possibilitats que ofereix R és immensa i recordar totes les normes i arguments que rep cada mètode és impossible. Per sort, R conté manuals d'instruccions sobre les diferents comandes que són fàcilment accessibles des de la mateixa consola. Podeu fer servir qualsevol d'aquestes instruccions

```
> ?save
> help(save)
> help.start()
> ? help.search
> help.search("rnorm")
> ?"if"
> ?apropos
> apropos("plot")
> ?demo
> demo(graphics)
> ?example
> example("boxplot")
```

Paquets

Tal i com hem dit anteriorment, R conté multitud de paquets que tenen declarades diferents objectes com funcions o data sets que permeten realitzar tasques diverses. Aquests paquets poden venir instal·lats i carregats a la sessió per defecte, només instal·lats i preparats per ser carregats o bé cal instal·lar-los del repositori corresponent per posteriorment carregar-los a la sessió.

Un exemple de càrrega i instal·lació de paquets:

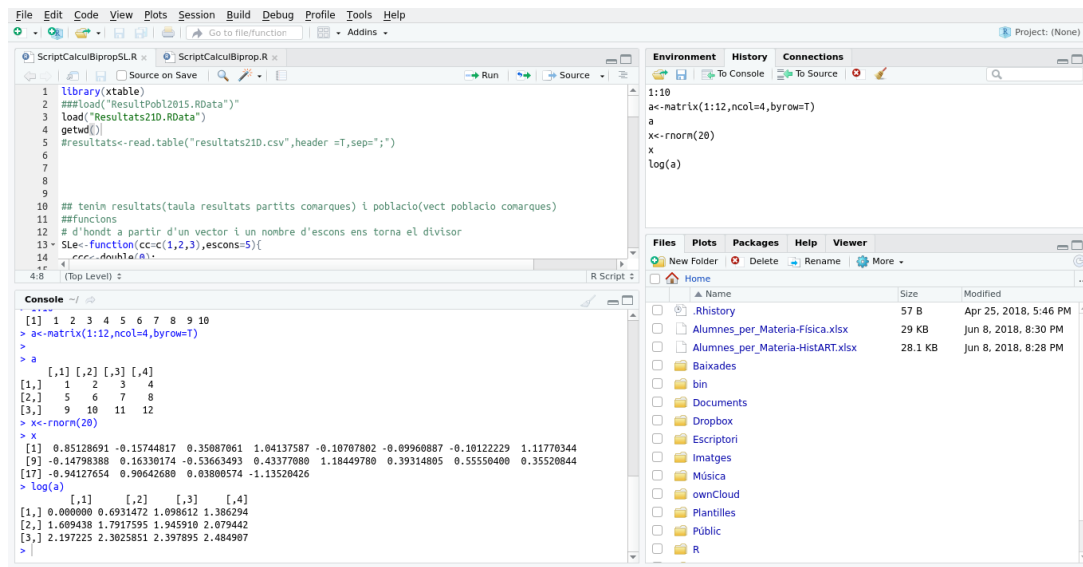
```
library(MASS) # Carrega el paquet MASS
help(package=MASS)
library(help=cluster) # Carrega l'ajuda del paquet cluster
search() # Llista els paquets carregats
install.packages("calendR") # Installa el paquet calendR
library("calendR") # Carrega el paquet
? whiteside
head(whiteside)
summary(whiteside)
names(whiteside )
dim(whiteside)
whiteside$Temp
plot(Gas ~ Temp, data=whiteside, pch =16)
```

Utilitzant un editor

Fins ara hem vist com utilitzar la línia de comandes de R. Tanmateix, durant les pràctiques es recomana utilitzar un editor de text que ens ajudi a escriure els scripts corresponents, així com també ens doni suport gràfic per visualitzar els nostres resultats. L'editor de text més conegut per treballar amb R s'anomena Rstudio i el teniu disponible a

<http://www.rstudio.com/>

A mode d'exemple:



A la figura observem al rectangle de la part superior esquerra el script que estem editant, que és el fitxer de text que conté el conjunt d'instruccions que volem executar. Rstudio proporciona diferents botons per executar una o totes les línies del fitxer ('Run', 'Source'). També podem executar qualsevol instrucció situant-nos sobre d'ella amb el cursor o bé seleccionant-les i apretant **Ctrl+Intro**. A la part inferior esquerra, disposem d'una consola de R com la que hem estat fent servir fins ara. La part de la dreta de la interfície es pot modificar segons les necessitats, però en general s'hi acostuma a posar una finestra amb la informació de totes les variables utilitzades i una altra on mostrar els gràfics que es generen.

Referencias

- [1] William N. Venables and David M. Smith. 2009. *An Introduction to R (2nd. ed.)*. Network Theory Ltd.