INTRODUCCIÓN AL MANEJO DEL PROGRAMA SPSS 12.0

Juan de Dios Luna del Castillo Francisco Requena Guerrero Pedro Femia Marzo Antonio Martín Andrés María Teresa Miranda León

Septiembre 2007

ÍNDICE DE CONTENIDOS.

0. INTRODUCCIÓN BÁSICA AL SISTEMA OPERATIVO WINDOWS 95	4
0.1. El escritorio de Windows	4
0.2. El explorador de Windows	5
GENERALIDADES	11
1. Generalidades sobre SPSS.	11
1.1. El procedimiento general de resolución de un problema con SPSS	11
1.2. La ventana principal de SPSS: el Editor de datos de SPSS.	11
1.3. Normas de funcionamiento de las ventanas de SPSS.	15
1.4. Uso de los botones del ratón	16
2. Cómo está organizado el manual.	16
2.1. Un ejemplo que nos acompañará a lo largo del manual.	18
3. Creación de un fichero de datos.	18
3.1. Definición de las variables.	18
3.1.1. El nombre de las variables.	. 19
3.1.2. El tipo de las variables	. 19
3.1.3. Las etiquetas de las variables	20
3.1.4. Los datos faltantes (valores perdidos).	21
3.1.5. El formato de columna de las variables.	22
3.1.6. La definición de la Escala de Medida de una variable.	. 22
3 1 7 Estructura del fichero de datos	22
3.2 Grabación de los datos	
3.2.1. Introducción de los datos por casillas, por filas y por columnas.	
3.2.2. Acciones de Cortar. Copiar y Pegar.	23
4. Edición de un fichero de datos.	24
4.1. Edición de casos completos	24
4.2. Edición de variables completas.	25
5. Operaciones con variables. Generación de nuevas variables.	26
5.1. Transformación de variables. Procedimiento Calcular.	26
5.2. Recodificación de variables. Procedimiento Recodificar.	29
6. Ordenación y selección de casos.	32
6.1. Procedimiento Ordenar casos.	32
6.2. Procedimiento Seleccionar casos	33
7. El Visor de Resultados.	35
7.1. La zona de resultados: el índice y los resultados propiamente dichos	35
7.2. Menús e Iconos del Visor de resultados.	37
8. Estadística Descriptiva con SPSS.	38
8.1. Procedimiento Frecuencias	39
8.2. Gráficos con SPSS	40
8.2.1. Gráficos de Barras	. 41
8.2.2. Gráficos de Líneas.	. 44
8.2.3. Gráficos de Sectores.	44
8.2.4. Histograma	44
9. Estimación y tests de hipótesis básicos con SPSS.	45
9.1. Procedimiento Explorar.	45
9.2. Procedimiento Binomial.	46
10. Tests de comparación de medias con SPSS.	48
10.1. Procedimiento Prueba T para muestras independientes.	48
10.2. Procedimiento Prueba T para muestras relacionadas.	50
10.3. Gráfico Barras de error.	51
10.4. Procedimiento Pruebas no-paramétricas para dos muestras independientes	52
10.5. Pruebas no-paramétricas para dos muestras relacionadas.	54

11. El test Chi-cuadrado con SPSS.	55
11.1. Procedimiento Tablas de contingencia.	55
Tablas de contingencia	57
12. Regresión y correlación lineal simple con SPSS	59
12.1. Gráficos de Dispersión (nubes de puntos).	59
12.2. El procedimiento Regresión lineal	60
12.3. El procedimiento Correlaciones bivariadas.	

0. INTRODUCCIÓN BÁSICA AL SISTEMA OPERATIVO WINDOWS 95

0.1. El escritorio de Windows

En la terminología de Windows, por *escritorio* se entiende la pantalla de fondo sobre la cual se disponen los *iconos de acceso a recursos* y se ejecutan las *aplicaciones*. Además del *fondo*, el escritorio posee otro componente que es la *barra de tareas*. Se trata de una barra que aparece normalmente en disposición horizontal en la parte inferior de la pantalla (aunque su localización se puede cambiar a gusto del usuario). En esta barra aparece en un extremo el *botón de inicio* que activa el menú principal de Windows. En el otro extremo es frecuente que aparezca un reloj y quizá una serie de iconos asociados a programas que de alguna manera siempre están activos (por ejemplo el antivirus, el indicador de correo electrónico, etc). En la figura siguiente se indican estos componentes del escritorio.



Elementos mas frecuentes del escritorio de Windows.

Simultaneidad de tareas. Obsérvese que en la barra de tareas aparecen botones que representan cada una de las aplicaciones que se tienen abiertas. En la figura, aunque hay dos ventanas visibles, una del Microsoft-Word y otra del SPSS, en realidad hay dos aplicaciones más que están en uso. Estas aplicaciones que tienen un botón en la barra de tareas pero que no tienen su ventana abierta en el escritorio se dice que están *minimizadas*. Para poder trabajar con ellas basta con hacer un *clic* con el botón principal del ratón (que normalmente es el izquierdo, aunque como todo, esto se puede cambiar a gusto del usuario) en el botón correspondiente de la barra de tareas. El efecto de esta acción es traer a un primer plano la ventana de la aplicación seleccionada y se dice entonces que se ha *maximizado* dicha aplicación. La idea entonces, es que Windows permite tener muchas aplicaciones abiertas al mismo tiempo y cada una de ellas tendrá su correspondiente botón en la barra de tareas. Cuando no vamos a utilizar una aplicación durante un tiempo pero no queremos cerrarla definitivamente, lo mejor es minimizarla (lo que se traduce en reducir los recursos que el ordenador otorga a dicha aplicación). Para ello, la mayoría de las ventanas tienen un conjunto de tres iconos en su esquina superior derecha:



El icono de la izquierda es precisamente el que sirve para *minimizar* la ventana. El central permite cambiar su tamaño: si este icono tiene el aspecto , indica que pulsándolo se conseguirá que la ventana ocupe toda la extensión de la pantalla (*maximización*). Al hacerlo, el icono cambiará a la forma , que permite *restaurar* las dimensiones de la ventana pero sin minimizarla. Finalmente, el icono de la derecha sirve para cerrar definitivamente la aplicación.

0.2. El explorador de Windows

Al pulsar el botón se despliega el *menú principal de Windows* que es el que nos va a permitir acceder a todas las aplicaciones y recursos disponibles. En este menú podemos encontrar un acceso a una aplicación de especial interés: *el explorador de Windows*, identificado con el icono (que con frecuencia se encuentra también en el escritorio). La ventana típica del explorador se muestra a continuación.



Ventana del explorador de Windows

Barras del explorador. Como toda ventana de Windows, la del explorador presenta en su parte superior una línea de título (con los iconos habituales de control de la ventana) y una barra de menús. En la parte inferior suele aparecer una barra de estado en la que se muestra diversa información, como la ocupación del disco, etc. Además de las ya indicadas pueden aparecer mas barras en la parte superior de la ventana, como son la barra de herramientas y la de dirección que se aprecian en la figura anterior. Para mostrar (u ocultar) estas barras debemos irnos al menú Ver/Barras de herramientas. Allí aparecen marcadas con el símbolo \checkmark aquellas barras que son visibles en cada momento. Para hacer que aparezca o desaparezca este símbolo basta con hacer clic en la opción deseada.

La ventana del explorador se divide verticalmente en dos áreas o, abusando del lenguaje, en dos ventanas. En la de la izquierda aparece una "estructura de árbol" conocida como *árbol de directorios*. La *raíz* del árbol de directorios (de donde "cuelga" todo lo demás) es el escritorio. Obsérvese que a la izquierda de muchos de los iconos aparece un signo \oplus , esto quiere decir que el objeto al que corresponde el icono puede *expandirse* para mostrar otros objetos que están contenidos en él. Para conseguirlo basta con hacer un clic con el botón principal del ratón en el signo \oplus . Cuando se ha *expandido la rama del* árbol el símbolo cambia

a 🖃, indicando que ahora la operación posible será la de *contraerla*. En la figura siguiente se muestran estos efectos



Cuando seleccionemos un objeto concreto en la ventana izquierda, su contenido aparecerá siempre en la ventana de la derecha. Si el objeto seleccionado era una carpeta, veremos cómo, además, su icono cambia de a a, de manera que en la ventana derecha aparecerán las carpetas y archivos que estén contenidos en ella. La vista de los objetos de la ventana derecha se puede establecer en alguno de los siguientes modos:

• Iconos grandes; • Iconos pequeños; • Lista; • Detalles

Una forma de seleccionar el modo deseado es a través del menú Ver. La opción **Detalles** es la mas interesante, pues informa no solo del nombre de cada objeto sino también de su *tipo* (si es una carpeta, un documento de texto, una aplicación, etc.), del tamaño que ocupa y de la fecha de la última modificación.

Nombre	Tamaño	Tipo	Modificado
<u> Sss1</u>		Carpeta de archivos	09/05/00 12:58
🚞 <u>epi6 new</u>		Carpeta de archivos	09/05/00 12:57
Datos		Carpeta de archivos	09/05/00 12:57
Epi6.cfg	1 KB		27/06/00 11:43
Spear.rec	8 K B		18/05/00 11:52
<u>ја Еріб</u>	1 KB	Acceso directo al progr	17/05/00 10:44

Vista en detalle de los elementos de una carpeta

En este modo es posible ordenar el contenido de la ventana por cualquiera de las propiedades indicadas y de forma ascendente o descendente. Por defecto, la ordenación es siempre de forma ascendente según el nombre, pero si hacemos un clic con el botón principal del ratón en la cabecera tentonces la ordenación de los objetos será de menor a mayor tamaño. Un nuevo clic en tenta ahora que la ordenación sea de mayor a menor. Lo mismo es valido para el nombre, el tipo o la fecha de modificación.

Menús contextuales. Si hacemos clic con el botón secundario del ratón (habitualmente el derecho) sobre cualquier objeto¹, entonces se despliega el denominado *menú contextual*, que es un menú con opciones particularizadas sobre dicho objeto. Esto ocurre en general en la mayoría de las aplicaciones Windows, incluido el SPSS. Por tanto es muy útil recordar que hacer un clic con el botón derecho del ratón sobre 'cualquier cosa' generalmente nos da una lista de opciones entre las cuales se encuentra aquella que estábamos necesitando.

Operaciones con carpetas y archivos. A continuación se van a presentar las operaciones básicas que se pueden realizar con el contenido de una carpeta. Una constante en el entorno Windows es que una misma operación pueda llevarse a cabo de formas distintas. De forma general, las operaciones más frecuentes suelen poder hacerse 1) mediante las opciones del menú principal; 2) mediante algún botón de la barra de herramientas; 3) mediante opciones de los menús contextuales y 4) mediante el arrastre de objetos con el ratón.

• **Operaciones con carpetas.** Para **crear** una carpeta lo primero que debemos hacer es seleccionar en la ventana izquierda del explorador la unidad de disco y la carpeta principal dentro de los cuales vamos a crear la nueva (bastará para ello ir haciendo clic con el botón

¹ Aquí la palabra *objeto* tiene un sentido abstracto y permite aludir a iconos, barras de herramientas, ventanas, archivos, etc

principal del ratón en las carpetas deseadas). El siguiente paso es seleccionar la opción **Nuevo** del menú **Archivo**. Allí la primera opción es **Carpeta**. Al seleccionarla aparecerá en la ventana derecha el icono **Nuevo**. Este es el momento en que podemos escribir el nombre de la nueva carpeta y pulsar [**Intro**] para concluir. Ya tenemos creada una nueva carpeta. La opción **Nuevo/Carpeta** aparece también en el *menú contextual* si pulsamos el botón derecho del ratón en el área en blanco de la ventana derecha.

Es posible que una vez creada la carpeta deseemos cambiar su ubicación. La mejor forma de **mover** una carpeta es arrastrarla (pulsando el botón izquierdo del ratón) y soltarla encima de la nueva ubicación. Para ello obviamente, debemos tener a la vista tanto la carpeta que deseamos mover como el destino donde queremos colocarla. Formas alternativas de hacerlo es utilizar la ventana izquierda del explorador o arrastrar desde la derecha hacia la izquierda. A continuación se reproduce el proceso considerando la primera de estas alternativas:



Finalmente, la forma de **eliminar** una carpeta es arrastrarla, del modo indicado, hasta el icono de la papelera de reciclaje (ya sea el que aparece en el escritorio o el del propio explorador), o también pulsar el botón \bowtie de la barra de herramientas, o seleccionarla con el ratón y pulsar la tecla [**Supr**], y por supuesto, al eliminarla se elimina también todo su contenido.

Con las carpetas se pueden hacer algunas cosas más, basta con pulsar el botón secundario del ratón sobre el icono de la carpeta deseada para ver las alternativas posibles. Veremos algunas de ellas aludiendo a los archivos en general.

