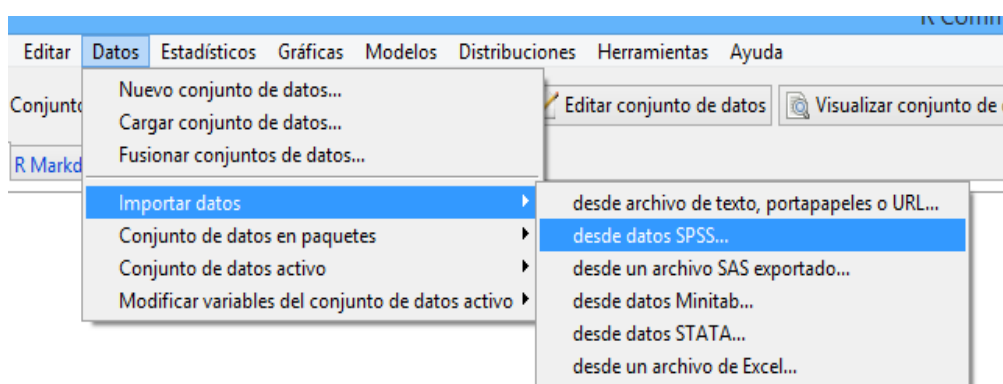


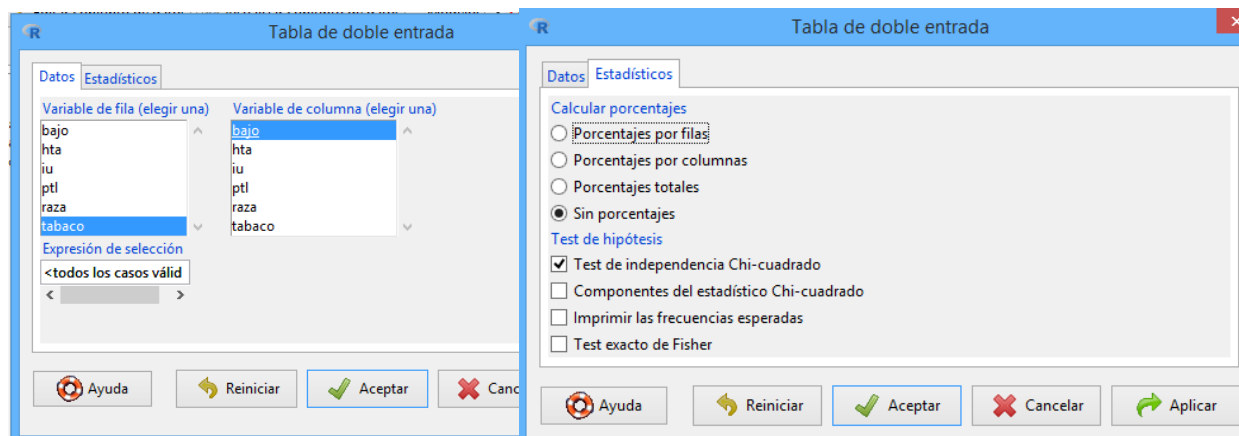
## CASO PRÁCTICO: APLICACIONES DE LA CHI-CUADRADO

En este ejercicio vamos a usar una base de datos sobre mujeres embarazadas donde se recogen las siguientes variables: edad de dichas mujeres, etnia, si fumaban o no, así como si el neonato al nacer era de peso normal o de bajo peso. Todos estos datos están recogidos en el archivo "Embarazo.sav". Hay que observar que el archivo es ".sav" este es el formato característico de los archivos de datos del programa SPSS, y por tanto tendremos que cargarlo en R como un archivo de SPSS.

Para ello en Rcmdr seleccionaremos **Datos Importar Datos → Desde datos SPSS...** damos nombre a nuestro archivo (embarazo) y nos aparecerá el explorador, buscamos nuestro archivo "Embarazo" y aceptamos la carga del archivo.



