

PRÁCTICAS DE ESTADÍSTICA CON SPSS

Autores:

Santiago Angulo Díaz-Parreño (sangulo@ceu.es)

José Miguel Cárdenas Rebollo (cardenas@ceu.es)

Anselmo Romero Limón (arlimon@ceu.es)

Alfredo Sánchez Alberca (asalber@ceu.es)

Curso 2012-2013



CEU

*Universidad
San Pablo*

Prácticas de Estadística con PASW

Santiago Angulo Díaz-Parreño, José Miguel Cárdenas Rebollo, Anselmo Romero Limón y Alfredo Sánchez Alberca (asalber@gmail.com).



Esta obra está bajo una licencia Reconocimiento-No comercial-Compartir bajo la misma licencia 2.5 España de Creative Commons. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/byncsa/2.5/es/> o envíe una carta a Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

Con esta licencia eres libre de:

- Copiar, distribuir y mostrar este trabajo.
- Realizar modificaciones de este trabajo.

Bajo las siguientes condiciones:



Reconocimiento. Debe reconocer los créditos de la obra de la manera especificada por el autor o el licenciador (pero no de una manera que sugiera que tiene su apoyo o apoyan el uso que hace de su obra).



No comercial. No puede utilizar esta obra para fines comerciales.



Compartir bajo la misma licencia. Si altera o transforma esta obra, o genera una obra derivada, sólo puede distribuir la obra generada bajo una licencia idéntica a ésta.

- Al reutilizar o distribuir la obra, tiene que dejar bien claro los términos de la licencia de esta obra.
 - alguna de estas condiciones puede no aplicarse si se obtiene el permiso del titular de los derechos de autor
 - Nada en esta licencia menoscaba o restringe los derechos morales del autor.
-

Índice general

| | |
|---|-----------|
| 1. Introducción a SPSS | 1 |
| 1.1. Introducción | 1 |
| 1.2. Funciones básicas | 2 |
| 1.2.1. Arranque | 2 |
| 1.2.2. Introducción de datos | 2 |
| 1.2.3. Guardar datos | 4 |
| 1.2.4. Recuperar datos | 4 |
| 1.2.5. Modificación de datos | 4 |
| 1.2.6. Transformación y generación de datos | 5 |
| 1.2.7. Recodificación de datos | 6 |
| 1.2.8. Impresión | 6 |
| 1.2.9. Salir del programa | 7 |
| 1.2.10. Ayuda | 7 |
| 1.3. Ejercicios resueltos | 8 |
| 1.4. Caso práctico | 10 |
| 2. Distribuciones de Frecuencias y Representaciones Gráficas | 13 |
| 2.1. Fundamentos teóricos | 13 |
| 2.1.1. Cálculo de Frecuencias | 13 |
| 2.1.2. Representaciones Gráficas | 14 |
| 2.2. Ejercicios resueltos | 19 |
| 2.3. Ejercicios propuestos | 20 |
| 3. Estadísticos Muestrales | 22 |
| 3.1. Fundamentos teóricos | 22 |
| 3.1.1. Medidas de posición | 22 |
| 3.1.2. Medidas de dispersión | 23 |
| 3.1.3. Medidas de forma | 24 |
| 3.1.4. Estadísticos de variables en las que se definen grupos | 25 |
| 3.2. Ejercicios resueltos | 26 |
| 3.3. Ejercicios propuestos | 29 |
| 4. Regresión Lineal Simple y Correlación | 32 |
| 4.1. Fundamentos teóricos | 32 |
| 4.1.1. Regresión | 32 |
| 4.1.2. Correlación | 35 |
| 4.2. Ejercicios resueltos | 39 |
| 4.3. Ejercicios propuestos | 41 |
| 5. Regresión No Lineal | 43 |
| 5.1. Fundamentos teóricos | 43 |
| 5.2. Ejercicios resueltos | 45 |
| 5.3. Ejercicios propuestos | 46 |

| | |
|--|------------|
| 6. Variables Aleatorias Discretas | 48 |
| 6.1. Fundamentos teóricos | 48 |
| 6.1.1. Variables Aleatorias | 48 |
| 6.1.2. Variables Aleatorias Discretas (v.a.d.) | 48 |
| 6.2. Ejercicios resueltos | 52 |
| 6.3. Ejercicios propuestos | 55 |
| 7. Variables Aleatorias Continuas | 57 |
| 7.1. Fundamentos teóricos | 57 |
| 7.1.1. Variables Aleatorias | 57 |
| 7.1.2. Variables Aleatorias Continuas (v.a.c.) | 57 |
| 7.2. Ejercicios resueltos | 63 |
| 7.3. Ejercicios propuestos | 68 |
| 8. Intervalos de Confianza para Medias y Proporciones | 69 |
| 8.1. Fundamentos teóricos | 69 |
| 8.1.1. Inferencia Estadística y Estimación de Parámetros | 69 |
| 8.1.2. Intervalos de Confianza | 69 |
| 8.2. Ejercicios resueltos | 74 |
| 8.3. Ejercicios propuestos | 76 |
| 9. Intervalos de Confianza para Comparación de Poblaciones | 77 |
| 9.1. Fundamentos teóricos | 77 |
| 9.1.1. Inferencia Estadística y Estimación de Parámetros | 77 |
| 9.1.2. Intervalos de Confianza | 77 |
| 9.2. Ejercicios resueltos | 82 |
| 9.3. Ejercicios propuestos | 84 |
| 10. Contraste de Hipótesis | 86 |
| 10.1. Fundamentos teóricos | 86 |
| 10.1.1. Inferencia Estadística y Contrastes de Hipótesis | 86 |
| 10.1.2. Tipos de Contrastes de Hipótesis | 86 |
| 10.1.3. Elementos de un Contraste | 86 |
| 10.2. Ejercicios resueltos | 93 |
| 10.3. Ejercicios propuestos | 95 |
| 11. Análisis de la Varianza de 1 Factor | 96 |
| 11.1. Fundamentos teóricos | 96 |
| 11.1.1. Notación, Modelo y Contraste | 96 |
| 11.2. Ejercicios resueltos | 100 |
| 11.3. Ejercicios propuestos | 102 |
| 12. Contrastes Basados en el Estadístico χ^2. Comparación de Proporciones | 103 |
| 12.1. Fundamentos teóricos | 103 |
| 12.1.1. Contraste χ^2 de Pearson para ajuste de distribuciones | 104 |
| 12.1.2. Contraste χ^2 en tablas de contingencia | 104 |
| 12.1.3. Test Exacto de Fisher | 105 |
| 12.1.4. Test de McNemar para datos emparejados | 105 |
| 12.2. Ejercicios Resueltos | 106 |
| 12.3. Ejercicios propuestos | 109 |
| 13. Contrastes de Hipótesis No Paramétricos | 111 |
| 13.1. Fundamentos teóricos | 111 |
| 13.1.1. Contrastes no paramétricos más habituales | 112 |
| 13.1.2. Aleatoriedad de una muestra: Test de Rachas | 113 |
| 13.1.3. Pruebas de Normalidad | 114 |
| 13.1.4. Test de la U de Mann-Whitney | 116 |