Operaciones con archivos. La palabra archivo o fichero en informática viene a ser equivalente, en un sentido amplio, a documento. Los datos de SPSS estarán almacenados en el disco en forma de archivos, los informes también y si escribimos una tesis doctoral su formato informático será también el de uno o muchos archivos. Toda la información que se guarde en un disco lo hace en forma de archivo informático. Es obvio entonces que un archivo es algo bastante mas complejo que una simple carpeta, pues esta última no es más que un contenedor (como ocurre con las carpetas que compramos en una papelería). En consecuencia, la creación de archivos implica el manejo de algún programa o aplicación informática -en el sentido de herramienta- que permita especificar y manipular la información deseada en el formato conveniente. Así, con la aplicación SPSS haremos archivos con nuestros datos, generaremos archivos con nuestros resultados, gráficos, informes, etc. Hay que destacar también que todo archivo tiene, además del nombre que lo identifica, lo que en informática se conoce como la extensión. Se trata de un conjunto de hasta tres caracteres que se colocan detrás del nombre separándolos de este por un punto. El usuario no debe preocuparse de escribir la extensión cuando le da nombre a un archivo, porque es el propio programa que está usando para generarlo quien lo hace. La idea es que la extensión viene a ser una especie de sello que cada programa pone a sus archivos, de forma que sea inmediato reconocerlos cuando están almacenados en el disco. Por ejemplo, los datos que se manejan con SPSS deberán ser archivos con extensión .sav, los resultados estadísticos serán archivos con extensión .spo, los archivos de imágenes tienen extensiones tales como .bmp, .jpg o .gif, los documentos escritos con Word son archivos con extensión .doc, las páginas que se publican en internet son archivos con extensión .htm, ... y así podríamos citar un buen número de extensiones interesantes. Una bastante importante es la extensión .exe. los archivos que la tienen disfrutan del privilegio de pertenecer a una clase muy especial de archivos: los programas. Un programa es entonces un archivo que sabe cómo poner a su servicio los recursos del ordenador constituyéndose en un intérprete capaz de traducir los deseos del usuario al frío lenguaje máquina. Un detalle final acerca de las extensiones: si no vemos en el explorador la extensión de nuestros archivos es por que Windows la oculta, no por perverso, sino porque existe una opción de configuración (un poquito más avanzada del objetivo de estas líneas) que consiste en mostrar u ocultar las extensiones de los archivos.

Una vez creado un archivo con el programa apropiado para ello, desde el explorador de Windows podemos realizar sobre él diversas acciones como son: Abrir, Copiar, Mover, Crear acceso directo, Cambiar el nombre, Eliminar, Imprimir, Veamos en primer lugar como se realizan estas acciones a través de las opciones de menú.

- Para Abrir un archivo es necesario, ante todo, que en el ordenador este instalado el programa que *sabe* como hacerlo. Por ejemplo, para abrir un archivo con extensión .SAV es necesario que esté instalado el SPSS. Entonces hay dos posibilidades: 1) abrir el programa y desde su propio menú abrir el archivo deseado y 2) En el menú archivo del explorador seleccionar la opción abrir. Windows sabrá entonces con que programa hacerlo gracias a la extensión del archivo.
- Para copiar un archivo en una carpeta diferente realizaremos los siguientes pasos: 1) seleccionar el archivo con el botón principal del ratón; 2) abrir el menú Edición del explorador de Windows y seleccionar la opción Copiar, con esto el archivo no está copiado, sino *señalado* para ser copiado, así que ahora tendremos que 3) ir al disco o la carpeta donde deseemos colocar la copia y 4) abrir de nuevo el menú Edición y seleccionar Pegar. Ahora ya tenemos nuestro archivo duplicado en dos localizaciones distintas. Todo esto merece un par de comentarios adicionales. Las opciones Copiar y Pegar se pueden activar también mediante el menú contextual o con los botones a (copiar) y (pegar) de la barra de herramientas. Si pegamos un archivo en la misma carpeta de donde lo hemos copiado, Windows creara correctamente la copia pero añadiéndole al nombre de esta última el prefijo "Copia de", ya que no puede haber dos objetos con el mismo nombre en una misma carpeta. A partir de la segunda las nuevas copias se numeran, y el efecto es el que se muestra a continuación:



- Mover un archivo significa copiarlo a otra ubicación y eliminarlo de su posición actual. Una forma de hacerlo es 1) seleccionar el archivo a mover; 2) seleccionar la opción <u>Cortar</u> en el menú <u>Edición</u>, o en el menú contextual o bien mediante el botón (cortar) de la barra de herramientas; 3) abrir la carpeta donde deseemos colocar el archivo y allí 4) seleccionar la opción <u>Pegar</u> (mediante el menú <u>Edición</u>, el contextual o pulsando el botón actual de la barra de herramientas).
- Para **cambiar el nombre** a un archivo (o a una carpeta) una de las formas de hacerlo es 1) señalar el archivo deseado con el ratón, 2) pulsar la tecla [F2]. Ahora podemos escribir el nuevo nombre y pulsar [**Intro**] para terminar. El mismo objetivo se consigue a través de la opción **Cambiar nombre** que aparece directamente en el menú contextual.
- Para **eliminar** un archivo seguiremos exactamente los mismos pasos que para eliminar una carpeta.
- En el menú contextual aparecen además otras operaciones que dependen del tipo específico de archivo. Algunas de ellas son **Abrir**, **Editar**, **Imprimir**,.... La disponibilidad de estas acciones depende normalmente de que tengamos instalado el programa que permite crear el archivo en cuestión. Obsérvese que al desplegar el menú contextual normalmente una de las opciones aparece destacada en negrita. Esta es la *opción por defecto*, que es aquella que se realiza al hacer doble clic con el botón principal del ratón sobre el icono del archivo en cuestión.

Muchas de las acciones descritas pueden hacerse también mediante el uso del ratón (y quizá de forma mas sencilla). Si hacemos doble clic con el botón principal del ratón, la acción que tiene lugar es la que esté predefinida para ese objeto. Normalmente esa acción es **abrir**. Así, el doble clic encima del icono de una carpeta la abre, si se hace encima del icono de un archivo de SPSS, entonces abre dicha aplicación y carga el archivo, y así en general ocurre con todos los archivos que Windows *sepa* como abrir (en principio todos aquellos cuyas aplicaciones que los manejan tengamos instaladas de forma que las extensiones asociadas serán conocidas por el sistema). Las acciones del tipo copiar o mover implican arrastrar el icono del archivo o de la carpeta deseada desde su ubicación actual a la carpeta destino. Veamos como se consiguen las distintas funciones. Si arrastramos el objeto con el botón derecho del ratón, al soltarlo aparecerá siempre el menú

<u>C</u> opiar aquí Crear iconos <u>d</u> e acceso directo aquí
Cancelar

Como vemos las funciones **copiar** y **crear acceso directo** resultan inmediatas, pues basta con seleccionar la deseada en este menú, donde también disponemos de la opción **cancelar** y aquí no ha pasado nada. Así pues, si queremos crear un acceso directo a un archivo o a una carpeta, en el escritorio de Windows por ejemplo, la forma mas fácil es (teniendo visible alguna parte del escritorio) arrastrar el icono del archivo o carpeta deseada (desde el explorador) al escritorio con el botón derecho del ratón y seleccionar la segunda opción del menú contextual que aparecerá al soltar dicho botón. Cuando el arrastre se efectúa con el botón principal del ratón entonces es conveniente atender a la forma que adopta el puntero durante el arrastre. A continuación se presentan las 4 alternativas posibles:



Como se puede deducir de la primera, cuando el puntero del ratón, aparece en forma de flecha, la acción que estamos realizando es mover (y no copiar, como se piensa con frecuencia). Esto ocurre con la mayoría de archivos y carpetas que se arrastren desde una ubicación de un determinado disco a otra ubicación del mismo disco. Sin embargo, cuando el arrastre sea desde un disco a otro distinto, la acción es copiar, aunque el puntero del ratón siga teniendo la misma apariencia. Como caso excepcional, cuando lo que se arrastra es un programa (archivos con extensión .exe) entonces por defecto lo que estamos haciendo es crear un acceso directo, pudiendo comprobarse que el puntero del ratón presenta el símbolo Ӣ a su derecha (propio de los accesos directos). Una forma de forzar la copia, cuando se arrastre un archivo a otra carpeta del mismo disco o cuando se arrastre el icono de un programa, es pulsar la tecla [Ctrl] al tiempo que se arrastra el icono con el botón principal del ratón. Veremos entonces como el puntero presenta a su derecha el símbolo 🖻. De forma análoga, si al arrastrar pulsamos simultáneamente las teclas [May] [Ctrl] entonces forzamos el acceso directo. Finalmente, cuando el puntero del ratón tenga la forma \otimes , estará indicando que en el área sobre la cual estamos arrastrando un icono no es posible soltarlo.

Selección de múltiples objetos. La mayoría de las operaciones descritas para archivos o carpetas individuales son válidas para conjuntos de estos objetos. Para definir un conjunto podemos utilizar alguna de las estrategias que se describen a continuación:

- dibujar, pulsando el botón principal del ratón, un cuadrado donde se englobe el conjunto de objetos deseado
- seleccionar el primer objeto de una lista, pulsar la tecla [**May**] y seleccionar el último de la lista. Todos los objetos intermedios aparecerán también seleccionados.
- seleccionar uno a uno los objetos deseados manteniendo pulsada la tecla [Ctrl] (si no pulsamos dicha tecla, al seleccionar un nuevo objeto se pierde la selección del antiguo)

La selección múltiple es posible en todas las ventanas salvo en la izquierda del explorador (la correspondiente al árbol de directorios).

GENERALIDADES

1. Generalidades sobre SPSS.

El paquete estadístico SPSS, responde al funcionamiento de todo programa que lleva a cabo análisis estadísticos: pasados los datos a analizar a un fichero con las características del programa, éste es analizado con una serie de órdenes, dando lugar a unos resultados de tipo estadístico que el investigador debe interpretar. Este camino a seguir es el que guiará la presentación de SPSS que efectúa este mini-manual, debido a que es el camino más natural de aprendizaje del mismo.

1.1. El procedimiento general de resolución de un problema con SPSS.

Los pasos a seguir para llevar a cabo un análisis de tipo estadístico son los siguientes: 1° Recoger la información del problema que se desee investigar y tenerla organizada generalmente en papel; 2° Grabar esa información en un archivo de datos correspondiente al programa que se va a usar, en el caso de SPSS en un archivo que tiene el nombre que deseemos, pero que posee forzosamente la extensión **.sav**; 3° Sobre tal archivo de datos se llevará a cabo el análisis con SPSS, usando diferentes procedimientos que como veremos se seleccionan de distintos menús; 4° Los resultados de tales análisis son volcados a un visor de resultados en el que su visualización y edición es más cómoda, y desde el que se pueden guardar en un fichero de nombre el que se desee, pero de extensión **.spo** . 5° El investigador interpreta los resultados y extrae las conclusiones que le parecen relevantes de los mismos, y con eso se cierra el ciclo sobre el que estamos trabajando.

El procedimiento que acabamos de especificar es el que se sigue siempre en este tipo de análisis, y es el que subyace a toda la presentación que aquí se hace, por tanto empezaremos a entrar en detalle de los pasos para llevarlo a cabo.

1.2. La ventana principal de SPSS: el Editor de datos de SPSS.

El paquete SPSS, desde la versión 7, es un paquete adaptado al entorno WINDOWS con lo cual la forma de ejecutarlo es a través de ventanas en las que se despliegan menús, de los que se pueden elegir distintas opciones y así sucesivamente; por tanto es a través de un entorno de tipo gráfico desde donde se solventan los problemas, y no a través de comandos (aunque también se puede hacer así) como se hacía antes en los paquetes estadísticos más usuales.

Por lo que acabamos de decir, la forma de iniciar la ejecución del programa SPSS es pinchando dos veces con el ratón ("pinchar" lo utilizaremos como sinónimo de "hacer clic" con el botón principal del ratón) en el icono de SPSS que es como el de la Figura 1.1, y que se suele



Figura 1.1 Icono del programa SPSS.

encontrar en el escritorio en forma de enlace o en el menú de inicio dentro del apartado de programas. Una de las primeras tareas que tendrá que hacer el usuario de SPSS será localizar la posición del icono y adaptarlo a su gusto para que la entrada al programa sea fácil.