Comenzaremos con un análisis bivalente, en el cual, queremos conocer si con los datos que tenemos podemos afirmar que el tabaco está asociado con el bajo peso de los recién nacidos, por lo que consideraremos que hay significación estadística cuando nuestra  $p < 0,05$ . Para ello, una vez que hemos cargado el archivo seleccionamos **Estadísticos → Tablas de contingencia → Tabla de doble entrada...** y nos aparece una ventana con dos pestañas que ya nos son familiares "Datos" en la cual debemos seleccionar en la columna de la izquierda la variable independiente que en este caso es el tabaco, en la columna de la derecha seleccionamos el peso de los neonatos. Si nos movemos a la pestaña de "Estadísticos" nos permite elegir qué tipo de test vamos a realizar con los datos



seleccionados en el apartado anterior, así como que nos permite diferentes formas de ver los datos de la tabla de 2x2, ya sea en porcentajes, en valores medidos etc.... Seleccionamos la opción "Sin porcentajes" para que los datos sean en valores obtenidos en el estudio. Y por defecto nos aparece una casilla con un recuadro que pone "Test de independencia Chi-cuadrado", de esta manera el programa sabe que a las variables que le hemos seleccionado nos debe crear una tabla de 2x2 con los valores de nuestra base de datos y le debe realizar un test de Chi-cuadrado. Al completar todos estos pasos en la consola de R nos debe aparecer el siguiente cuadro:

```

Frequency table:
      bajo
tabaco  Peso normal Bajo peso
  NO Fuma      86      29
   Fuma       44      30

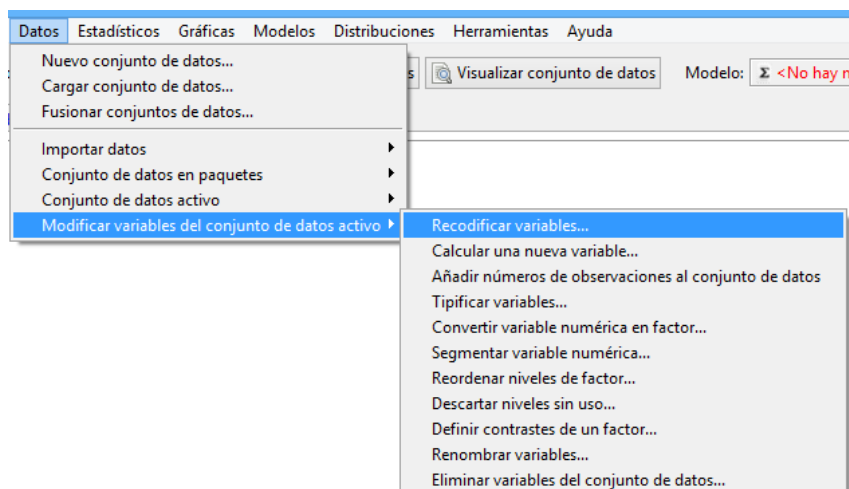
      Pearson's Chi-squared test

data: .Table
X-squared = 4.9237, df = 1, p-value = 0.02649

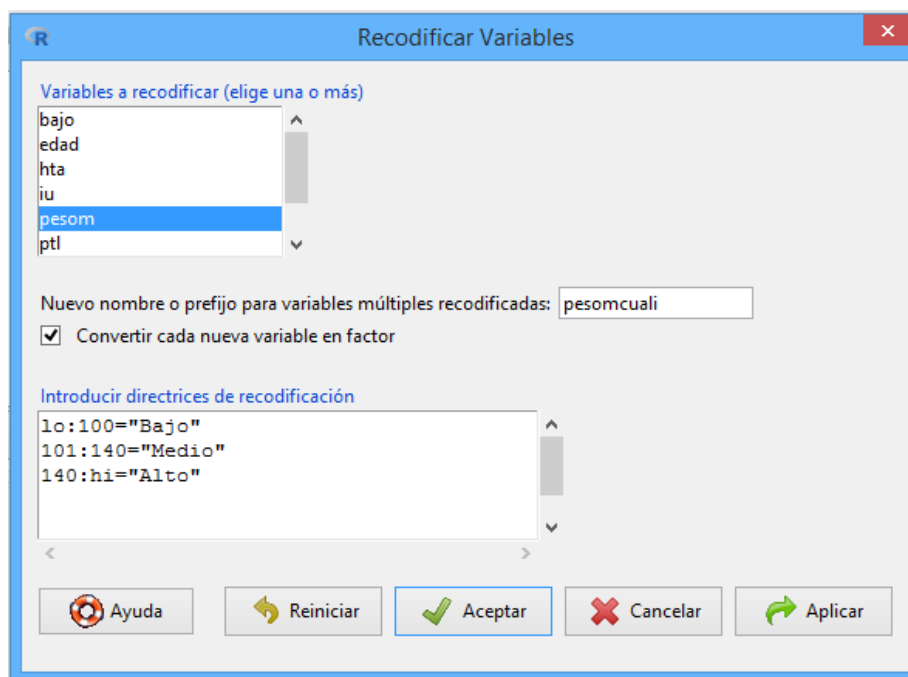
```