| | |
|---|------------|
| 13.1.5. Test de Wilcoxon para datos emparejados | 117 |
| 13.1.6. Test de Kruskal-Wallis: comparación no paramétrica de k medias independientes | 118 |
| 13.1.7. Test de Friedman: equivalente no paramétrico del ANOVA con medidas repetidas | 119 |
| 13.1.8. Test de Levene para el contraste de homogeneidad de varianzas | 120 |
| 13.1.9. El coeficiente de correlación de Spearman | 121 |
| 13.2. Ejercicios resueltos | 121 |
| 13.3. Ejercicios propuestos | 126 |
| 14. Análisis de Concordancia | 128 |
| 14.1. Fundamentos teóricos | 128 |
| 14.1.1. Introducción | 128 |
| 14.1.2. Análisis de la Concordancia entre dos Variables Cuantitativas | 128 |
| 14.1.3. Análisis de la Concordancia entre dos Variables Cualitativas | 129 |
| 14.2. Ejercicios resueltos | 130 |
| 14.3. Ejercicios propuestos | 133 |
| 15. ANOVA de Dos o Más Factores, ANOVA de Medidas Repetidas y ANCOVA | 134 |
| 15.1. Fundamentos teóricos | 134 |
| 15.1.1. ANOVA de dos o más factores | 135 |
| 15.1.2. ANOVA de medidas repetidas | 141 |
| 15.1.3. ANOVA de medidas repetidas + ANOVA de una o más vías | 143 |
| 15.1.4. Análisis de Covarianza: ANCOVA | 144 |
| 15.2. Ejercicios resueltos | 145 |
| 15.3. Ejercicios propuestos | 154 |

Introducción a SPSS

1 Introducción

La gran potencia de cálculo alcanzada por los ordenadores ha convertido a los mismos en poderosas herramientas al servicio de todas aquellas disciplinas que, como la estadística, requieren manejar un gran volumen de datos. Actualmente, prácticamente nadie se plantea hacer un estudio estadístico serio sin la ayuda de un buen programa de análisis estadístico.

SPSS®* es uno de los programas de análisis estadísticos más utilizados, sobre todo en el ámbito de las ciencias biosanitarias.



El objetivo de esta práctica es introducir al alumno en la utilización de este programa, enseñándole a realizar las operaciones básicas más habituales. A lo largo de la práctica, los alumnos aprenderán a crear variables, introducir datos de las muestras, transformar variables, filtrar datos y fundir e importar archivos de datos.

*Esta practica está basada en la versión 20.0 de SPSS® para Windows en español.

2 Funciones básicas

2.1 Arranque

Como cualquier otra aplicación de Windows, para arrancar el programa hay que hacer click sobre la opción correspondiente del menú **Inicio->Programas**, o bien sobre el icono de escritorio



Cuando el programa arranca, aparece la ventana del editor de datos (figura 1.1).

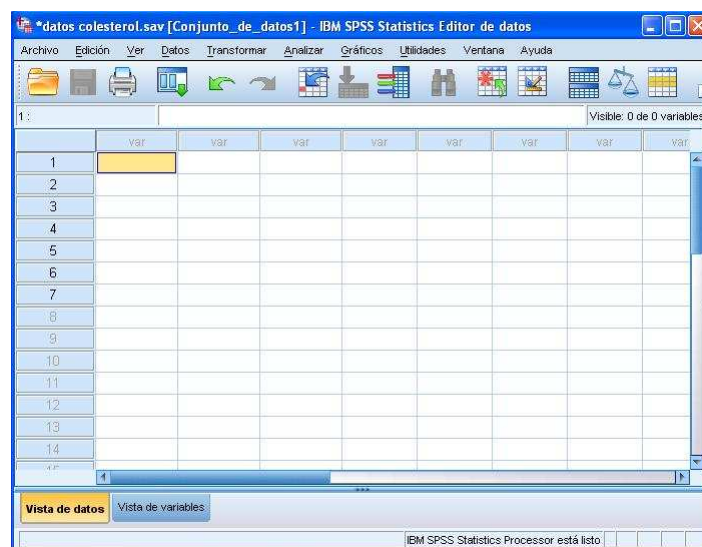


Figura 1.1 – Ventana del editor de datos.

