Estadística Aplicada al Medio Ambiente

Sesión 02. El paquete R

10 de octubre de 2023

Listado de los contenidos de la presentación.

Descarga del paquete R

El paquete R en la materia

Algunas indicaciones para realizar la PEC

Carga de los datos de la PEC en R

Representación gráfica de los datos

Cálculo de medidas de forma y dispersión

Descarga del paquete R

El paquete R se puede descargar desde la siguiente url:

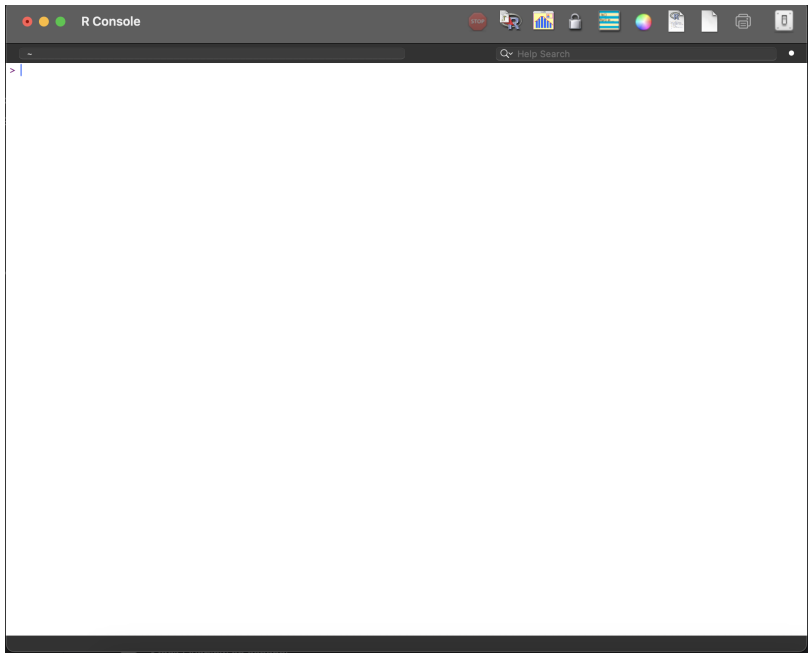
`https://www.r-project.org`

Más concretamente, desde la url:

`https://cran.rediris.es`

Seleccionados el sistema operativo y descargamos e instalamos el paquete.

Una vez descargado, al ejecutarlo, tendría que aparecer una ventana como la siguiente en donde podemos ver el símbolo $>$



El paquete R en la materia

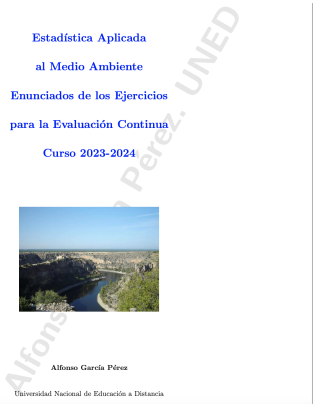
Emplearemos el paquete R como:

- Herramienta básica para poner en práctica los contenidos teóricos de la materia.

  En el primer capítulo del libro tenemos una introducción básica de R.

- Para la resolución de la PEC **(Nos centraremos en la PEC. El examen presencial se hace sin ayuda de R)**.

  Ya está disponible en el Curso Virtual **(Fecha límite 10 de enero)**


Estadística Aplicada

al Medio Ambiente

Enunciados de los Ejercicios

para la Evaluación Continua

Curso 2023-2024

Alfonso García Pérez. UNED

Alfonso García Pérez

Universidad Nacional de Educación a Distancia

Algunas indicaciones para realizar la PEC

- Es recomendable leer el primer capítulo del libro, sobre todo en lo que respecta a los tipos de datos, representaciones gráficas y funciones.
- En la práctica aparecen algunos conceptos que iremos viendo a lo largo del curso y sobre los que iremos trabajando.
- En lo que respecta al tipo de datos que emplearemos en el primer ejercicio, nos centraremos en la estructuras de datos o data frames. Estas estructuras permiten trabajar con diferentes tipos de datos a la vez (numéricos, caracteres, etc.).
- Para crear estructuras de datos, podemos emplear ficheros en los que tengamos almacenados los datos que necesitamos.
- Por ejemplo, en el primer ejercicio de la PEC, tenemos los datos en el fichero **AU.txt** que podemos encontrar en el Curso Virtual y, en el segundo, tenemos los datos en el fichero **Precipitaciones.txt**.
- Lo que podemos hacer hoy es cargar dichos ficheros en R y hacer algunas manipulaciones básicas.