Donde se puede observar la tabla que enfrenta el consumo de tabaco con el peso de los recién nacidos, y los resultados del test Chi-cuadrado, el cual obtiene un valor en este caso de 4,9237, a este valor con un grado de libertad (df=1) le corresponde una probabilidad de 0,026, esta probabilidad es menor que nuestra probabilidad límite que hemos considerado de 0,05. Luego se podría concluir que hay significación estadística con los datos que tenemos en cuanto a la relación entre los niños que nacen con bajo peso al nacer y las madres que fuman durante el embarazo.

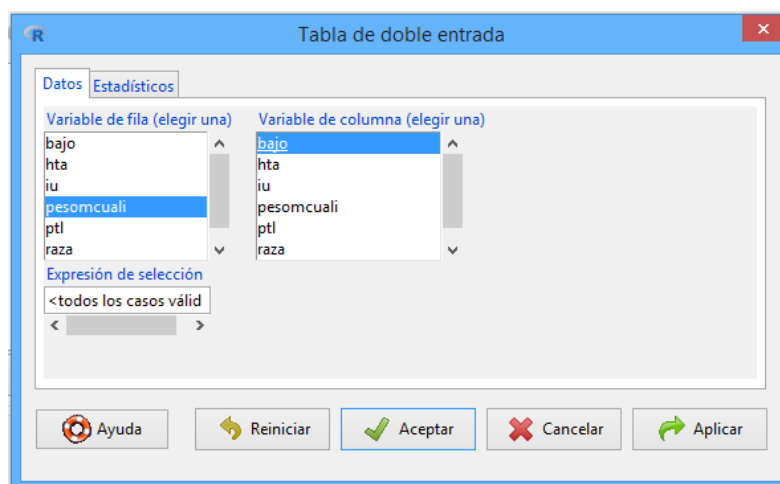
Para el siguiente ejercicio primero debemos realizar una reconversión de una variable dentro de nuestra base de datos, queremos comparar si con los datos que tenemos podemos afirmar que haya significación estadística entre el peso de la madre y el peso de recién nacido mediante un contraste de hipótesis Chi-cuadrado, pero para ello tenemos un problema previo que solucionar, ya que la variable "Peso" en nuestra base de datos esta codificada como una variable numérica, no como una variable cualitativa. De ahí que sea necesario recodificar la variable, y lo haremos de la siguiente manera:



Una vez hayamos cliqueado esto nos aparecerá la siguiente ventana, en la cual tenemos que configurar la nueva recodificación de la variable, y lo haremos de la siguiente manera:



Seleccionamos la variable que queremos recodificar (en este caso "pesom") y nombramos la nueva variable que vamos a crear (en este caso "pesomcuali"), para recodificarla debemos escribir exactamente lo mismo que aparece en el ejemplo, ya que R es una consola de programación, y si los comandos no se escriben exactamente probablemente tendremos algún problema. En este sistema de codificación le estamos diciendo a R que los pesos desde el más bajo (low) hasta 100 libras sea "Bajo" en nuestra nueva variable, entre 101 y 140 libras que sea un peso medio y de 140 libras al peso más alto (high) que las consideremos como pesos altos. De esta manera aparece una nueva tabla de datos con nuestra variable recodificada, la cual vamos a usar a continuación, repitiendo el procedimiento anterior. [Estadísticos](#) → [Tablas de contingencia](#) → [Tabla de doble entrada...](#) y ahora seleccionamos las dos variables que vamos a comparar.



Y obtenemos los siguientes resultados:

```
Frequency table:
      bajo
pesomcuali  Peso normal  Bajo peso
Alto          35          12
Bajo          13          11
Medio         82          36

      Pearson's Chi-squared test

data:  .Table
X-squared = 3.1229, df = 2, p-value = 0.2098
```

Donde podemos observar que para el valor de Chi-cuadrado con dos grados de libertad el valor de  $p=0,209$ , si consideramos nuestro punto de corte en  $0,05$ , podemos asegurar que el peso de la madre no ha influido en el peso de los recién nacidos de acuerdo con la recodificación que hemos realizado.

---