Como cualquier otra ventana de aplicación de Windows, la ventana del editor de datos tiene una barra de título, una barra de menús con las distintas funciones que puede hacer SPSS, entre ellas los análisis estadísticos de datos, una barra de botones que son atajos a las opciones más habituales de los menús, y una barra de estado en la parte inferior que nos indica lo que hace el programa en cada instante. Además, en la parte inferior aparecen dos pestañas que permiten pasar a la **Vista de datos** o a la **Vista de variables**.

2.2 Introducción de datos

Para realizar cualquier análisis, la ventana del editor de datos debe contener la matriz de datos a analizar. Una vez que el usuario obtiene los datos muestrales, estos deben introducirse en esta ventana. Para ello, lo primero es definir las variables que se han considerado en el estudio. Cada variable se corresponderá con una columna de la matriz de datos.

Para definir una variable debemos pasar a la **Vista de variables** haciendo click sobre la correspondiente pestaña (figura 1.2).

En esta otra ventana, debemos definir cada variable en una fila, rellenando los siguientes campos:

Nombre El nombre de la variable puede ser cualquier cadena de caracteres que comience por una letra y que no contenga espacios en blanco ni caracteres especiales como $?$, $;$, $*$, etc. Cada nombre de variable debe ser único y no se distingue entre mayúsculas y minúsculas.

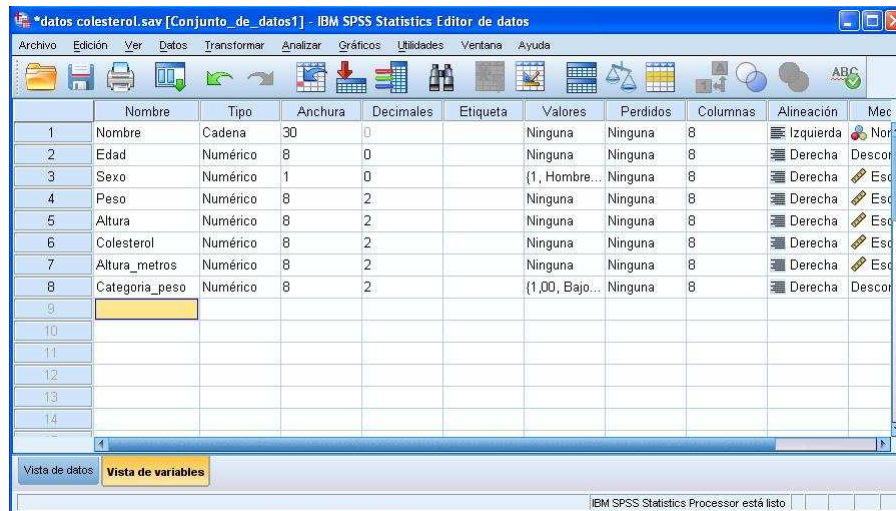


Figura 1.2 – Vista de definición de variables.

Tipo Los tipos más comunes son **Numérico** (formato numérico estándar), **Coma** (con comas de separación cada tres cifras y punto para la parte decimal), **Punto** (con puntos de separación cada tres cifras y coma para la parte decimal), **Notación Científica** (utiliza la E para la exponenciación), **Cadena** (para datos alfanuméricos) y **Fecha**.

Anchura Es el número máximo de caracteres que pueden tener los valores de la variable.

Decimales Para las variables numéricas es el número de cifras decimales que podrán escribirse.

Etiqueta Es una descripción de la variable. Si el nombre de la variable es suficientemente descriptivo se puede omitir.

Valores Permite asignar etiquetas a los distintos valores que puede tomar la variable. No es obligatorio pero puede ser útil en algunos casos.

Al hacer click sobre la casilla aparece un cuadro de diálogo para asignar etiquetas a valores. Para ello basta con escribir un valor en el cuadro de texto **Valor** y la correspondiente etiqueta en el cuadro de texto **Etiqueta**. Después hay que hacer click sobre el botón **Añadir** y repetir los mismos pasos para todos los valores de la variable. Para finalizar hay que hacer click en el botón **Aceptar**.

Perdidos Permite definir qué valores se utilizarán para representar los datos perdidos por el usuario. Es útil para distinguir datos que se han perdido por distintas causas. Por ejemplo, puede ser interesante distinguir el dato perdido correspondiente a un entrevistado que se niega a responder, del dato perdido debido a que la pregunta no le afectaba al entrevistado. Los valores de datos especificados como perdidos por el usuario se excluyen de la mayoría de los cálculos.

Al hacer click sobre la casilla aparece un cuadro de diálogo donde deben indicarse los valores discretos que representarán valores perdidos (pueden introducirse hasta tres), o bien, el rango de valores que se representarán como valores perdidos.

Columnas Permite especificar el ancho de la columna en la que se introducirán los datos correspondientes a la variable.

Alineación Permite especificar la alineación de los datos correspondientes a la variable. Puede ser Izquierda, Derecha o Centrado.

Medida Permite especificar el tipo de escala utilizada para medir la variable. Puede ser **Escala** cuando la variable es numérica y la escala es de intervalo, **Ordinal** cuando los valores de la variable representan categorías con un cierto orden o **Nominal** cuando los valores representan categorías sin orden.

Rol Permite especificar la función que una variable tiene en el análisis. Puede ser **entrada**, cuando se trata de una variable independiente, **objetivo**, cuando es una variable dependiente, **ambos**, cuando la variable puede ser dependiente e independiente, **ninguna** si la variable no tiene ninguna función asignada, **particion**, cuando la variable se utilizará para dividir los datos en muestras separadas y **segmentar**, cuando se trata de una variable introducida para asegurar la compatibilidad en SPSS.