Cuando se ha pinchado dos veces sobre el icono, se abre la ventana principal de SPSS que es el **Editor de datos de SPSS**, aunque también la llamaremos ventana principal de SPSS. Esta ventana tiene dos versiones o vistas: vista de datos y vista de variables. En la figura 1.2 podemos ver la **vista de datos**; en ella aparecen ya incorporados los datos del fichero de datos **osteo.sav** (que es el fichero que nos acompañará a lo largo de todo el manual; ver el Apartado 2.1). En la figura 1.3 aparece la **vista de variables**; en ella se muestran las características de todas las variables del fichero de datos. De una vista a otra se cambia pinchando con el ratón en la pestaña correspondiente en la parte inferior izquierda de la ventana Dentro de la vista de datos se pueden distinguir varias zonas. La primera zona (parte más alta de la ventana,

🛅 oste	o.sav - Ed	litor de da	tos SPSS								×
Archivo	Edición	Ver Datos	Transform	nar Analizar	Gráficos U	tilidades	Ventana ?				
	26]] 46] 15] 15] 15] 15] 15] 15] 15] 15] 15] 15										
	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida]-
1	num	Numérico	3	0	Número del	Ninguno	Ninguno	6	Derecha	Escala	
2	edad	Numérico	2	0	Edad del pa	Ninguno	Ninguno	7	Derecha	Escala	
3	sexo	Numérico	1	0	Sexo del pa	{1, Varón}	Ninguno	6	Derecha	Escala	
4	peso	Numérico	6	1	Peso del pa	Ninguno	Ninguno	7	Derecha	Escala	
5	talla	Numérico	3	0	Talla del pac	Ninguno	Ninguno	6	Derecha	Escala	
6	imc	Numérico	5	1	Índice de ma	Ninguno	Ninguno	7	Derecha	Escala	-
7	tevol	Numérico	2	0	<u>Tiempo de e</u>	Ninguno	Ninguno	7	Derecha	Escala	┛
▲ ► \ Vis	sta de datos j	λVista de varia	ables /		4					•	١
				SPSS El p	rocesador lestá pre	eparado					

Figura 1.3 Ventana principal de SPSS (vista de variables).

🛅 osteo	.sav - E	ditor de	datos S	PSS								×
Archivo	Edición	Ver Dat	tos Tran	sformar	Analizar	Gráficos	Utilidades	Ventana	?			
	ð 🖳 🖌		🏪 😰 d	M <u>*</u>		\$ 0						
1 : peso		72	1,3									
	num	edad	sexo	peso	talla	imc	tevol	tabaco	alcohol	ingca	acfis	
1	1	23	1	72,3	175	23,6	2	1	2	1	1	
2	2	35	1	82,5	173	27,6	4	2	2	1	2	2
3	3	51	1	61,0	170	21,0	17	1	2	1	1	T
4	4	44	1	67,0	167	25,0	17	1	2	2	1	
5	5	54	2	51,0	168	18,1	30	1	2	1	1	
6	6	36	2	61,5	158	24,7	2	1	2	1	1	-
▲ ► \Vist	ta de datos	s 🖌 Vista de v	/ariables			•						•
					SPSS El pro	ocesador está	preparado					- //.

Figura 1.2 Ventana principal de SPSS (vista de datos).

generalmente en color azul) está formada por la barra que contiene el nombre de la ventana, incluyendo el nombre del fichero de datos activo si existe, en nuestro caso el **osteo.sav**. La segunda zona a destacar (justo debajo de la anterior) es la zona de los menús en la que aparecen los nombres de los menús desplegables que nos servirán para llevar a cabo nuestras tareas; cuando se coloca el cursor sobre uno de los rótulos y se pincha con el ratón se despliega un menú sobre el que desplazando el ratón se remarcarán las acciones que se pueden ejecutar y de la que nosotros escogeremos una; los menús que se pueden desplegar figuran en la siguiente tabla con una breve descripción de lo que se puede hacer con ellos, descripción que corresponde en muchos casos con la de cualquier programa Windows que presenta tales menús.

Menú	Función
Ar <u>c</u> hivo	Todas las funciones que se pueden hacer con archivos: Abrir, cerrar,
	guardar, importar, exportar, imprimir, etc.
<u>E</u> dición	Realiza todas las funciones típicas de la edición como son: cortar,
	copiar, eliminar, buscar, reemplazar, etc
<u>V</u> er	Permite controlar la forma en la que se ve la pantalla principal,
	controlando las barras que aparecen en ella así como la forma en la
	que se presentan los datos.
<u>D</u> atos	Contiene el conjunto de acciones que se pueden llevar a cabo con los
	datos: definir propiedades de las variables, seleccionar casos,
	ordenar casos y muchas más.
<u>T</u> ransformar	Permite realizar cualquier función conducente a crear nuevas
	variables a partir de otras existentes o no: transformar, recodificar,
	asignar rangos, etc
<u>A</u> nalizar	Acceso al conjunto de programas de SPSS, que van desde la
	generación de una tabla de frecuencias a análisis multivariantes
	complejos.
<u>G</u> ráficos	Acceso al conjunto de gráficos estadísticos que van desde un simple
	histograma al dibujo de una curva ROC.
<u>U</u> tilidades	Acceso a la descripción de las variables del fichero activo, creación
	de grupos de variables, así como edición de los menús.
Ve <u>n</u> tana	Acceso rápido a las ventanas de datos, de resultados, de sintaxis.
<u>?</u>	Ayuda en línea sobre todo el paquete SPSS.

Tabla 1.1. Menús de la ventana principal de SPSS

Como se puede observar en la tabla, bajo una de las letras del rótulo del menú, aparece un subrayado, indicando tal cosa que combinando la tecla Alt con la letra subrayada se puede desplegar el menú correspondiente, pudiendo desplazarse uno por él sin más que usar las teclas de flecha presentes en el teclado.

Abrir archivo	? ×
Buscarien: 🔄 Spss	💿 🖻 💆 📑 🗐
C Acrobat	Alquileres comerciales
acroread	Amor 🗧
Config	🛗 Ansiedad 2
🗀 Looks	🛅 Ansiedad
Ddbc	🛗 Coches
C Scripts	🛅 Datos arteria coronaria
	Þ
Nombre de archivo:	Abrir
<u>T</u> ipo de archivos: SPSS (*.sav)	Pegar
	Cancelar

Muchas de las opciones de estos menús las vamos a ir estudiando y utilizando a lo largo de este manual. Comentaremos aquí sólo algunas opciones básicas de uso general. La opción **Abrir**, del menú **Archivo**, nos permite abrir un archivo (generalmente de datos) para empezar a

trabajar con él. Hoy día, esta opción es común (y funciona igual) en casi todos los programas Windows. Al pinchar sobre ella se abre una ventana como la de la Figura 1.4. En la primera caja de esta ventana (justo a la derecha de **Buscar en**) se coloca la carpeta donde está el archivo que queremos abrir; pinchando en la pestaña $\mathbf{\nabla}$ podemos acceder al árbol de carpetas y seleccionar la que corresponda (por defecto aparece la carpeta SPSS). En la última caja y pulsando la pestaña $\mathbf{\nabla}$, podemos seleccionar el tipo de archivo que queremos abrir (por defecto aparece archivos de datos de SPSS, esto es, con extensión **.sav**); entonces en la segunda caja (la más grande de la ventana) aparecerá una lista con todos los archivos del tipo seleccionado que existen en la carpeta seleccionada, y de ahí seleccionaremos el que deseemos abrir, que aparecerá escrito en la tercera caja de la ventana. Finalmente, pulsando el botón **Abrir**, se abrirá el archivo. Éste también se puede abrir haciendo doble clic sobre su nombre en la segunda caja.

También en el menú **Archivo** tenemos dos opciones para guardar el fichero de datos activo. La opción **Guardar** permite almacenar el fichero activo en un disco. Si el fichero activo ha sido leído previamente se guardará con el mismo nombre que tenía (el fichero original que existía en el disco se perderá). Por el contrario si el fichero ha sido creado sin que exista ninguna imagen de él en el disco, se nos pedirá que demos un nombre al nuevo fichero en el que se va a guardar la información. Debe quedar claro que esta opción siempre guarda un fichero de datos de SPSS, es decir, aquél que tiene una extensión **.sav.** Por el contrario la opción **Guardar como...** permite guardar el fichero activo con otro nombre y/o con formato de otras aplicaciones informáticas, como bases de datos u hojas de cálculo.

base	S	de	datos	u	nojas
: Guard	lar da	tos como)		? 🔀
Guardar e	n: 🔁 9	PSS	•	· 🗢 🖻 🖻	* 💷 •
🛅 Alqu	ileres	comercial	es.sav		🛅 carpet.s
Amo	r.sav				Coches.
anore	ectic.sa	av			🗰 coffee.s
Ansi Ansi	edad 2.	sav			🔳 Datos at
🔳 Ansi	edad.sa	Ŵ			🔳 Datos d
1					
					<u> </u>
	Lonser	vando 24 de 24	4 variables.		Variables
Nombre:					Guardar
Tipo:	SPSS	(*.sav)			- Pegar
	🔽 Esc	ribir nombres d	le variable en hoia de	cálculo	Cancelar
	🗖 Gua	ardar etiquetas	de valor donde se haj	yan definido en	vez de valores de dato
	🗖 Gua	ardar etiquetas	de valor en un archiv	0.888	

Cuando se selecciona la opción se abre una ventana como la presente en la Figura 1.5, en la que se seleccionará la carpeta donde se va a guardar el archivo, se escribirá el nombre del archivo donde se desea guardar la información, se seleccionará el tipo de archivo deseado y se pulsará el botón **Guardar**.

La opción Imprimir, del menú **Archivo**, nos permite sacar por impresora el fichero activo (todo o una parte de él) y funciona de forma similar a otras aplicaciones Windows. Finalmente, la opción **Deshacer**, del menú **Edición**, nos permite deshacer, si es posible, la última acción que hemos realizado.

La tercera zona de la ventana principal de SPSS, que está formada por iconos, es la que se denomina barra de herramientas y en ella aparecen los iconos que representan acciones muy comunes que se pueden llevar a cabo con SPSS (y de las que se desea un acceso más rápido); cuando uno pasea el cursor por encima de los iconos se despliega un texto que indica la acción que lleva a cabo. Los iconos aparecerán resaltados o no según estén o no accesibles en un momento determinado. La barra de herramientas puede configurarse y añadir o eliminar iconos en ella, pero los más comunes son los que aparecen en la Figura 1.2 y que serán descritos cuando se utilicen por primera vez. Por ahora, sólo mencionar que los iconos 1, 2, 3 y 5 realizan, respectivamente, las funciones básicas de abrir, guardar, imprimir y deshacer, ya descritas en los párrafos anteriores.

La cuarta zona es la línea de edición de datos en la que figura información de la casilla de datos seleccionada; en la Figura 1.2 aparece seleccionada la información correspondiente a la variable "peso" del primer caso, que tiene el valor 72.3; en esa zona aparece el contenido que se le está dando a una variable para un caso concreto cuando se están grabando datos o cuando se están modificando éstos.

La quinta zona, que es la que denominaremos <u>matriz de datos</u>, tiene forma de tabla con filas y columnas, y representa al fichero de datos activo. Las filas representan generalmente los casos (individuos) de nuestra muestra, y las columnas las diferentes características (variables) de esos individuos que son objeto de nuestro estudio. Así, en la Figura 1.2 se puede ver que algunas de las variables que hemos grabado para cada uno de los individuos de nuestra muestra, son num, edad, sexo, peso, talla, imc, etc.; además se puede ver que SPSS, en la primera columna de todas, asigna un número (empezando por el 1) a cada fila o individuo: es el número de registro. Así podemos ver que para el paciente número 2 el valor de la talla es de 173 cm.

Tanto en la parte baja de la pantalla como en el lateral derecho de la misma aparecen dos barras desplazadoras que nos permiten movernos hasta abarcar todas las variables del fichero activo (en el caso de la barra en la parte inferior de la pantalla) o para abarcar todos los casos del fichero activo (en el caso de la barra de la derecha de la pantalla). Por último, debajo de la barra desplazadora inferior, aparece la barra de estado en la que se nos informa del estado de ejecución de la tarea que se ha encomendado a SPSS.

En la vista de variables se cambia la matriz de datos por una descripción detallada de las características (nombre, tipo de variable, anchura, etiqueta, etc.) de cada una de las variables del fichero de datos activo. En esta vista podremos cambiar cualquiera de estas características, pero esto lo veremos más adelante.



1.3. Normas de funcionamiento de las ventanas de SPSS.

Para ejecutar cualquier acción con SPSS se empieza siempre escogiendo tal acción de un menú (o de un submenú, en su caso) (véase la Figura 1.6) que nos lleva a la ventana correspondiente a la acción o procedimiento elegido, en la que hay que elegir los componentes fundamentales de la acción y en la que

hay la posibilidad de elegir otros componentes accesorios, como se puede ver en la Figura 1.7. Esta ventana tiene una estructura como la que sigue: en la parte izquierda suele haber una caja en la que aparecen todas las variables presentes en el fichero activo y que son susceptibles de ser manejadas con el procedimiento elegido; a la derecha de esa caja suele haber otra u otras cajas en las que pondremos las variables seleccionadas para la acción que deseamos llevar a cabo; la forma de establecer tal selección es pinchar, en la caja de la izquierda, la variable de interés (inmediatamente aparecerá remarcada) y presionar en un botón con una flecha que señala hacia la derecha (►), lo que llevará la variable a la caja de las variables seleccionadas para el análisis. También suele haber unos botones de particularización, o en la parte baja de la ventana o en la parte derecha de la misma (en el caso de la Figura 1.7 están en la parte baja) que particularizan el conjunto de acciones que dentro del procedimiento elegido se pueden llevar a cabo: en el caso de la Figura 1.7, las acciones a llevar a cabo son: elegir los estadísticos que aparecerán junto con las tablas de frecuencias, elegir los gráficos que acompañarán a las tablas de frecuencias y por último elegir el formato en que aparecerán las tablas de frecuencias; todos estos botones dan lugar a ventanas en las que se harán las elecciones que hemos señalado. También, en las ventanas de SPSS aparecen cinco botones de acciones inmediatas, que suelen estar a la derecha de la ventana en vertical. Tales botones aparecen en la tabla 1.2 y serán explicados en ella para no tener que hacerlo más.

