### **Ejercicio Propuesto 2 (Resuelto)**

**En Andalucía se quiere contrastar la independencia global entre las variables sexo y utilización de instalaciones deportivas de los municipios de la Comunidad. Para ello, se toma una muestra de 854 individuos cuyas observaciones figuran en la siguiente tabla aqui**



[Tabla 11. Datos del Ejercicio propuesto 2(.docx)](https://wpd.ugr.es/~bioestad/wp-content/uploads/Propuesto2.docx)

**Considerando un nivel de significación del 5%, estudia si las dos variables si las dos variables están relacionadas o si, por el contrario, son independientes.**

#### ****Solución****

En primer lugar, para trabajar con *R\_Commander* escribimos la siguiente sentencia en *R*

> library(Rcmdr)

Introducimos los datos en *R-Commander*. Para ello, tenemos que crear un fichero de texto como el que aparece en la Imagen



Como puede verse, en la primera fila introducimos el nombre de las variables entre comillas y separados por un espacio. A continuación, en las siguientes filas se van introduciendo los datos que aparecen en la tabla. Hay que tener en cuenta que las modalidades de las variables cualitativas hay que escribirlas entre comillas.

A continuación, cargamos el fichero seleccionando:

***Datos🡪Importar datos🡪desde archivo de texto, portapapeles o URL…***



Se muestra la siguiente ventana, en la cual vamos a introducir el nombre que queremos asignarle al conjunto de datos con el que vamos a trabajar; en nuestro caso, escribiremos Ejemplo\_Instalaciones. El resto de opciones las dejamos por defecto, ya que el archivo de texto que hemos creado cumple con todas ellas.



Una vez le damos a *Aceptar* se abre una ventana para que seleccionemos el archivo de texto que hemos creado y guardado anteriormente en nuestro ordenador. Cuando abrimos el archivo, podemos ver que en *Conjunto de datos* aparece el nombre que le hemos asignado a nuestro conjunto de datos.



Para transformar la tabla de frecuencias en un conjunto de datos (data.frame) con elque R-Commander pueda trabajar hay que escribir las siguientes instrucciones en la ventana *R Script*, seleccionar las tres a la vez y darle a *Ejecutar*:

P<-rep(Ejemplo\_Instalaciones$Sexo,Ejemplo\_Instalaciones$Frecuencia)

Q<-rep(Ejemplo\_Instalaciones$Usuario,Ejemplo\_Instalaciones$Frecuencia)

Sexo\_Usuario<-data.frame(P,Q)

El nuevo conjunto de datos creado en forma de lista lo hemos llamado *Sexo\_Usuario*, y es el que tenemos que seleccionar en la pestaña de *Conjunto de datos*.



El contraste de hipótesis que se debe resolver es:

**H0 ≡ El sexo y se usuario de instalaciones deportivases independiente**

**H1 ≡ El sexo y se usuario de instalaciones deportivas noes independiente**

Para ello, en el menú seleccionamos

**Estadísticos🡪Tablas de Contingencia🡪Tabla de doble entrada…**



En la pestaña *Datos* tenemos que seleccionar las variables que queremos que aparezcan tanto por filas (en nuestro caso, la variable P a la que hemos asignado el sexo) como por columnas (seleccionamos la variable Q a la que hemos asignado si es usuario o no).



En la pestaña *Estadísticos* ya aparece seleccionado el *Test de independencia Chi-cuadrado*, que es el que nos interesa. Por tanto, no modificamos ninguna de las opciones



Pulsamos *Aceptar* y se muestra la siguiente salida



El p-valor asociado a este contraste es **0.003258**. Como este p-valor es menor que 0.05, se rechaza la hipótesis nula del contraste, por lo que concluimos que el sexo y la utilización de instalaciones deportivas son variables dependientes. Esto es, los valores de una dependen de los valores de la otra.