Una vez definidas las variables se procede a introducir los datos de la muestra. Para ello hay que volver a la **Ventana de datos** haciendo click en la correspondiente pestaña. Ahora aparecerán en las cabeceras de columna los nombres de las variables definidas. Cada individuo de la muestra se corresponde con una fila de la matriz de datos. Para introducir el valor de una variable en un individuo determinado, nos situamos en la celda de la fila de dicho individuo y de la columna de la variable, bien haciendo click sobre la misma, o bien desplazándonos por la matriz de datos con las flechas de movimiento del cursor del teclado, y se teclea el valor seguido de la tecla **Intro** (figura 1.3).

| | Nombre | Edad | Sexo | Peso | Altura |
|----|------------------|------|------|-------|--------|
| 1 | José Luis | 18 | 1 | 85,00 | 179,00 |
| 2 | Rosa Díez | 32 | 2 | 65,00 | 173,00 |
| 3 | Javier García | 24 | 1 | 71,00 | 181,00 |
| 4 | Carmen López | 35 | 2 | 65,00 | 170,00 |
| 5 | Cristobal Campos | 44 | 1 | 70,00 | 178,00 |
| 6 | Marisa López | 46 | 2 | 51,00 | 158,00 |
| 7 | Antonio Ruiz | 68 | 1 | 66,00 | 174,00 |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |
| 11 | | | | | |
| 12 | | | | | |

Figura 1.3 – Introducción de datos en la matriz de datos. Cada columna corresponde a una variable y cada fila a un individuo de la muestra.

2.3 Guardar datos

Una vez introducidos los datos, conviene guardarlos en un fichero para no tener que volver a introducirlos en futuras sesiones. Para ello, se selecciona el menú **Archivo->Guardar**. Si el fichero ya existe, se actualizará su información, y si no, aparecerá un cuadro de diálogo en el que hay que introducir el nombre que queremos darle al fichero y la carpeta donde lo queremos ubicar. Los ficheros de datos de SPSS tienen por defecto extensión ***.sav**. Cuando los datos estén guardados en un fichero, el nombre del fichero aparecerá en el título de la ventana de datos (figura 1.3).

2.4 Recuperar datos

Si los datos con los que se pretende trabajar ya están guardados en un fichero, entonces tendremos que abrir dicho fichero. Para ello, se selecciona el menú **Archivo->Abrir->Datos** y se selecciona el fichero que se desea abrir. Automáticamente, los datos aparecerán en la vista de datos.

2.5 Modificación de datos

En ocasiones es necesario modificar los datos de la matriz de datos para corregir errores, añadir nuevos datos o eliminarlos. Para corregir un valor basta con seleccionar la celda que contiene el valor y teclear el nuevo. Otras operaciones habituales son:

- Insertar una variable nueva entre otras ya existentes. En la vista de variables se selecciona la fila que contiene la variable por encima de la cual queremos insertar la nueva, y se selecciona el menú **Edición->Insertar variable**.
- Eliminar una variable. En la vista de variables se selecciona la fila que contiene la variable a eliminar y se pulsa la tecla **Supr.**
- Insertar un individuo entre otros ya existentes. En la vista de datos se selecciona la fila que contiene los datos del individuo por encima del cual queremos insertar el nuevo, y se selecciona el menú **Edición->Insertar caso**.
- Eliminar un individuo. En la vista de datos se selecciona la fila que contiene los datos del individuo a eliminar y se se presiona la tecla **Supr.**

Cada vez que realicemos modificaciones en la matriz de datos, conviene volver a guardar los datos para que se actualice el fichero que los contiene.

¡Importante!: Cuando por equivocación realicemos una operación no deseada, podemos deshacerla mediante el menú **Edición->Deshacer**.

2.6 Transformación y generación de datos

En muchos análisis estadísticos se suelen transformar los datos de las variables originales en otros más convenientes para el análisis que se vaya a efectuar. Para generar una nueva variable mediante una transformación de otra ya existente o bien mediante funciones ya predefinidas se selecciona el menú **Transformar->Calcular Variable...** Entonces aparece la ventana de transformación de variables tal y como se muestra en la figura 1.4.

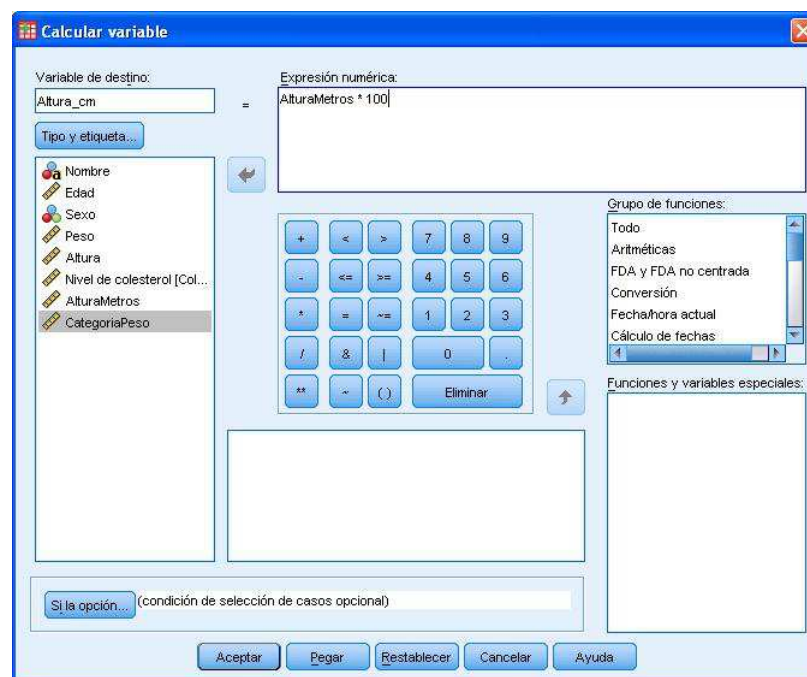


Figura 1.4 – Ventana de transformación de variables. A la izquierda aparecen las variables ya definidas, a la derecha las funciones predefinidas que pueden utilizarse, y en el centro los operadores aritméticos y relacionales más comunes.