Por último, si se observa la Figura 1.7 se puede ver en ella que hay un pequeño cuadradito en el que figura una marca; tal marca puede estar presente o no, significando su presencia que la acción que aparece está seleccionada y se llevará a cabo, mientras que su ausencia indicará que no se llevará a cabo tal acción. Esta forma de seleccionar las acciones secundarias a llevar a cabo es muy frecuente en SPSS.

Hasta aquí las explicaciones generales sobre las ventanas de SPSS, aunque cuando se estudien los procedimientos concretos estudiaremos las ventana oportunas con todo detalle.

Botón	Función
Aceptar	Llevar a cabo el análisis seleccionado.
Pegar	Escribir en el editor de sintaxis los comandos correspondientes a las
_	acciones elegidas en la ventana (no lo veremos en este manual).
Restablecer	Borrar todas las elecciones hechas en la ventana para poder elegir
	otras nuevas.
Cancelar	Salir de la ventana, dejándola como está, sin llevar a cabo ninguna
	acción.
Ayuda	Desplegar la ayuda contextual del conjunto de acciones que aparecen
	en la ventana.

Tabla 1.2 Botones presentes en las ventanas de los distintos procedimientos de SPSS

1.4. Uso de los botones del ratón.

Como en todo programa de Windows el ratón o puntero es un accesorio fundamental en el manejo de SPSS. La selección de cualquier acción se puede hacer con el ratón desplazándolo a lo largo de la ventana y pulsando el botón izquierdo del mismo sobre el símbolo o el texto que representa la acción.

El botón derecho del ratón también tiene un uso importante aquí. Más allá de usos particulares (que se explicarán en su lugar), pulsar el botón derecho sobre cualquier icono o texto desplegará una leyenda de ayuda contextual sobre la acción seleccionada; la leyenda aparecerá sobre la ventana en un recuadro, generalmente amarillo. Se podrá hacer desaparecer dicha leyenda sin más que pulsar el ratón en otra opción o moverlo convenientemente. Pulsar el botón derecho dentro de la matriz de datos (por ejemplo sobre el nombre de una variable) despliega un menú contextual que contiene acciones ligadas a los objetos señalados.

2. Cómo está organizado el manual.

Este manual pretende enseñar SPSS de una manera paralela a como se hace el análisis de unos datos; por eso su organización será la de los pasos que se siguen para llevar a cabo tal análisis. Los apartados que se presentarán primero serán todos aquellos referentes al manejo de

datos (apartados 3 al 6); en ellos se incluye todo lo referente a creación de variables, grabación y edición de datos, operaciones con variables y selección y ordenación de casos. Después aparecerán los capítulos referentes al análisis estadístico propiamente dicho (desde el apartado 7 en adelante); en ellos se incluirá desde la Estadística Descriptiva básica, con sus gráficos correspondientes, a los tests de comparación de medias, los tests Chi-cuadrado y problemas de Regresión y Correlación Lineal Simple .

Debe quedar claro que este no es un manual de Estadística, sino de un paquete de programas, el paquete SPSS, que hace los cálculos estadísticos; por ello siempre se explicará el funcionamiento del programa y no los conceptos estadísticos que subyacen al mismo y a los resultados que éste genera; por ello se recomienda al lector que estudie de manera previa un manual de estadística, ya que aquí no encontrará las explicaciones que son propias de aquel.

En cuanto a la presentación de las explicaciones y de las indicaciones, sólo expondremos que cuando queramos indicar de manera breve el camino a seguir para llevar a cabo una acción cualquiera, seleccionando en secuencia objetos de SPSS (menús, ventanas, etc.), se presentarán los nombres de los objetos de SPSS en cursiva separados por flechas (\rightarrow); así la indicación *Analizar* \rightarrow *Frecuencias* \rightarrow *Gráficos....*, indicará que para seleccionar la acción que deseemos habrá que desplegar el menú Analizar, de él elegir la opción Frecuencias y de esa ventana elegir el botón Gráficos, eligiendo de allí las opciones de interés que no hemos expuesto en este caso (esto lo hemos representado por los puntos suspensivos). A veces observará el lector que en la secuencia que se presenta aparece una barra (/); con ello queremos indicar que se pueden elegir varias opciones que pueden ser excluyentes o simultáneas, según aclarará el contexto en el que estemos trabajando.

La forma que hemos elegido de explicar el paquete SPSS obligará en muchas ocasiones a pasar por el mismo menú para fijarnos en características distintas de él (si bien siempre que se acceda por primera vez a tal menú se explicará detalladamente en la máxima extensión posible). Además el lector debe acostumbrarse a ver características en los menús que no son abordadas en el manual; si así lo hiciéramos este se convertiría en un manual de gran extensión que desborda los objetivos de esta introducción.

Descripción de la variable	Nombre de la variable
Número del paciente	NUM
Sexo(1= varón;2= mujer)	SEXO
Edad (años)	EDAD
Peso (Kgr)	PESO
Talla (cm)	TALLA
Índice de masa corporal	IMC
Tiempo de evolución de la enfermedad (años)	TEVOL
Consumo de tabaco (1=si; 2=no)	TABACO
Consumo de alcohol (1=no; 2=sí, pero no en exceso; 3=sí, en	ALCOHOL
exceso)	
Ingesta de calcio (1=suficiente; 2=insuficiente)	INGCA
Actividad física (1=sí; 2=no)	ACFIS
Presencia de retinopatía (1=no; 2=leve; 3=grave)	RETIN
Presencia de nefropatía (1=no; 2=leve; 3=grave)	NEFRO
Presencia de neuropatía (1=no; 2=leve; 3=grave)	NEURO
Hemoglobina glicoxilada	HBA1C
Calcio	CA
Fósforo	Р
Creatinina	CR
Pth media	PTHM
Pth intacta	TPI
Densidad de masa ósea en el cuello del fémur	BMDCUE
Valor tipificado de la densidad de masa ósea en L24	SZL24
Valor tipificado de la densidad de masa ósea en triángulo de Ward	SZTRI
Valor tipificado de la densidad de masa ósea en el cuello del fémur	SZCUE

Tabla 2.1 Variables del fichero osteo. sav

En todas las situaciones la explicación se acompañará de un ejemplo que nos ayudará a explicar de manera aplicada los detalles de los procedimientos que estemos estudiando, tal ejemplo es el que se expone a continuación.

2.1. Un ejemplo que nos acompañará a lo largo del manual.

El ejemplo que emplearemos tiene sus datos grabados en un fichero denominado **osteo.sav**; en este fichero aparecen los datos referentes a un estudio acerca de la densidad mineral ósea en pacientes diabéticos insulin-dependientes. La descripción de las variables se muestra en la Tabla 2.1, que pretende ser suficientemente explicativa como para que el lector entienda la información grabada allí. El lector debería familiarizarse con las variables presentes en el fichero, pues se hará continua referencia en el texto a tales variables en cada uno de los apartados en los que se estudie.

FASE DE GENERACIÓN Y MANEJO DE DATOS

3. Creación de un fichero de datos.

Como se ha venido diciendo ya, el programa SPSS trabaja con los datos previamente grabados en un fichero al que denominamos fichero de datos y que ya sabemos que tiene la extensión obligatoria **.sav**. Cuando se ejecuta el programa SPSS tal fichero deberá estar activo (esto es, estar presente en la **vista de datos** del **Editor de datos SPSS**) para que sobre él se lleven a cabo los cálculos oportunos. Esto se conseguirá o bien leyendo un archivo de datos previamente grabado o bien creando uno en el propio **Editor de datos SPSS**.

La creación de un fichero de datos tiene dos pasos fundamentales: la definición de las variables y la grabación de los datos. Lo común, y así será siempre en este manual, es que para todos los individuos se disponga de información sobre las mismas variables, aunque, ocasionalmente, pueda faltar en algunos individuos información de alguna(s) variable(s); la matriz de datos que obtendríamos así se denomina matriz rectangular (fichero rectangular). Empecemos definiendo las variables de un nuevo fichero de datos, después veremos la grabación de los datos.

3.1. Definición de las variables.

Cuando se abre SPSS aparece la matriz de datos vacía indicando en ese caso que no hay ningún fichero seleccionado para trabajar con él; es en esta situación en la que debemos estar para proceder a crear la estructura (definir las variables) de un nuevo fichero de datos de SPSS. A esa situación también se puede llegar después de haber trabajado con SPSS sin más que desplegar el menú Archivo y elegir de él la opción Nuevo y dentro de ella datos (*Archivo* \rightarrow *Nuevo* \rightarrow *Datos*), lo que hará que se elimine el fichero activo y se deje limpia la matriz de datos. Ahora cada columna será una nueva variable en potencia hasta que sea definida, en cuyo caso pasará a ser una variable en la nueva matriz de datos.

Una variable queda automáticamente definida si en alguna casilla de su columna se introduce algún dato: SPSS define una variable adecuada al primer dato introducido. A estas variables, SPSS les asigna por defecto un nombre constituido por las letras VAR seguidas de un código de cinco dígitos (00001, para la primera; 00002 para la segunda; ...). Siempre podremos cambiar después el nombre y las características de la(s) variable(s) así definida(s).

De todas formas, la manera más natural de crear las variables de un fichero es definirlas antes de introducir dato alguno. Para ello hemos de situarnos en la vista de variables de la ventana principal de SPSS, pinchando en la pestaña correspondiente o haciendo doble clic en la cabecera de la columna. En la vista de variables, cada fila corresponde a una variable, y las columnas representan las características de la(s) variable(s). Para cada variable habrá que ir especificando cada una de sus características, empezando por el nombre (primera columna).

3.1.1. El nombre de las variables.

Se pincha (o se hace doble clic) sobre la casilla correspondiente al nombre de la variable que se está definiendo, y se escribe el nombre que deseemos que tenga, con un máximo de 64 caracteres, sin espacios en blanco, empezando siempre por una letra o con el carácter @ y no terminando con un punto. Los caracteres que formen el nombre deben ser: cualquier letra o dígito, o los signos #.@\$ (éste último no es recomendable como último carácter).

3.1.2. El tipo de las variables.

Tipo de variable			? ×		
Numérica		_	Aceptar		
C Punto	Anchura: Cifras decimales:	Anchura:	Anchura:	8	Cancelar
O Notación científica		14	Ayuda		
C Fecha					
O Dólar					
O Moneda personalizada					
🗢 Cadena					

Una vez escrito el

nombre de la variable, si se pincha en la casilla correspondiente al tipo de variable (o en cualquier otro sitio), aparecerá, para dicha variable, "numérico" en la columna Tipo, un 8 en la columna Anchura y un 2 en la columna Decimales. Esto es lo que SPSS ofrece por defecto para esa variable. Para cambiar el tipo de la variable se pincha en la casilla de tipo de variable y se pincha sobre el pequeño recuadro gris situado en dicha casilla, aparece entonces una ventana como la de la Figura 3.1; en ella se puede definir el tipo de la variable de entre los que aparecen: numérico, cadena, fecha, etc.... El formato de la ventana es muy simple, a la izquierda aparecen los diferentes tipos que pueden tener las variables y a la derecha aparece el tamaño o características particulares de cada tipo que también son definibles; así en la figura 3.1, el tipo es numérico y la variable tendrá una anchura total de 8 caracteres, dos de los cuales son decimales: esto es lo que por defecto coloca SPSS. Pero describamos los diferentes tipos.

El tipo Numérico es para una variable cuyos valores son números. Los valores se muestran en el formato numérico estándar, utilizando como delimitador decimal el especificado en la Configuración Regional del Panel de control de Windows. El tipo Coma define una variable numérica cuyos valores se muestran con la coma de separación de miles y con un punto como separador de la parte decimal. El tipo **Punto** define una variable numérica cuvos valores se muestran con el punto de separador de miles y con una coma como separador de la parte decimal. El tipo Notación científica define una variable numérica cuyos valores se muestran con una E intercalada y un exponente con signo que representa una potencia de base diez. El exponente puede ir precedido de E o D con un signo opcional, o por el signo solamente. Por ejemplo, 123, 1,23E2, 1,23D2, 1,23E+2, o incluso 1,23+2. El tipo Fecha define una variable numérica cuyos valores se muestran en uno de los diferentes formatos de fecha-calendario u hora-reloj. Para ello, habrá que seleccionar un tipo de fecha de la lista desplegable que aparece a la derecha cuando se selecciona el tipo fecha. Se pueden introducir las fechas utilizando como delimitadores: barras, guiones, puntos, comas o espacios en blanco. El tipo Dólar sirve para definir una variable numérica (con diferentes formatos) cuyos valores contienen un signo de dólar y una coma para la separación de los miles. Moneda personalizada sirve para definir una variable numérica cuyos valores se muestran en uno de los formatos de moneda personalizados que se hayan definido previamente en la pestaña Moneda del cuadro de diálogo Opciones dentro del menú Edición. Los caracteres definidos en la moneda personalizada no pueden emplearse para la introducción de datos, pero sí los mostrará el Editor de Datos. Por último el

tipo **Cadena** define una variable cuyos valores no son numéricos y, por ello, no se utilizan en los cálculos. Pueden contener cualesquiera caracteres hasta la longitud definida. Estas variables son conocidas como variables alfanuméricas.

