
Curso OpenCourseWare

**Aprendizaje del Software Estadístico R: un entorno
para simulación y computación estadística**

Alberto Muñoz García

1. Introducción al análisis de datos y al lenguaje S



Este es un curso práctico sobre análisis de datos, utilizando un paquete estadístico concreto, el R, como software de apoyo. El análisis de datos puede ser exploratorio o confirmatorio. Aquí trataremos de mostrar herramientas para ambos enfoques. La idea es introducir un lenguaje muy adaptado a las tareas estadísticas, y una vez dominados los aspectos básicos del lenguaje, introducir diversas herramientas del análisis de datos. El objetivo de este curso es que el alumno, al terminarlo, sea capaz de analizar a fondo sus datos con una herramienta que proporciona la más amplia variedad de técnicas, con el máximo grado de control sobre las mismas, que conocemos en el mercado.

¿Por qué R? R es un paquete estadístico de libre acceso basado en el lenguaje S, desarrollado en AT&T por Becker, Chambers y Wilks. S ha tenido desde su origen dos continuaciones paralelas: SPlus y R. SPlus es un paquete comercial, en activo desarrollo, y de extraordinaria calidad. R tiene facilidades similares, el apoyo de un magnífico equipo de profesionales, y además es gratuito. Cualquier manual o libro sobre S ó Splus sirve perfectamente para aprender R. Gran parte de los comandos de ambos lenguajes son iguales, o cuando menos, similares.

La filosofía de R se basa en el análisis interactivo de los datos a base de comandos. A diferencia de otros programas, no funciona por menús desplegables donde se van eligiendo las opciones del análisis que se desea realizar. Esto hace que el aprendizaje pueda resultar algo duro al principio, o por decirlo de otra manera, menos cómodo. A cambio, una vez acostumbrados, tendremos un control infinitamente mayor sobre nuestros análisis que el que tendríamos siguiendo el otro enfoque.

Para seguir el curso es necesario instalar previamente el paquete estadístico R. Para ello, podemos acudir a la siguiente dirección:

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

Se pueden encontrar páginas por todo el mundo de donde bajarse el programa.

Para instalar R en nuestra máquina:

1. Entramos en una página de CRAN, por ejemplo en <http://cran.r-project.org/>
2. Elegimos en la parte izquierda de la pantalla la opción R Binaries, en el apartado de Software.
3. Una vez hecho esto, aparecen en la parte derecha una serie de carpetas con nombres de sistemas operativos. Elegimos Windows.
4. Una vez ahí aparece un fichero ReadMe (conviene leerlo) y nuevas carpetas. Entraremos en la carpeta base.
5. Una vez ahí dentro, nos bajaremos el fichero SetupR.exe. Al ejecutarlo en nuestra máquina, se instalará el programa R, así como una serie de librerías que amplían el número de funciones disponibles.

Si alguien utiliza el sistema operativo LiNuX (es mi caso), deberá elegir la carpeta Linux, y proceder en similar manera.

Una vez instalado el programa, podremos comenzar a trabajar.