En esta ventana se debe introducir el nombre de la nueva variable en el cuadro **Variable de destino**, y la expresión cuyo resultado será el contenido de la nueva variable en el cuadro **Expresión numérica**. Para ello aparecen toda una serie de operadores y funciones para realizar la transformación, así como

la lista de variables ya definidas que pueden utilizarse como argumentos de las distintas funciones de transformación.

Los operadores más habituales para construir expresiones son los aritméticos +, -, *, /, ** (potenciación), los relacionales =, <, >, ~=, <=, >= y los lógicos & (Y), | (O) y ~ (negación). Y algunas de las funciones más habituales son: **ABS** (valor absoluto), **SQRT** (raíz cuadrada), **EXP** (exponencial), **LN** (logaritmo neperiano), **SIN** (seno), **COS** (coseno), **TAN** (tangente), **SUM** (suma), **MEAN** (media aritmética), **SD** (desviación estándar), **RND** (redondeo al entero más cercano), **TRUNC** (parte entera de un número).

Haciendo click en el botón **Si la opción...** se pueden establecer condiciones de aplicación de la transformación. Para establecer una condición debemos activar la opción **Incluir si el caso satisface la condición** y después introducir una condición lógica como por ejemplo **Sexo=1**. De este modo, la transformación sólo se aplicará a los individuos que cumplan dicha condición.

Una vez definida la expresión hay que hacer click sobre el botón **Aceptar** y automáticamente aparecerá en la vista de datos una nueva columna con los datos transformados de la nueva variable.

2.7 Recodificación de datos

Otra forma de transformar una variable es crear otra cuyos valores sean una recodificación de los de la primera, por ejemplo agrupando en intervalos. Esta recodificación podemos hacerla tanto en la misma variable como en variables diferentes. Para ello se selecciona el menú **Transformar->Recodificar en distintas variables**. Automáticamente aparece la ventana de recodificación de variables tal y como se ve en la figura 1.5.

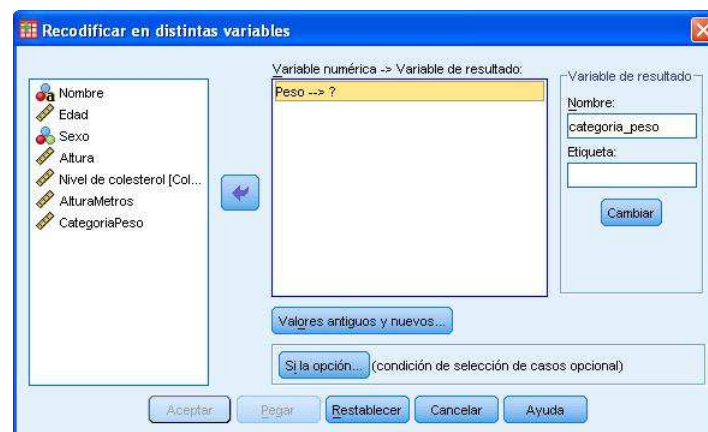


Figura 1.5 – Ventana de recodificación de variables. A la izquierda aparecen las variables ya definidas, a la derecha deben especificarse las reglas de recodificación.

Para recodificar una variable en otra nueva, primero debemos seleccionar la variable que queremos recodificar y hacer click sobre el botón con una flecha que aparece al lado. Después hay que escribir el nombre de la nueva variable en el cuadro **Nombre** y hacer click sobre el botón **Cambiar**. A continuación hay que establecer las reglas de recodificación. Para ello hay que hacer click en el botón **Valores antiguos y nuevos** para que aparezca la ventana de definición de reglas (figura 1.6). Las reglas pueden establecer la conversión del valor de la variable original que introduzcamos en el cuadro **Valor antiguo** en el valor de la variable nueva que introduzcamos en el cuadro **Valor nuevo**, o bien la conversión de todo un intervalo de valores de la variable original en un valor de la variable nueva. Una vez definidos dichos valores hay que hacer click sobre el botón **Continuar**, y después sobre **Aceptar**.

2.8 Impresión

Para imprimir se utiliza el menú **Archivo->Imprimir**. Al instante aparece un cuadro de diálogo para la impresión donde debemos indicar si queremos imprimir todo o bien la selección que hayamos hecho. Tras esto se hace click sobre el botón **Aceptar** y la información se envía a la impresora.

Antes de imprimir conviene hacer una previsualización de lo que se va a enviar a la impresora para estar seguros de que es eso lo que se quiere. Para ello se utiliza el menú **Archivo->Presentación**

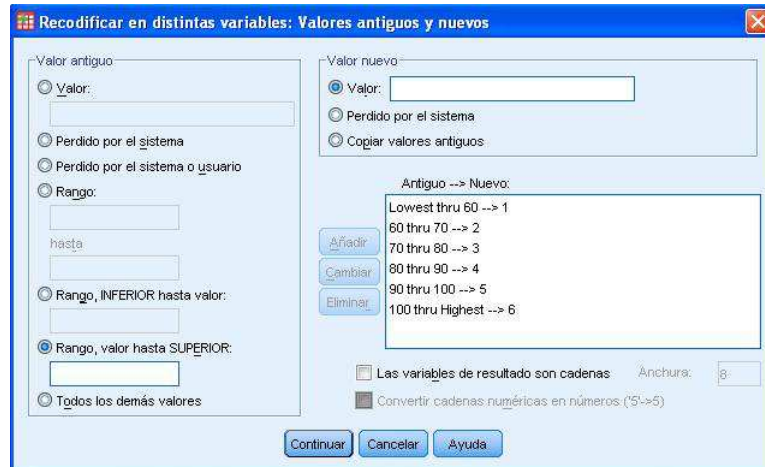


Figura 1.6 – Ventana de definición de reglas.

preliminar. Entonces aparece un visor donde se ve la página, tal y como se enviará a la impresora. Si todo parece correcto se puede hacer click sobre el botón **Imprimir** y aparecerá el cuadro de diálogo de impresión desde el que se puede enviar a la impresora definitivamente.