Como hemos dicho antes, el tamaño y el formato de cada tipo se expresa en los campos que aparecen en la parte de la derecha de la ventana. Habrá que especificar el tamaño total y el número de decimales en los tipos Numérico, Coma, Punto y Notación Científica, colocándose sobre los huecos oportunos y escribiendo ahí un número para el tamaño global y otro para el número de decimales. La anchura máxima en todos ellos es de 40 posiciones, siendo 16 el número máximo de cifras decimales. Para el tipo Cadena habrá que especificar la anchura total que no podrá sobrepasar los 255 caracteres. Para los tipos de Fecha, Dólar y de Moneda Personalizada habrá que elegir el formato de entre los que aparecen a la derecha en la ventana de definición de tipos.

Por último, la anchura y número de decimales se pueden cambiar igualmente en las columnas Anchura y Decimales de la vista de variables.

3.1.3. Las etiquetas de las variables.

Generalmente el nombre de las variables suele tener pocos caracteres, lo que hace que dicho nombre sea en muchos casos un acrónimo de difícil traducción para el que no lo ha escrito (así, no es fácil de adivinar que con el nombre ACR124 estoy representando a la variable aclaramiento de creatinina en el primer día de seguimiento en orina de 24h.); por ello se suele poner una etiqueta de variable a la variable, que la identifique de una manera más precisa y que nos permita reconocerla cuando se presentan los resultados. Esa etiqueta puede ser como máximo de 256 caracteres. Pero no sólo se puede poner una etiqueta general para la variable, sino que en las variables que sean cualitativas y cuyas categorías suelen estar representadas por códigos numéricos, pueden establecerse etiquetas de valor que permitan identificar a las categorías con ellas (haciéndolas más explicativas) en vez de con los códigos numéricos. Estas etiquetas pueden ser de hasta 60 caracteres. Estas etiquetas no son viables para variables de cadena de más de 8 caracteres de amplitud.

Etiquetas de valor			
Etiquetas de valor Valor: 3		Aceptar	
Etiqueta de valor: Gra	ive	Cancelar	
Añadir 1 = Cambiar Eliminar	"No" "Leve"	Ayuda	

La etiqueta de la variable se escribe directamente en la casilla correspondiente de la columna Etiqueta en la vista de variables. Para las etiquetas de valor hay que pinchar sobre la casilla correspondiente de la columna Valores (en la vista de variables) y luego pinchar sobre el pequeño recuadro gris que aparece en dicha casilla, se despliega entonces una ventana como la de la Figura 3.2. En esta ventana, la forma en que se asignan etiquetas a los códigos de las variables, es escribir un código en la zona en la que aparece "Valor" y la etiqueta correspondiente en la zona de "Etiqueta de valor" y después pulsar el botón de **Añadir**; así, en la Figura 3.2, al código 1 se le ha asignado la etiqueta "no", al código 2 se la ha asignado la etiqueta "moderado" y al código 3 se le añadirá la etiqueta "grave" en cuanto pulsemos el botón **Añadir**. El código al que no se le asigne etiqueta sorrespondientes son guardados para que queden a la vista del usuario en una zona en la que son susceptibles de ser seleccionados. Cuando se selecciona una de estas parejas "código = etiqueta" vuelve a ser colocada en la zona de etiquetas de valor asuma de estas parejas "código = etiqueta" vuelve a ser colocada en la zona de etiquetas de valor se una de proceder a su modificación; si después de hacer las

modificaciones oportunas se desea guardar la información bastará pulsar en **Cambiar**; si una vez seleccionada la pareja se desea borrar bastará pulsar el botón **Borrar** y la pareja será eliminada; el programa permite seleccionar una pareja cada vez. Cuando se hayan escrito todas las etiquetas bastará con pulsar el botón **Aceptar** y se volverá a la vista de variables.

Las etiquetas, tanto de las variables como de los códigos de las variables, no son obligatorias pero si son muy útiles a la hora de entender y presentar los resultados.

3.1.4. Los datos faltantes (valores perdidos).

Se produce un dato faltante cuando la información de una variable de un individuo no existe; ello puede ser debido a varias causas: a que tal información se ha perdido, a que por las características particulares del individuo no puede existir o a que el individuo no ha querido proporcionar dicha información. Cuando para una variable existen datos faltantes frecuentemente se escoge un código para representarlos, debiendo proporcionarle tal código a SPSS para que él no los incluya en los análisis posteriores; a esta representación de los datos faltantes se le denomina datos faltantes del usuario, para distinguirlos de los datos faltantes del sistema (que se consiguen sin más que dejar en blanco el espacio reservado para ellos, donde, si la variable es numérica, SPSS colocará una "coma" para identificarlos). No obstante, estas reglas tienen excepciones: en las variables de cadena no existen datos faltantes del sistema puesto que el blanco lo interpreta SPSS como un valor; tampoco se pueden definir datos faltantes del usuario en las variables de cadena de longitud superior a 8 caracteres. Es importante que el usuario sepa que la identificación de datos faltantes es crucial pues, si no se identifican, estos serán empleados con los valores que tengan, dando lugar a resultados erróneos. Los datos faltantes (valores perdidos) del usuario se especifican pinchando sobre la casilla correspondiente de la columna Perdidos en la vista de variables, y volviendo a pinchar sobre el pequeño recuadro gris que aparece en dicha casilla; se despliega entonces una ventana como la de la Figura 3.3 que comentaremos a continuación.

Valores perdidos	? 🗙
O No hay valores perdidos	Aceptar
Valores perdidos discretos	Cancelar
	Ayuda
C Rango más un valor perdido discreto opcio	nal
Mínimo: Máximo:	
Valor discreto:	

Figura 3.3 Ventana de definición de valores perdidos.

Como se ve hay tres formas de indicar la presencia de datos faltantes. La primera es la de indicar que no existen datos faltantes o perdidos empleando la opción **No hay valores perdidos;** esta es la opción que de oficio ofrece SPSS indicando que no hay datos faltantes por parte del usuario, pudiendo haber datos faltantes del sistema sin más que hayamos dejado en blanco el espacio oportuno. La segunda manera de identificar los valores que representan a los datos faltantes es presentarlos como **Valores perdidos discretos**, indicando con ello que son valores aislados los que representarán a los datos faltantes estén representados por valores que aparezcan seguidos, de manera que sea más fácil identificarlos a través de un máximo y un mínimo quedando caracterizados como datos faltantes todos los que están en medio. Esto es lo que permite hacer la última opción, **Rango más un valor perdido discreto** opcional, en la que además tenemos la posibilidad de indicar un valor perdido discreto adicional. Pulsando el botón **Aceptar** volveremos a la vista de variables.

3.1.5. El formato de columna de las variables.

Ya hemos visto como se hace la definición del tipo y la anchura de la variable; ahora bien, esas características son características internas que no tienen que corresponderse con las características que manejan su presentación en la matriz de datos. Para ello tenemos la opción de **Formato de columna**. Este formato de presentación de la columna tiene dos características: la anchura total de la columna y la alineación que tendrá la información dispuesta en esta columna. El Ancho de la columna puede ser de hasta 256 caracteres, y sólo afecta a la presentación del editor de datos, de manera que si el ancho asignado a una variable fuera inferior al necesario para un dato concreto de dicha variable, éste será guardado internamente, pero no será presentado correctamente en el editor de datos; si la variable es numérica, el dato aparecerá redondeado, si se puede, o en notación científica, si se puede, o relleno de asteriscos; si la variable es de cadena, el dato aparecerá truncado. De oficio SPSS asigna un ancho de columna igual a 8 caracteres.

La alineación del texto en la columna puede ser a la izquierda, centrada o a la derecha, siendo esta última la asignada por SPSS en el caso de variable numérica y a la izquierda en el caso de variable de cadena. Tanto el ancho de la columna como la alineación pueden ser cambiadas fácilmente en la vista de variables; para ello pulsar convenientemente sobre las flechitas que aparecen al pinchar en las casillas correspondientes de las columnas: Columnas y Alineación. El ancho de columna también se puede cambiar en la vista de datos, pinchando y arrastrando los bordes de la columna. La definición de las características de las columnas no son obligatorias, y en muchos casos nos bastaremos con las definiciones que SPSS hace de oficio.

3.1.6. La definición de la Escala de Medida de una variable.

La escala de medida de una variable determina las operaciones que se pueden realizar con ella. Así, una variable como el peso será susceptible de ser sometida a todas las operaciones, por lo que se podrán hacer muchos análisis estadísticos con ella; sin embargo, una variable como el grupo sanguíneo ABO no será susceptible de ser sometida a operaciones de cálculo de media, por ejemplo, ya que es claramente un carácter cualitativo.

Para definir la escala de medida de una variable en SPSS se pincha en la casilla correspondiente de la columna Medidas en la vista de variables, y se pulsa sobre la flechita que aparece, eligiendo la opción que se desee. En general tendremos tres opciones excluyentes: Escala, Ordinal v Nominal. En Escala los valores de los datos son considerados como valores numéricos sobre una escala de intervalo o de razón (por ejemplo, la edad, el peso, el número de hermanos); cuando definimos una variable de tipo Numérico, Coma, Punto o Notación Científica, SPSS asigna Escala a la escala de medida de la variable. En el caso de Ordinal, los valores de los datos representan categorías con algún orden intrínseco (por ejemplo, bajo, medio, alto; peor, igual, mejor); las variables ordinales pueden ser cadenas (alfanuméricas) o valores numéricos que representen categorías diferentes (por ejemplo, 1=bajo, 2=medio, 3=alto); la escala ordinal corresponde a datos cualitativos ordinales. En la escala Nominal, los valores de los datos representan categorías sin un orden intrínseco (por ejemplo, el grupo sanguíneo ABO, el tipo de trabajo de una persona); las variables nominales pueden ser cadenas (alfanuméricas) o valores numéricos que representen categorías diferentes (por ejemplo, 1= varón, 2= mujer). De la definición de la escala de medida depende, como hemos dicho previamente, los análisis que se pueden hacer con los diferentes datos.

3.1.7. Estructura del fichero de datos.

Con todo lo visto hasta aquí podemos definir las variables en SPSS; la definición cuidadosa y detallada de las variables ayudará mucho en los análisis y en la interpretación de los resultados, por lo que recomendamos al usuario que gaste el tiempo necesario en tales definiciones (pues no será tiempo perdido). Cada variable definida aparece en la vista de datos con su nombre en la cabecera de la columna y se puede seleccionar pulsando el botón izquierdo del ratón sobre ella una vez.

El conjunto de variables definidas, junto con las características que les hayamos asignado, forman lo que podemos llamar la **estructura del fichero de datos**; esto sería una de

las dos partes de un fichero de datos de SPSS. Esta estructura se puede guardar en un fichero **.sav** que aparecerá sin los datos, pero donde se han guardado las variables y sus características, y la podemos ver resumida en una ventana mediante **Utilidades**—**Variables...** Sobre dicha estructura se puede añadir el otro ingrediente, los datos propiamente dichos, y juntos conformarán el archivo de datos de SPSS. Para guardar el trabajo actual, efectuar **Archivo**—**Guardar** (ó usar el botón "guardar") asignando el nombre deseado, por ejemplo, prueba.sav.

3.2 Grabación de los datos.

Una vez que el usuario ha definido las variables de su fichero de datos, tendrá una rejilla en la que por columnas aparecerán dichas variables con el nombre de cada una de ellas. Es claro que tal rejilla estará vacía y que el siguiente paso será el rellenarla a base de grabar la información de los diferentes casos (que aparecen en el Editor de Datos por filas). Por ello lo primero que explicaremos es cómo se graban los datos. Desde luego que antes de la grabación de casos se debería conocer cómo se mueve uno dentro de la rejilla de datos. El movimiento más simple es desplazar el cursor mediante el ratón a la posición deseada y pinchar en ella con el botón izquierdo. Otras formas más lentas es usando las teclas de flechas en las cuatro direcciones que funcionan como ellas indican.

3.2.1. Introducción de los datos por casillas, por filas y por columnas.

La primera forma en que se puede grabar un dato es situarse en una casilla (cruce de una variable, columna, con un caso, fila) pinchando en ella con el botón izquierdo del ratón y escribir el valor deseado seguido de la tecla **Intro.** Tras ello, el cursor aparecerá en la casilla inmediatamente debajo. Evidentemente esto nos permitirá la grabación de los datos por columnas (o variables): cuando lleguemos al final de una columna pasamos al principio de la siguiente. Sin embargo, esta forma de grabar los datos, no es la más común para proceder a la grabación de un fichero de datos.