Carga de los datos de la PEC en R

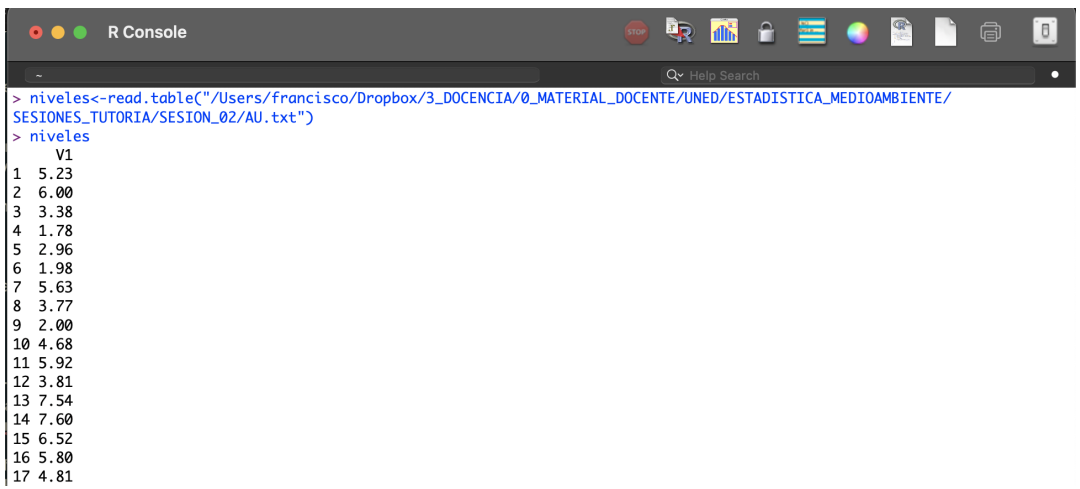
Comencemos cargando el primer fichero de datos de la PEC en R. Para ello, tecleamos la siguiente instrucción (OJO, lo que está entre las comillas es la ruta en donde tienes almacenado el fichero, la sentencia tiene que estar en una sola línea, la he partido para que entrase en la presentación)

```
> niveles<-read.table("/Users/francisco/Dropbox/3_DOCENCIA/
0_MATERIAL_DOCENTE/UNED/ESTADISTICA_MEDIOAMBIENTE/
SESIONES_TUTORIA/SESION_02/AU.txt")
```

Si ahora tecleamos

```
>niveles
```

Obtendremos lo siguiente:



Observamos que tenemos dos columnas, la primera con la numeración del individuo y la segunda con el nivel de ácido úrico en sangre.

Carga de los datos de la PEC en R

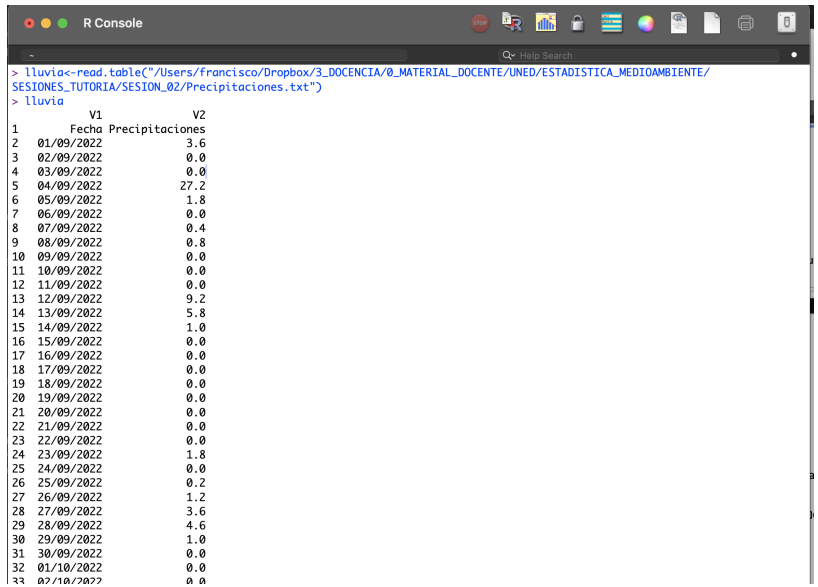
Veamos ahora como podemos cargar el segundo fichero de datos de la PEC en R. Para ello, tecleamos lo siguiente:

```
> lluvia<-read.table("/Users/francisco/Dropbox/3_DOCENCIA/0_MATERIAL_DOCENTE/
UNED/ESTADISTICA_MEDIOAMBIENTE/SESIONES_TUTORIA/
SESION_02/Precipitaciones.txt")
```

Si ahora tecleamos

```
>lluvia
```

Obtenemos lo siguiente:



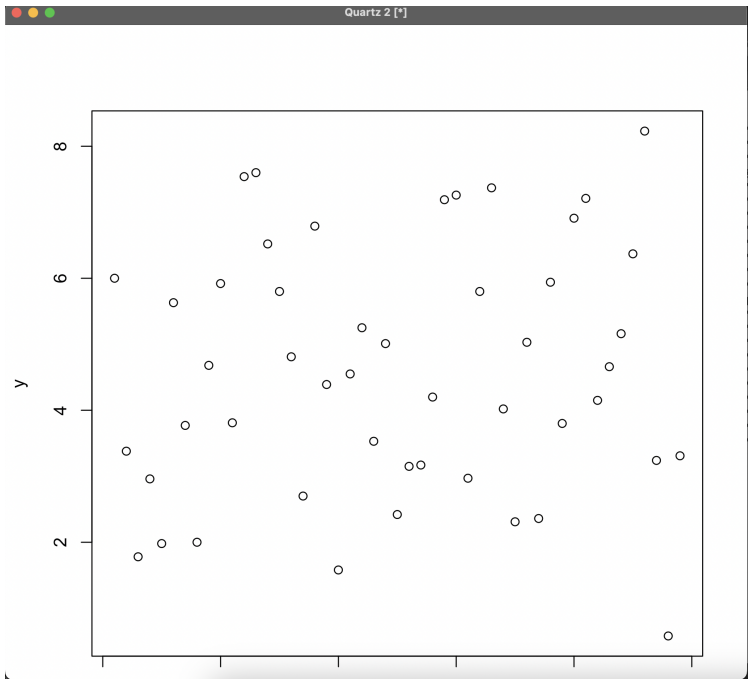
Observamos que ahora la estructura de datos es mas rica pues, además de la cantidad de lluvia diaria, nos aparece la fecha en la que se ha realizado la medición.

Representación gráfica de los datos

En el primer problema nos piden que representemos gráficamente los datos, para ello, podemos emplear las siguientes instrucciones de R:

```
> x<-seq(1,nrow(niveles)-1,len=nrow(niveles)-1)
> y<-niveles$V1[2:nrow(niveles)]
> plot(x,y)
```

Obteniendo el siguiente diagrama de dispersión:



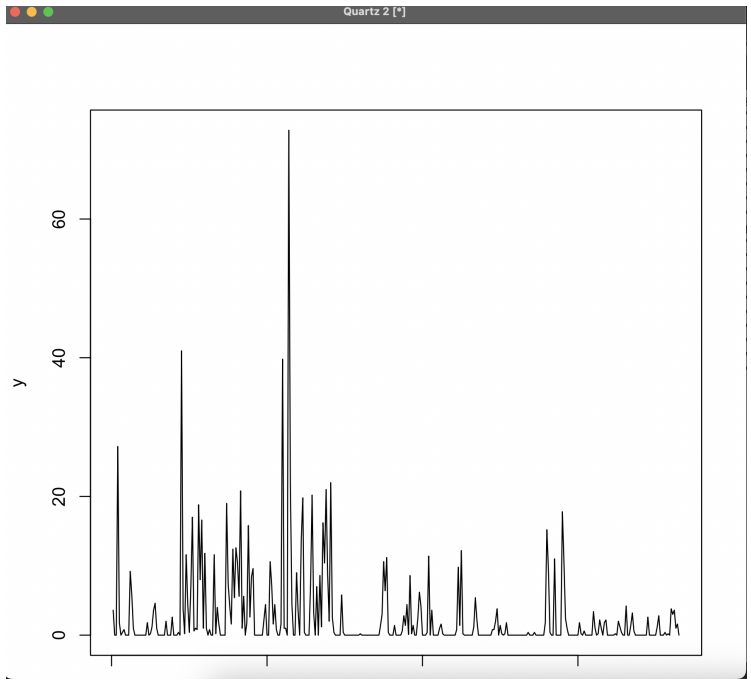
Puede ser interesante realizar algún otro tipo de representación gráfica como, por ejemplo, un histograma agrupando los datos por intervalos.

Representación gráfica de los datos

Aunque en el segundo ejercicio no nos pidan la representación gráfica de los datos, puede ser interesante representarlo. En este caso tecleamos las siguientes instrucciones:

```
> x<-seq(1,nrow(lluvia)-1,len=nrow(lluvia)-1)
> y<-lluvia$V2[2:nrow(lluvia)]
> plot(x,y)
> plot(x,y,type="l")
```

Obtenemos la siguiente representación gráfica:



Cálculo de medidas de forma

En el primer ejercicio nos piden que calculemos algunas medidas de forma y dispersión de los datos. En particular:

- Media.

```
> y<-niveles$V1[2:nrow(niveles)]
> mean(y)
[1] 4.587551
```

- Mediana.

```
> median(y)
[1] 4.55
```

- Moda.

```
# Definimos la función moda
moda <- function(x) {return(as.numeric(names(which.max(table(x)))))}
# Calculamos la moda empleando la función anterior
> moda(y)
[1] 5.8
```

- Percentiles 10 y 90:

```
> quantile(y,probs=0.1)
  10%
2.248
> quantile(y,probs=0.9)
 90%
7.22
```