2.9 Salir del programa

Para terminar una sesión de trabajo se utiliza el menú **Archivo->Salir**, o bien se hace click sobre el aspa para cerrar la ventana del programa. Si quedan datos o resultados que no se han guardado, el programa nos preguntará antes de salir si deseamos guardarlos.

2.10 Ayuda

En esta práctica sólo hemos descrito las operaciones básicas en una sesión de trabajo. Pero quedan por describir todos los análisis estadísticos que pueden realizarse con los menús de la barra de menús. Aunque muchos de estos menús se explicarán en las siguientes prácticas, el programa dispone del menú de ayuda **Ayuda** en el que podemos encontrar una descripción de todos estos menús y al que podemos recurrir cada vez que tengamos dudas.

3 Ejercicios resueltos

1. Introducir en la matriz de datos los datos de la siguiente muestra y guardarlos en un fichero con el nombre `datos_colesterol.sav`.

| Nombre | Sexo | Peso | Altura | Colesterol |
|---------------------------------|------|------|--------|------------|
| José Luis Martínez Izquierdo | H | 85 | 179 | 182 |
| Rosa Díaz Díaz | M | 65 | 173 | 232 |
| Javier García Sánchez | H | 71 | 181 | 191 |
| Carmen López Pinzón | M | 65 | 170 | 200 |
| Marisa López Collado | M | 51 | 158 | 148 |
| Antonio Ruiz Cruz | H | 66 | 174 | 249 |
| Antonio Fernández Ocaña | H | 62 | 172 | 276 |
| Pilar Martín González | M | 60 | 166 | 213 |
| Pedro Gálvez Tenorio | H | 90 | 194 | 241 |
| Santiago Reillo Manzano | H | 75 | 185 | 280 |
| Macarena Álvarez Luna | M | 55 | 162 | 262 |
| José María de la Guía Sanz | H | 78 | 187 | 198 |
| Miguel Angel Cuadrado Gutiérrez | H | 109 | 198 | 210 |
| Carolina Rubio Moreno | M | 61 | 177 | 194 |

Indicación

- En la ventana de **Vista de variables**, crear las variables **Nombre**, **Sexo**, **Peso**, **Altura** y **Colesterol** e introducir los datos anteriores, siguiendo las indicaciones del apartado 1.2.2.
- Una vez introducidos los datos, se guardan en un fichero de nombre `datos_colesterol` siguiendo lo indicado en el apartado 1.2.3.

2. Sobre la matriz de datos del ejercicio anterior realizar las siguientes operaciones:

- Insertar detrás de la variable **Nombre** una nueva variable **Edad** con las edades de todos los individuos de la muestra.

| Nombre | Edad |
|---------------------------------|------|
| José Luis Martínez Izquierdo | 18 |
| Rosa Díaz Díaz | 32 |
| Javier García Sánchez | 24 |
| Carmen López Pinzón | 35 |
| Marisa López Collado | 46 |
| Antonio Ruiz Cruz | 68 |
| Antonio Fernández Ocaña | 51 |
| Pilar Martín González | 22 |
| Pedro Gálvez Tenorio | 35 |
| Santiago Reillo Manzano | 46 |
| Macarena Álvarez Luna | 53 |
| José María de la Guía Sanz | 58 |
| Miguel Angel Cuadrado Gutiérrez | 27 |
| Carolina Rubio Moreno | 20 |

Indicación

- En la **Vista de variables** seleccionar la fila correspondiente a la variable **Sexo** haciendo click con el ratón sobre la cabecera de la misma y a continuación seleccionar el menú **Edición->Insertar variable**, con lo que aparece una nueva fila entre las variables **Nombre** y **Sexo**.
- En la nueva fila definir la variable **Edad** e introducir los datos anteriores.
- En la **Vista de datos** rellenar los datos de la columna correspondiente a la **Edad**.

- Insertar entre los individuos 4º y 5º los datos correspondientes al siguiente individuo

Nombre: Cristóbal Campos Ruiz.

Edad: 44 años.

Sexo: Hombre.

Peso: 70 Kg.

Altura: 178 cm.

Colesterol: 220 mg/dl.

Indicación

- 1) Seleccionar la fila correspondiente al 5º individuo, haciendo click con el ratón sobre la cabecera de la misma y a continuación seleccionar el menú **Edición->Insertar caso**, con lo que aparece una nueva fila entre las correspondientes a los individuos 4º y 5º.
- 2) Introducir en la nueva fila los datos que se indican.

- c) Cambiar el valor de la variable **Peso** de Macarena Álvarez Luna por 58.

Indicación

Hacer click con el ratón en la casilla cuyo contenido se desea modificar, escribir 58 y pulsar **Enter**.

- d) Transformar la variable **Altura** para que aparezca expresada en metros.

Indicación

- 1) Seleccionar el menú **Transformar->Calcular variable**.
- 2) En la ventana de transformación de datos introducir el nombre **Altura_metros** en el cuadro **Variable de destino**.
- 3) Introducir la expresión **Altura/100** en el cuadro **Expresión numérica**.
- 4) Hacer click sobre el botón **Aceptar**.