La forma más común es grabar los datos por filas, es decir, por casos. Para ello se coloca el cursor en la casilla más a la izquierda de una fila (de un caso) y se introduce el dato correspondiente a esa casilla, seguido de la pulsación de la tecla de flecha hacia la derecha, o pulsando la tecla **Tab**. Obsérvese que si se pulsa la tecla Intro el cursor bajará una fila abajo con lo que estará fuera del caso, habiendo que volver a él mediante la tecla hacia arriba, lo que alarga claramente la grabación. Cuando uno haya grabado un caso, puede saltar al siguiente pulsando la tecla **Tab**, y así aparecerá al inicio del nuevo caso. SPSS identifica cada caso con un número que es interno y que no existe como variable del fichero, por lo que si el usuario desea una identificación para cada caso debe crear, y grabar, una variable a tal efecto, lo cual es conveniente.

Desde luego que cuando uno está grabando datos puede equivocarse, pudiendo, en cualquier momento, proceder a la corrección de los errores colocándose sobre ellos (en la celda en la que esté el error), eliminándolos (con la tecla **Supr** o con la tecla **\leftarrow BackSpace**) y reescribiendo los datos correctamente. Ni que decir tiene que el posicionamiento sobre el error se puede hacer de las múltiples maneras que hemos explicado anteriormente.

Siguiendo el procedimiento que hemos explicado, y con un poco de paciencia, se conseguirá grabar todo el fichero de datos. No obstante, hay algunos pequeños atajos que permiten hacer algo menos tedioso el proceso o que permiten la corrección pronta de errores que se repiten; a ello dedicaremos la siguiente sección.

3.2.2. Acciones de Cortar, Copiar y Pegar.

En la grabación de datos podemos utilizar las de acciones (familiarizadas por Windows) de **Cortar, Copiar y Pegar.** Estas tres acciones se basan en la selección del objeto sobre el que van a ser aplicadas. En el caso que nos ocupa el objeto puede ser una casilla, varias casillas, una fila, varias filas, una columna o varias columnas; una casilla se selecciona sin más que colocarse en ella (pulsando el botón izquierdo del ratón sobre ella o mediante una serie de

desplazamientos con las flechas); varias casillas se seleccionan sin más que pinchar con el ratón en una de ellas y arrastrar el mismo (moverlo con el botón izquierdo pulsado) hasta la última casilla que deseemos seleccionar, apareciendo estas remarcadas para indicar que están seleccionadas; es claro que el arrastrado que hemos mencionado puede ser en cualquier sentido por lo que el conjunto de casillas seleccionadas puede ser cualquiera siempre que sea en forma rectangular. La forma de seleccionar una fila (caso) es pulsar con el ratón sobre el número de caso que le ha asignado SPSS, mientras que si se desean seleccionar varios casos consecutivos deberá arrastrarse el ratón sobre ellos. De manera análoga, la selección de una variable se hará como en los casos pero pulsando sobre la fila de nombres de las variables. Resumiendo, supongamos que hemos seleccionado el objeto de nuestras acciones y pasemos a describirlas.

La acción de **Cortar** hace desaparecer el objeto seleccionado que es guardado en el Portapapeles para estar disponible en una acción de Pegado. La acción de **Copiar** mantiene el objeto seleccionado a nuestra vista pero es simultáneamente llevado al Portapapeles para estar disponible en una acción de Pegado. La acción **Pegar** trae desde el Portapapeles el objeto allí colocado y lo presenta a partir de la zona seleccionada.

La forma de invocar esas acciones, siempre a partir de un objeto seleccionado, es desplegar el menú **Editar** y seleccionar de él la acción deseada. Hay una manera alternativa que sería desplegando un llamado menú contextual, que aparece cuando se pulsa el botón derecho del ratón sobre el objeto seleccionado. Cuando se lleva a cabo dicha acción sobre un dato o conjunto de datos seleccionados, el menú contextual que aparece es como el de la Figura 4.1. En

Cortar
Copiar
Pegar
Borrar
Fuente de cuadrícula

Figura 4.1 Menú contextual para un conjunto de datos seleccionados.

este menú aparecen las acciones antes citadas además de la función **Borrar** que permite borrar el objeto seleccionado (en este caso los datos) sin colocarlo en el Portapapeles por lo que no estará disponible para ser pegado. La eliminación, también, se puede llevar a cabo, a partir del objeto seleccionado, pulsando la tecla **Supr**.

La combinación de las acciones **Cortar, Copiar, Pegar y Borrar** permite llevar a cabo acciones complejas con pocos toques de ratón. Así por ejemplo, para borrar un conjunto de datos habrá que seleccionarlos y elegir **Borrar**, y para repetir el mismo dato a partir de una posición determinada habrá que copiar el dato en el Portapapeles (acción **Copiar**), seleccionar la zona en la que se debe repetir el dato y elegir la opción **Pegar.**

4. Edición de un fichero de datos.

Sobre un fichero de datos ya creado podemos efectuar los cambios que se deseen, tanto en las variables como en los propios datos. Si deseamos cambiar las características de algunas de las variables, lo hacemos directamente en la **vista de variables** del **Editor de datos SPSS**.

De la misma manera, cambios en los datos los haremos directamente en la **vista de datos** del **Editor de datos SPSS**. A estos efectos sería aplicable todo lo dicho en el apartado anterior 3.2.2. La realización de cambios sobre casos completos y variables completas las vemos a continuación.

4.1. Edición de casos completos.

Para llevar a cabo las acciones que vamos a describir se necesita seleccionar un caso o un conjunto de casos. La selección se hace pinchando en el número de registro correspondiente. Las acciones se llevan a cabo o con el menú **Edición**, o con el menú contextual, el más cómodo.



Figura 4.2 Menú contextual para la edición de casos.

El menú contextual que se despliega en la selección de casos es como el que aparece en la Figura 4.2. En él aparecen las cuatro opciones que aparecían para los datos como son Cortar, Copiar, Pegar y Borrar; tales acciones han sido suficientemente explicadas anteriormente por lo que lo único que diremos aquí es que se comportan de manera análoga a lo visto y el lector ya sabe como funcionan. No obstante aparece debajo una acción que es propia de los casos: **Insertar casos.** Cuando se selecciona aparece un caso nuevo inmediatamente antes del caso seleccionado en el que toda la información estará en blanco esperando para ser rellenada con los datos de un nuevo caso. También puede ser rellenada con los datos de un caso ya existente en la matriz, para lo cual no tenemos más que **Cortar** dicho caso completo y **Pegar** en el caso en blanco; con ello habríamos cambiado la posición de ese caso dentro de la matriz de datos. Otra forma de cambiar la posición de un caso (o grupo de casos) es seleccionarlo y, después, arrastrarlo con el ratón hasta la posición deseada. Si no se rellena el caso en blanco insertado se entendería que en ese nuevo caso todas las variables tienen datos faltantes.

4.2. Edición de variables completas.

Para llevar a cabo las acciones que vamos a describir se necesita seleccionar una variable o un conjunto de variables. La selección de una variable se hace pulsando el botón izquierdo sobre el nombre de la misma y así quedará seleccionada. Para seleccionar una serie de variables consecutivas bastará con seleccionar la primera y arrastrar el ratón hasta la última, ese arrastre se puede hacer con el botón izquierdo (en cuyo caso para llevar a cabo las acciones habrá que desplegar el menú **Editar** o el menú **Datos**), o con el botón derecho (y en este caso, soltando el botón derecho al final del arrastre, las acciones se podrán llevar a cabo mediante el menú contextual que aparece).

El menú contextual que se despliega en la selección de variables es como el que aparece en la Figura 4.3. En él aparecen las cuatro opciones que aparecían para los datos como son



Figura 4.3 Menú contextual para la edición de variables.

Cortar, Copiar, Pegar y Eliminar; tales acciones han sido suficientemente explicadas anteriormente por lo que lo único que diremos aquí es que se comportan de manera análoga a lo visto y el lector ya sabe como funcionan. No obstante aparecen, debajo, una serie de acciones que son propias de las variables. Explicaremos aquí la de **Insertar variables**.

Cuando se selecciona **Insertar variables**, aparece una variable nueva inmediatamente antes de la variable seleccionada en la que toda la información estará en blanco esperando para ser rellenada. El nombre de la variable habrá sido asignado por SPSS de oficio, según las normas que vimos en la definición de variables, así como su tipo, su presentación y su asignación de valores faltantes. Además, los datos correspondientes a los casos existentes los rellenará automáticamente con datos faltantes del sistema, pudiendo introducirse los datos en la variable posteriormente. Finalmente, cambiar la posición de una variable dentro de la matriz de datos se lleva a cabo de forma completamente similar a como se ha visto para los casos. Las acciones explicadas se pueden llevar a cabo en la **vista de datos** o en la **vista de variables**.

5. Operaciones con variables. Generación de nuevas variables.

Una vez que se ha creado un fichero de datos con SPSS, y antes de entrar en el análisis, podemos llevar a cabo la tarea de crear nuevas variables a partir de las ya disponibles. Así, se generan nuevas variables esencialmente de dos maneras distintas: como resultado de operaciones aritméticas y/o lógicas que generalmente combinan una o más variables, y como resultado de una recodificación (o cambio de códigos) de la variable original. De lo dicho hasta aquí se deduce que siempre tendremos una variable de destino (que puede ser nueva o existente) que recibirá el resultado de la operación realizada y, generalmente, tendremos al menos una variable origen sobre la que se producirá la operación; ni que decir tiene que la variable de origen y la variable de destino pueden ser la misma (en cuyo caso habremos realizado simplemente una transformación o recodificación de la variable de origen). Los diferentes



Figura 5.1 Ventana de Transformación de variables mediante el cálculo.

procedimientos para llevar a cabo este tipo de operaciones sobre las variables están en el menú **Transformar**. Aquí solo veremos dos de ellos.

5.1. Transformación de variables. Procedimiento Calcular.

La manera de invocar el procedimiento que dará lugar al nuevo contenido de una variable es usando *Transformar →Calcular...*. Cuando se hace eso, aparece una ventana como la de la Figura 5.1. Lo primero que haremos será describir en términos generales las distintas zonas de la ventana para especificar después el funcionamiento de la transformación.

Observando la ventana se ve que en el ángulo superior izquierdo de la ventana aparece una zona para la **Variable de destino**; en ella se escribirá el nombre de la variable que recibirá el resultado de la transformación u operación realizada; esa variable puede ser una de las existentes, en cuyo caso el resultado machacará el contenido anterior de ella, o una variable nueva cuyo nombre tendrá las características de los nombres de las variables de SPSS; en el caso de que la variable sea nueva, se le pueden definir las características de tipo y etiqueta sin mas que pulsar la tecla **Tipo y Etiqueta** que aparece debajo de la zona del nombre de la variable. En nuestro caso la variable que vamos a obtener como resultado de la transformación se denomina "lpthm".

Inmediatamente debajo del nombre de la variable de recepción, aparece una caja deslizante donde están las variables del fichero de datos para que pinchando sobre ellas nos sea más fácil crear la expresión que, una vez calculada, nos dé lugar a la nueva variable. Cuando se selecciona una variable de la caja, se la pasará a la zona de **Expresión numérica** pinchando el botón \blacktriangleright .

Siguiendo a la derecha y en la parte alta de la ventana aparece la caja de **Expresión numérica** dónde se escribirá la expresión que, ejecutada, nos dará lugar a la nueva variable. Las expresiones se escriben combinando nombres de variables, números y funciones, con operadores matemáticos y/o lógicos. Para ello podemos utilizar los operadores, dígitos, y funciones que aparecen en la ventana. Las funciones son pegadas en la caja de **Expresión numérica** pulsando sobre ellas dos veces o marcándolas y pulsando el botón \blacktriangle , mientras que los operadores y los dígitos son pegados simplemente pulsando una sola vez sobre ellos.

Los operadores matemáticos disponibles son los clásicos: + (adición), - (sustracción), * (multiplicación), / (división) y ** (exponenciación). Los operadores lógicos se describen en la Tabla 5.1. Estos operadores son funciones que dan como resultado Verdadero (V=1) o Falso (F=0), ayudándonos a realizar acciones en función de uno u otro resultado. Los operadores lógicos permiten acciones selectivas sobre determinadas variables, pues ellas se llevarían a cabo sólo sobre los casos en los que la expresión lógica, definida a tal efecto, sea verdadera. Veamos a continuación algunos ejemplos de expresiones lógicas. La expresión A=2, será verdadera en los casos en los que la variable A contenga el valor 2. La expresión (A>=2 | A=0) será verdad cuando en A haya un valor mayor o igual que 2 o cuando haya el valor 0. La expresión (A=1 & B>3) será verdad cuando en A haya el valor 1 y simultáneamente en B haya un valor mayor que 3. El bloque de operadores se completa con los paréntesis de agrupación (). Los operadores y funciones que aparecen dentro de los paréntesis se evaluarán antes que los que están fuera. Si se selecciona un texto en **Expresión numérica** y se pulsa (), dicho texto queda encerrado entre paréntesis.