- e) Recodificar la variable peso en las siguientes cuatro categorías, teniendo en cuenta el sexo:

| Categoría | Hombres | Mujeres |
|-----------|-----------|-----------|
| Bajo | ≤ 70 | ≤ 60 |
| Medio | (70,85] | (60,70] |
| Alto | (85,100] | (70,80] |
| Muy Alto | > 100 | > 80 |

Indicación

- 1) Seleccionar el menú **Transformar->Recodificar** en distintas variables.
- 2) En la ventana de recodificación de datos seleccionar la variable **Peso** y hacer click sobre el botón con una flecha que aparece al lado.
- 3) Escribir el nombre de la variable recodificada **Categoria_Peso** en el cuadro **Nombre** de la **Variable de resultado** y hacer click sobre el botón **Cambiar**.
- 4) Hacer click en el botón **Valores antiguos y nuevos** para abrir la ventana de definición de reglas de recodificación.
- 5) Para definir las reglas de recodificación de los hombres,
 - a' Seleccionar la opción **Rango INFERIOR hasta valor** del cuadro **Valor antiguo** e introducir 70 en el cuadro correspondiente. Introducir 1 en el cuadro de la opción **Valor** del cuadro **Valor nuevo** y hacer click en el botón **Añadir**.
 - b' Seleccionar la opción **Rango** del cuadro **Valor antiguo** e introducir 70 en el cuadro correspondiente y 85 en el cuadro **hasta**. Introducir 2 en el cuadro de la opción **Valor** del cuadro **Valor nuevo** y hacer click en el botón **Añadir**.
 - c' Seleccionar la opción **Rango** del cuadro **Valor antiguo** e introducir 85 en el cuadro correspondiente y 100 en el cuadro **hasta**. Introducir 3 en el cuadro de la opción **Valor** del cuadro **Valor nuevo** y hacer click en el botón **Añadir**.
 - d' Seleccionar la opción **Rango valor hasta SUPERIOR** del cuadro **Valor antiguo** e introducir 100 en el cuadro correspondiente. Introducir 4 en el cuadro de la opción **Valor** del cuadro **Valor nuevo** y hacer click en el botón **Añadir**.
- 6) Hacer click en el botón **Continuar** para cerrar la ventana.
- 7) Hacer click en el botón **Si la opción...** para abrir la ventana de definición de condiciones.
- 8) Seleccionar la opción **Incluir si el caso satisface la condición** e introducir la condición **Sexo='H'** en el cuadro correspondiente.
- 9) Hacer click en el botón **Continuar** para cerrar la ventana.
- 10) Hacer click en el botón **Aceptar**.
- 11) Repetir los mismos pasos para establecer las reglas de codificación de las mujeres.
- 12) En **Vista de variables** hacer clic en **Valores de Categoria_Peso**, y en **Etiquetas de valor** ir asignando a los valores 1, 2, 3 y 4 las etiquetas **Bajo**, **Medio**, **Alto** y **Muy alto** respectivamente, haciendo click en el botón **Añadir** después de cada asignación, y una vez terminado hacer click en el **Aceptar**.

f) Volver a guardar los cambios en el fichero anterior y salir del programa.

Indicación

- 1) Seleccionar el menú **Archivo->Guardar**.
- 2) Seleccionar el menú **Archivo->Salir**.

4 Caso práctico

Se ha diseñado un ensayo clínico aleatorizado, doble-ciego y controlado con placebo, para estudiar el efecto de dos alternativas terapéuticas en el control de la hipertensión arterial. Se han reclutado 100 pacientes hipertensos y estos han sido distribuidos aleatoriamente en tres grupos de tratamiento. A uno de los grupos (control) se le administró un placebo, a otro grupo se le administró un inhibidor de la enzima convertidora de la angiotensina (IECA) y al otro un tratamiento combinado de un diurético y un Antagonista del Calcio. Las variables respuesta final fueron las presiones arteriales sistólica y diastólica.

Los datos con las claves de aleatorización han sido introducidos en una base de datos que reside en la central de aleatorización, mientras que los datos clínicos han sido archivados en dos archivos distintos, uno para cada uno de los dos centros participantes en el estudio.

Las variables almacenadas en estos archivos clínicos son las siguientes:

- CLAVE Clave de aleatorización
- NOMBRE Iniciales del paciente
- F_NACIM Fecha de Nacimiento
- F_INCLUS Fecha de inclusión

- SEXO Sexo (0: Hombre 1: Mujer)
- ALTURA Altura en cm.
- PESO Peso en Kg.
- PAD_INI Presión diastólica basal (inicial)
- PAD_FIN Presión diastólica final
- PAS_INI Presión sistólica basal (inicial)
- PAS_FIN Presión sistólica final

El archivo de claves de aleatorización contiene sólo dos variables.

- CLAVE Clave de aleatorización
- FARMACO Fármaco administrado (0: Placebo, 1: IECA, 2:Ca Antagonista + diurético)

Se pide:

1. Leer los datos del centro con 10 pacientes, incluidos en el archivo Hipertensos HA.xls, este hospital trabaja con la hoja de cálculo Excel.

Indicación

Desplegar el menú **Archivo->Abrir->Datos**, e ir hasta el directorio que contiene el archivo Hipertensos HA.xls, escogiendo como tipo de archivo el formato de Excel. Una vez tenemos delante el nombre del archivo, para abrirlo será suficiente con un doble clic, pero teniendo en cuenta que hay que activar la opción **Leer nombres de variables** si queremos que utilice la primera fila del archivo de Excel para dar nombre a las variables en el archivo de datos de SPSS.

2. Añadir a estos, los datos de los pacientes 11 al 100, incluidos en el archivo Hipertensos HB y guardar los datos como Hipertensos totales.

Indicación

- a) Teniendo en cuenta que lo que pretendemos es añadir nuevos casos a un archivo de datos ya abierto, el menú a utilizar es **Datos->Fundir archivos->Añadir casos**. Después activamos la opción **Un archivo de datos de SPSS Statistics externo**, y con el botón **Examinar** accedemos hasta la carpeta que contenga el archivo Hipertensos HB. Una vez seleccionado, utilizamos el botón **Continuar**, y posteriormente el botón **Aceptar** en el siguiente cuadro de diálogo que aparece.
- b) Una vez generado el archivo de datos, para guardarlo podemos utilizar el menú **Archivo->Guardar como**.