Junto a los dígitos está el punto decimal (.) (hay que hacer notar que el separador decimal aquí es el punto, y no la coma), y debajo de ellos aparece el botón **Eliminar** que permite borrar el trozo de texto seleccionado en el cajón de la expresión numérica. Ello nos indica que ese cajón donde se va escribiendo la expresión numérica es de hecho una ventana de edición a todos los efectos en la que el usuario puede escribir la expresión que desee.

Nos falta hablar de las funciones, que aparecen en una caja deslizable a la derecha de la zona de los dígitos, y, como ya se ha indicado, lo haremos sólo de las esenciales dejando las demás para que el usuario las revise. Sin embargo antes de enumerarlas y explicar su funcionamiento, digamos que cuando una función es seleccionada y pegada en la caja de las expresiones, aparece escrita y entre paréntesis figuran tantos signos de interrogación como argumentos (variables o constantes) tenga la función, por ello el usuario deberá cambiar tales signos por los valores correspondientes. En la Tabla 5.2 aparecen las funciones más comunes usadas en las expresiones numéricas o lógicas, así como una breve descripción de las mismas. Desde luego que en la ventana, según se desplaza la barra de ella, se pueden ver muchísimas más funciones que hacen referencia a múltiples aspectos del trabajo que se puede hacer con SPSS; sin embargo para el usuario principiante las expuestas son suficientes. Esta claro que cuando una función no pueda ser calculada (porque el valor del caso al que pretenda aplicarse no sea correcto para ella) no se aplicará, y en la variable aparecerá un blanco indicando el valor faltante, avisándonos explícitamente SPSS de la inconsistencia existente. En nuestro caso concreto, en la variable "lpthm" guardaremos el resultado de calcular el logaritmo neperiano de la variable "pthm"; es claro que no se calculará siempre que el valor que aparezca en la variable "pthm" sea menor o igual que cero, donde no está definido el logaritmo.

Operador	Función que realiza
<	Operador Menor Que. Para términos numéricos, da como resultado verdadero si el término que precede al operador es menor que el término que sigue. Para términos alfanuméricos, es verdadero si el término que precede aparece antes que el término que sigue en la secuencia de clasificación (por orden alfabético).
>	Operador Mayor Que. Lo mismo que el anterior pero con mayor.
<=	Operador Menor O Igual Que. Para términos numéricos, da como resultado verdadero si el término que precede al operador es menor o igual que el término que sigue. Para términos alfanuméricos, es verdadero si el término que precede aparece antes que el término que sigue en la secuencia de clasificación (por orden alfabético), o si ambos son iguales.
>=	Operador Mayor O Igual Que. Lo mismo que el anterior pero con mayor o igual.
=	Operador Igualdad lógica. Da como resultado verdadero si los dos términos son exactamente iguales. Si dos términos alfanuméricos son de distinta longitud, el más corto se completa con espacios a la derecha, hasta igualar en longitud al más largo, antes de la comparación.
~=	Operador Desigualdad lógica. Da como resultado Verdadero si los términos no son iguales. Si dos términos alfanuméricos son de distinta longitud, el más corto se completa con espacios a la derecha, hasta alcanzar la longitud del más largo, antes de la comprobación.
&	Operador Y lógico. Da como resultado verdadero si tanto el término precedente como el que le sigue son verdaderos. Los términos pueden ser de tipo lógico o numérico; los términos numéricos diferentes de 0 ó 1 se tratan como valores perdidos.
I	Operador O lógico. Da como resultado verdadero si el término precedente o el siguiente también lo son. El resto es como en el caso del operador Y lógico.
~	Negación lógica. Da como resultado Verdadero si el término que sigue es Falso. El resto como en el operador Y lógico.

Función	Tipo y Acción que realiza
ABS(expr_num)	Numérico. Devuelve el valor absoluto de expr_num, que debe ser un valor numérico. Ej: ABS(A-B)
EXP(expr_num)	Numérico. Devuelve e elevado a la potencia expr_num, siendo e la base de los logaritmos neperianos y expr_num un valor numérico.
LG10(expr_num)	Numérico. Devuelve el logaritmo en base 10 de la expr_num, la cual debe ser numérica y mayor que 0.
LN(expr_num)	Numérico. Devuelve el logaritmo en base e de la expr_num, la cual debe ser numérica y mayor que 0.
MAX(valor, valor,)	Numérico. Compara los argumentos y devuelve como resultado el mayor de los que tengan valores válidos. Esta función requiere al menos dos argumentos.
MIN(valor, valor,)	Numérico. Compara los argumentos y devuelve como resultado el menor de entre los que tengan valores válidos. Esta función requiere al menos dos argumentos.
MISSING(variable)	Lógico. Devuelve 1 (verdadero) si la variable tiene un valor perdido. El argumento debe ser un nombre de variable del archivo de datos de trabajo.
SQRT(expr_num)	Numérico. Devuelve la raíz cuadrada positiva de expr_num, la cual debe ser numérica y no negativa.

Tabla 5.1. Operadores lógicos para la creación de nuevas variables

Tabla 5.2. Algunas funciones	s usadas en	las transforma	aciones de	variables
------------------------------	-------------	----------------	------------	-----------

Como ya hemos dicho, una vez construida nuestra expresión numérica, el valor que esta dé para cada caso sustituirá al valor existente en la variable de destino para dicho caso. La ejecución que así se conseguirá será para todos los casos de la base de datos. Sin embargo, a veces el usuario quiere restringir el cálculo de la expresión numérica a un subconjunto de casos que verifiquen una condición, no llevándose a cabo el mismo para los casos que no cumplen tal condición. Para ello habrá que especificar la condición que se debe cumplir, y eso se hace pulsando el botón **Si...** que aparece justo debajo del conjunto de operadores de la ventana (en la Figura 5.1). Al pulsar tal botón aparece una nueva ventana como la de la Figura 5.2.

Calcular variable: Si los casos		×
 acfis presencia de retino netro neuro hba1c ca ca p cr pthi pthi bmdcue szl24 sztri szcue 	 Incluir todos los casos Incluir si el caso satisface la condición: pthm > 0 	

Figura 5.2 Ventana de especificación de la condición para el cómputo condicional.

Lo primero es decir que en esta ventana estableceremos la condición que permita extender el cálculo al conjunto de casos que la cumplan. Lo primero que se ve en la ventana son las dos opciones: **Incluir todos los casos e Incluir si el caso cumple la condición;** ambas opciones son incompatibles y realmente la que a nosotros nos interesa es la segunda, por lo que habrá que pincharla, pues SPSS ofrece siempre la que no supone ninguna restricción. En la ventana, aparece a la izquierda una caja con barra deslizante dónde están todas las variables del fichero activo para que, seleccionándolas, nos permita escribir la condición. A la derecha de esta caja aparece otra caja en la que ha de escribirse la condición lógica. Inmediatamente debajo aparecen los mismos operadores y las mismas funciones ya vistas anteriormente, por tanto, siguiendo lo dicho en párrafos anteriores, podemos construir la condición que deseemos. En nuestro caso hemos exigido para el cálculo del logaritmo neperiano de la "pthm" que ésta tenga un valor mayor que 0. Escrita la condición bastará pulsar el botón **Continuar** para volver a la ventana **Calcular variable**, la de la Figura 5.1.

En esa ventana ya sólo bastará con pulsar el botón **Aceptar** y se llevará a cabo el cálculo. Por último reiterar que en ventanas como esta, que son complejas, la ayuda contextual (que se consigue pulsando el botón derecho del ratón sobre el tema de interés) de SPSS es de gran calidad y puede ser una gran guía para un usuario novel.

Como se ve la función de crear variables a partir de cómputos que combinan otras variables es muy potente, quedando para el usuario su práctica más detallada.

5.2. Recodificación de variables. Procedimiento Recodificar.

Ya habíamos anunciado antes que la creación de nuevas variables se podía llevar a cabo por el procedimiento de Recodificación, que, como su nombre indica, cambia los valores de una variable por otros mediante un proceso que no es de tipo numérico sino de especificación de las reglas que producirán el cambio. La forma de invocar la recodificación es a partir del menú *Transformar* \rightarrow *Recodificar...* Colocados sobre la opción **Recodificar**, aparece un nuevo menú desplegable para darnos a elegir entre: **En las mismas variables...** o **En distintas variables...** Estas dos alternativas nos permiten grabar el resultado de la recodificación en la misma variable fuente que la produce o en una variable distinta (nueva o no). La recodificación sobre la misma variable tienen el inconveniente, si es que lo es, de que el resultado de la misma destruye lo que ya había en ella, de manera que la vuelta a la codificación primitiva sólo será posible mientras se pueda aplicar la acción de **Deshacer**; no obstante, a veces eso es lo deseado por el usuario. La recodificación en una variable distinta conserva las dos, la codificación primitiva y la recodificación **En distintas variables...**, ya que generalmente es la más aconsejable (la recodificación **En las mismas variables...** se hace de forma similar).

La ventana correspondiente es la de la Figura 5.3; dicha ventana nos permite elegir las variables que van a ser recodificadas; lo que significa que varias variables pueden ser recodificadas simultáneamente con las mismas reglas. Según se observa en la ventana hay una caja a la izquierda que tiene todas las variables que aparecen en el fichero activo para

Recodificar e	🖬 Recodificar en distintas variables 🛛 🔀				
 Actividad física (1= Presencia de retino Presencia de nefro Presencia de neuro Presencia de neuro Hemoglobina glico Calcio [ca] Fósforo [p] Creatinina [cr] 	Var. numérica -> Var. de resultado: szl24> szl24_rc Variable de resultado Nombre: szl24_rc Etiqueta: szl24 recodificada Cambiar				
 Pth media [pthm] Pth intacta [pthi] Densidad de masa Valor tipificado de Valor tipificado de 	Valores antiguos y nuevos Si Aceptar Pegar Restablecer Cancelar Ayuda				

seleccionar aquella o aquellas que van a ser recodificadas; la forma de seleccionarlas es pinchar sobre ellas con el ratón y pulsar el botón ▶, para que sean incluidas en el cajón que hay a la derecha (Var. Numérica -> Var. de Resultado) en el que aparecen las variables a ser recodificadas. En dicho cajón debe aparecer obligatoriamente, para cada una de ellas, una variable destino de la recodificación. Así, seleccionada una variable para recodificar y llevada al cajón de Var. Numérica → Var. de resultado, aparece un símbolo de interrogación señalando que debe escribirse la variable de destino. Para ello bastará escribir, en la zona de Nombre, el nombre de una variable de SPSS, que puede ser nueva o ya existente; en cuanto esté escrito el nombre se iluminará el botón Cambiar que nos permitirá asociar la variable destino con la variable fuente va seleccionada. Si la variable destino existe, los valores de ella quedarán machacados por el resultado de la recodificación y SPSS nos pedirá una confirmación para ello. En el caso de que no exista, en la variable nueva aparecerá el resultado de la recodificación, teniendo en cuenta que si hay algún valor fuente que no ha sido recodificado, aparecerá en la variable de destino como un dato faltante del sistema. En la ventana se ve que se puede especificar una etiqueta para la variable resultado de la recodificación, sujetándose a las normas va descritas para las etiquetas. En la Figura 5.3 aparece un ejemplo en el que la variable "szl24" va a ser recodificada en una nueva variable que se denominará "szl24 rc" y que tendrá de etiqueta "szl24 recodificada. Una vez seleccionadas las variables a recodificar se puede especificar una recodificación condicional pulsando el botón Si.. que funciona exactamente igual que el Si.. condicional va visto en la ventana Calcular variable (Figura 5.1). Finalmente habrá que pasar a la especificación de las normas que definen la recodificación; eso se hace pulsando el botón **Valores antiguos y nuevos..** que nos llevará a una ventana como la de la Figura 5.4.

La ventana de la Figura 5.4 está claramente dividida en dos zonas: la zona en la que se

Recodificar en distintas variables: Val	ores antiguos y nuevos	×
Valor antiguo Valor: Perdido por el sistema Perdido por el sistema o usuario Rango: 2 <u>h</u> asta 3 C Rango: Del menor hasta	Valor nuevo Valor: 1 C Perdido por el sistema Copiar valores antiguos Antiguo> Nuevo: Añadir Borrar	
C Rango: hasta el mayor C Todos los demás valo <u>r</u> es	Las variables de resultado son cadenas Ancho: 8 Convertir cadenas numéricas en números (5->5) Continuar Cancelar Ayuda	

Figura 5.4 Especificación de valores antiguos/nuevos en una recodificación

especifican los valores que van a cambiar (los valores "antiguos") y la zona dónde se especificarán los valores "nuevos". En principio lo lógico es que a todo valor antiguo le corresponda uno nuevo, pero, si no es así, sino que algún valor antiguo no fuera recodificado, su correspondiente valor nuevo será el de perdido por el sistema, si la variable destino es numérica, o blanco, si es de cadena. Las formas de especificar los valores antiguos que van a ser recodificados figuran en la Tabla 5.3.