3. Fusionar los datos clínicos con las claves de aleatorización. El fichero con las claves se denomina Claves aleatorización. Grabar el archivo resultante con el nombre Hipertensos Datos Claves

Indicación

De forma similar al apartado anterior pero teniendo en cuenta que ahora lo que pretendemos es añadir variables. Para ello utilizamos el menú: **Datos->Fundir archivos->Añadir variables**, para acceder después hasta el archivo Claves aleatorización, y posteriormente vamos aceptando en todos los cuadros de diálogo que aparecen. Una vez generado, guardamos el nuevo archivo de datos y claves: menú **Archivo->Guardar como**.

4. Crear un archivo para cada uno de los grupos de tratamiento. Denominar a estos archivos Hipertensos placebo, Hipertensos IECA e Hipertensos Ca respectivamente.

Indicación

Se podría lograr el mismo resultado de múltiples maneras. Por ejemplo:

- a) Segmentando el archivo mediante el menú **Datos->Segmentar archivo**, y **Organizar los resultados por grupos** basados en la variable **farmaco**. Una vez segmentado el archivo, podemos marcar los casos correspondientes a cada uno de los fármacos y hacer un cortar y pegar en un nuevo archivo específico para cada fármaco.
- b) También podemos generar los nuevos archivos utilizando el sistema de filtros de SPSS. Para ello el menú a utilizar es **Datos->Seleccionar Casos**, con la opción **Si se satisface la condición**, botón **Si la op...**, y entramos en un cuadro de diálogo en el que damos forma a la condición, que en primera instancia será **farmaco=0**, para volver al cuadro de diálogo anterior y escoger la opción **Copiar casos seleccionados a un nuevo conjunto de datos** y poner en **Nombre de conjunto de datos**: **Hipertensos_placebo**. Igualmente, repetiríamos el proceso para los otros dos fármacos.

5. Calcular, para cada paciente, la edad en años el día de la incorporación al estudio (redondeando al entero más próximo). Denominar la nueva variable **edad** y etiquetarla correspondientemente.

Indicación

Internamente SPSS trabaja con las variables en formato fecha almacenando el número de segundos transcurridos desde el comienzo del Calendario Gregoriano en el año 1582 hasta la fecha que introducimos. Por lo tanto, si restamos dos variables en formato fecha, no nos da el número de años transcurridos entre una y otra, sino el número de segundos. Por ello, la nueva variable obtenida como resultado de la resta hay que dividirla entre el número de segundos que tiene un año a razón de 3600 segundos la hora, 24 horas el día, y 365,25 días el año, aproximadamente. Teniendo en cuenta lo anterior, el proceso a utilizar es: Menú **Transformar**→**Calcular variable**, y en **Variable de destino**, escribimos **edad**. Como expresión numérica para su cálculo introducimos:

$$(F_INCLUS - F_NACIM) / (365.25 * 24 * 3600)$$

6. Recodificar dicha edad de forma que la nueva variable, de nombre **grupoedad**, tome los siguientes valores y etiquetas de valor:

| Edad en años | Grupo edad | Etiqueta |
|----------------------------|------------|--------------------|
| Edad ≤ 37 | 1 | Hasta 37 inclusive |
| $37 < \text{Edad} \leq 44$ | 2 | De 37 a 44 |
| $44 < \text{Edad} \leq 51$ | 3 | De 44 a 51 |
| Edad > 51 | 4 | Mayores de 51 |

Indicación

Para recodificar una variable, se utiliza el proceso ya explicado en la práctica de Introducción a SPSS, con el menú **Transformar**→**Recodificar en distintas variables**, escogiendo **edad** como **Variable de entrada**, y dando el nombre **grupoedad** a la **Variable de resultado**, utilizando el botón **Cambiar**, y posteriormente el botón **Valores antiguos y nuevos** para delimitar las categorías de la nueva variable.

7. Calcular, para cada paciente, el índice de masa corporal (se obtiene dividiendo el peso, expresado en kg, entre la altura, expresada en m, elevada al cuadrado) y almacenar el resultado en la variable **masacorp**.

Indicación

Con el menú **Transformar**→**Calcular variable**, escogiendo como variable de destino **masacorp**, y como expresión numérica la indicada en el enunciado.

8. Recodificar dicho índice de masa corporal, de forma que la nueva variable, de nombre **obesidad**, tome los siguientes valores y etiquetas de valor según el sexo del paciente.

| Sexo | Masacorp | Obesidad | Etiqueta |
|------|----------------|----------|------------|
| 0 | ≤ 21 | 1 | Desnutrido |
| | (21, 26,94] | 2 | Normal |
| | (26,94, 32,94] | 3 | Sobrepeso |
| | (32,94, 43,94] | 4 | Obeso |
| | $> 43,94$ | 5 | Muy obeso |
| 1 | ≤ 19 | 1 | Desnutrido |
| | (19, 24,94] | 2 | Normal |
| | (24,94, 29,94] | 3 | Sobrepeso |
| | (29,94, 39,94] | 4 | Obeso |
| | $> 39,94$ | 5 | Muy obeso |

Indicación

Se trata de un problema de recodificación, muy parecido al explicado en la indicación del punto 6, con la única novedad de que ahora hay que hacer una doble recodificación: por un lado para hombres y por el otro para mujeres. Para ello, después de escoger las variables de entrada y de resultado, haciendo uso del botón condicional Si la opción podemos escoger únicamente los casos que cumplen la condición impuesta por la variable **Sexo**; es decir, con la condición **Sexo=0**, recodificamos en primera instancia el índice de masa corporal de los hombres, y con **Sexo=1** recodificamos el de las mujeres.