A cada especificación de valores antiguos habrá que hacerle corresponder un valor nuevo. Los tres valores nuevos posibles son: un único valor que especifica el usuario, o el valor de perdido por el sistema o copiar el valor antiguo existente. En cuanto hayamos elegido una pareja valores_antiguos/valor_nuevo se iluminará el botón **Añadir** y pulsándolo añadiremos la regla de recodificación a la caja de las mismas que aparece debajo de la zona de valores nuevos. Con este proceso se van añadiendo las reglas de recodificación hasta la última. Si alguna regla no estuviera correcta se podría pulsar sobre ella y volvería a ser colocada sobre los valores antiguos y el valor nuevo para que hagamos las modificaciones oportunas pulsando el botón

Opción	Valores a los que se refiere		
Valor	Valor antiguo individual que se va a recodificar en un valor nuevo.		
Perdido por el sistema	Valores antiguos que son valores perdidos (faltantes) del sistema.		
Perdido por el sistema	Valores antiguos que, o bien son valores perdidos del sistema o bien		
o usuario	son valores declarados como perdidos (faltantes) por el usuario.		
Rango: Desde-Hasta	Un rango que incluye a los dos valores. No está disponible para variables de cadena. Incluirán cualesquiera valores definidos por el usuario como perdidos que se encuentren dentro del rango.		
Rango: Del menor hasta	Rango que incluye a todos los valores desde el menor valor registrado para la(s) variable(s) hasta un valor especificado. Incluye cualquier valor perdido definido por el usuario que se encuentre dentro del rango.		
Rango: hasta el mayor	Rango que incluye a los valores desde un valor especificado hasta el mayor valor registrado para la(s) variable(s). Se incluye cualquier valor perdido definido por el usuario que se encuentre dentro del rango.		
Todos los demás valores	Se identifica así cualquier valor no incluido en una de las especificaciones incluidas en la caja (lista) Antiguo->Nuevo.		

Tabla 5.3.	Formas	de especificad	ción de los	valores	antiguos pa	ra una i	recodificació	ón.
I usia cici	1 01 11145	ac especificat		value es	uninguos pu	1	i ecounicació	/

Cambiar para que sea reintegrada en la caja en la nueva versión. Si deseamos borrar una regla de recodificación basta con seleccionarla, pulsando con el ratón sobre ella, y pulsando el botón **Borrar** queda eliminada del conjunto de reglas. En el ejemplo de la Figura 5.4 las reglas de recodificación escritas son: el valor 1 será cambiado por el valor 0 y los valores de 2 a 3 (ambos inclusive) son recodificados con el valor 1. Por último y justo debajo de las reglas, aparecen dos nuevas opciones. La primera nos permite especificar el tipo cadena (y su ancho) para la nueva variable, lo que nos permite recodificar una variable numérica a variable de cadena. La segunda nos permite lo contrario: recodificar una variable de cadena a una variable numérica (los datos de la variable de cadena que son números, son recodificados automáticamente a verdaderos números en la nueva variable; los datos que contienen otros caracteres pueden ser recodificados como números o datos faltantes del sistema). Al final basta con pulsar el botón **Continuar** y volveremos a la ventana de la Figura 5.3 estando ya listos para que se produzca la recodificación. Para ello bastará con pulsar el botón **Aceptar**.

6. Ordenación y selección de casos.

Explicaremos a continuación una serie de procedimientos que suponen ya un manejo más sofisticado de un fichero de datos. En primer lugar explicaremos cómo se puede ordenar un fichero de datos en función de los valores que presenten una o varias variables. A continuación explicaremos como se pueden seleccionar casos de un archivo para llevar a cabo en ellos análisis estadísticos y no hacerlos con los casos no seleccionados. También veremos la ponderación de casos, lo que nos permitirá poder introducir y analizar datos con frecuencias. Por último explicaremos como se divide un fichero en subconjuntos de casos, en cada uno de los cuales se lleva a cabo el análisis elegido; a este procedimiento se le denomina Segmentación. Los tres procedimientos que vamos a explicar se encuentran dentro del menú **Datos** de la ventana de Edición de Datos.

6.1. Procedimiento Ordenar casos.

Una tarea muy frecuente, cuando uno ha acabado de grabar datos en un fichero, es la de ordenar los mismos por algunas variables que nos permitan revisarlos de una manera fácil. Eso se lleva a cabo con el Procedimiento **Ordenar casos**, que se invoca como **Datos →Ordenar casos**. Este procedimiento ordena los casos (las filas) del archivo de datos activo basándose en los valores de una o más variables de ordenación. Puede ordenar los casos en orden ascendente o descendente. Si selecciona más de una variable de ordenación, los casos se ordenarán respecto a cada variable dentro de las categorías de la variable anterior de la lista **Ordenar por**. Por ejemplo, si selecciona "sexo"como la primera variable de ordenación y "edad" como la segunda, los casos se ordenarán por sexo, y dentro de cada categoría del sexo se ordenarán por edad. Para las variables de cadena, las letras mayúsculas preceden a las minúsculas correspondientes en orden. Por ejemplo, el valor de cadena "Sí" precede a "sí". Cuando se invoca **Ordenar casos** aparece una ventana como la de la Figura 6.1.



La ordenación se hace en función de los contenidos de las variables por lo que en la

Figura 6.1 Ventana de ordenación de casos por variables.

ventana aparece una caja con las variables del fichero activo a la izquierda de la ventana; de ella se seleccionan las variables (por el procedimiento que tantas veces hemos explicado) y caen a la caja de **Ordenar por**; para cada variable se puede seleccionar el **Orden de clasificación** de entre las dos alternativas: **Ascendente** o **Descendente**. Hecha la selección y pulsando **Aceptar** se lleva a acabo la ordenación deseada. En nuestro ejemplo, los casos se ordenarán por los códigos de la variable "sexo". Si se ha hecho una reordenación, la forma de volver al orden original es ordenando respecto a una variable en la que tengamos guardado dicho orden original. La ordenación suele ser útil para seleccionar casos determinados o para búsquedas muy particulares, o para listar los casos en el orden deseado.

6.2. Procedimiento Seleccionar casos.

Seleccionar casos proporciona varios métodos para seleccionar un subgrupo de casos

🚯 Seleccionar casos	×
 num edad sexo peso talla inc tevol tabaco alcohol ingca acfis presencia de retino netro neuro hba1c aci 	 Seleccionar I odos los casos Si se satisface la condición Si Muestra aleatoria de casos Muestra Basándose en el rango del tiempo o de los casos Flango Usar variable de filtro: Los casos no seleccionados son Eltrados Eltrados
Estado actual: No filtrar casos	
Aceptar P <u>e</u> gar	Restablecer Cancelar Ayuda

Figura 6.2 Ventana de selección de casos de un fichero de SPSS

basándose en criterios que incluyen variables y expresiones complejas. También se puede seleccionar una muestra aleatoria de casos. Los criterios usados para definir un subgrupo pueden incluir: Valores y rangos de las variables, Rangos de fechas y horas, Números de caso (filas), Expresiones aritméticas, Expresiones lógicas y Funciones.

La forma de invocar la selección de casos es **Datos** \rightarrow Seleccionar casos. Cuando se hace tal cosa aparece una ventana como la de la Figura 6.2. La ventana muestra, a la izquierda, una caja con las diferentes variables del fichero activo para que sean utilizadas en las distintas formas en que uno quiere seleccionar los casos. A la derecha aparecen las cinco formas de seleccionar casos; aquí describiremos las dos primeras.

Todos los casos desactiva el filtrado y utiliza todos los casos (es realmente una noselección de casos); ella está activa cuando se entra por primera vez a la ventana de selección de casos.

Si se satisface la condición utiliza una expresión condicional para seleccionar los casos. Si el resultado de la expresión condicional es verdadero, el caso se selecciona. Si el resultado es falso o perdido, entonces el caso no se selecciona. La forma de escribir la expresión condicional es la misma que ya hemos descrito anteriormente y la forma de acceder a ella, una vez seleccionada la opción Si se satisface la condición, es pinchar en el botón Si... Una vez que

hayamos escrito la condición y vuelto a la ventana de **Seleccionar casos..** la condición aparecerá escrita a la derecha del botón **Si...** Cuando se ejecuta esta opción se crea una variable de filtro (o de selección), con el nombre FILTER_\$, que está activa; una variable de filtro no es mas que una variable que tendrá el valor 0 para los valores no seleccionados y el valor 1 para los valores seleccionados; redundando en que la variable de filtro es una variable cualquiera, podemos decir que se le puede cambiar el nombre y sus etiquetas, aunque hay algún problema más con el tipo.

Los casos no seleccionados son.. permite elegir entre mantener los casos no seleccionados en el archivo de datos o borrarlos por completo del mismo; si nos decantamos por Filtrados, los casos no seleccionados permanecerán en el archivo activo (quedando marcados con una barra diagonal sobre el número de fila) pero no serán incluidos en ningún análisis mientras esté activa la selección. Por el contrario, si elegimos Eliminados los datos serán borrados del fichero activo, y si éste es guardado en esa situación serán perdidos para siempre. SPSS ofrece el filtrado como opción de oficio pues es menos peligrosa que la opción de borrado para la que exige una acción positiva del usuario.

Escrita la forma en la que deseamos seleccionar los datos y cómo queremos que sea ella, bastará con pulsar el botón **Aceptar** para que la selección sea llevada a cabo, pudiendo deshacerse tal selección, con las restricciones que ya hemos citado, sin más que volver a la ventana y elegir **Todos los casos.**

Se ha de tener precaución cuando se especifica una selección pues, aunque aparece la palabra Filtrado en la barra de estado del Editor de datos y aparecen las barras tachando los casos no seleccionados, con frecuencia se olvida que tal selección está activa.

FASE DE ANÁLISIS DE DATOS

Haremos ahora una exposición de cada uno de los métodos estadísticos básicos de SPSS y de los resultados que proporcionan. Pero antes, describiremos la forma general en que se presentan dichos resultados; SPSS tiene una aplicación especial que maneja los resultados y que se denomina el Visor de resultados.

7. El Visor de Resultados.

Para ejecutar cualquier análisis estadístico habrá que invocar el menú análisis que se encuentra, entre otras, en la ventana del Editor de datos. Al desplegarse el menú se elegirá el análisis deseado y el resultado aparecerá en una ventana como la de la Figura 7.1. Esa ventana (llamada <u>Visor SPSS</u>) es siempre la misma, aunque en el ejemplo que hemos descrito se refiere a los resultados de la descriptiva de unas variables de nuestro fichero activo. Lo primero que hay que decir de los resultados es que son una entidad aparte de los datos y que como tal son susceptibles de imprimirse, editarse y ser guardados en un fichero aparte, que tendrá la extensión **.spo**. Como toda ventana de SPSS, está compuesta por una zona de menús desplegables y barra de herramientas, y una zona dónde se presentan los resultados, zona que se divide a su vez en dos partes, una en la que aparece un índice de tales resultados y otra en la que aparecen los resultados propiamente dichos. A continuación describiremos la zona de resultados con detalle.

7.1. La zona de resultados: el índice y los resultados propiamente dichos.



Figura 7.1 Visor de resultados de SPSS

La zona de resultados es la que ocupa la mayor parte de la ventana, siendo de ella la zona de índice la más estrecha y que está a la izquierda y la de resultados la de la derecha más ancha; la zona de índices se puede agrandar y achicar a voluntad (achicando y agrandando la zona de resultados) sin más que pinchar sobre la barra que separa ambas ventanas y arrastrar el ratón. Como se ve en la figura, tanto la ventana de índices como la de resultados tienen barras horizontales que se desplazan para abarcar el conjunto de la información allí expuesta, si es que

no cabe en la ventana visible, apareciendo barras verticales, con la misma función cuando así se necesite.

El índice de los resultados aparece a la izquierda y como su propio nombre indica funciona como catálogo de los mismos. Lo primero que llama la atención es que tiene una estructura jerárquica (arborescente) de manera que los distintos resultados y sus diferentes estructuras son presentados como las diferentes ramas de una jerarquía. Así, en el ejemplo que nos ocupa, tendremos un objeto de Resultados (el más general), dentro de él está el análisis realizado, que es Descriptivos, y dentro de él los objetos Titulo, Notas y Estadísticos descriptivos. Cada una de las entradas del índice son susceptibles de ser seleccionadas (al pinchar sobre ellas con el ratón) y eso conlleva la selección de los resultados correspondientes. Seleccionada una entrada se puede borrar (sin más que pulsar la tecla **Supr**) lo que conllevará la desaparición del resultado propiamente dicho. Seleccionada una entrada del índice ser puede mover en el árbol sin mas que desplazarla con el ratón hasta dónde deseemos, moviéndose el resultado correspondiente a dicha posición. Pero la función más importante del índice será la de posicionamiento rápido sobre los resultados, pues al seleccionar una de las entradas del índice, automáticamente en la zona de resultados aparece el resultado correspondiente a dicha entrada.

Como se ve el índice tiene una relación muy estrecha con los resultados y cumple una función de facilitar su presentación y manejo. El índice puede ser manejado como una entidad independiente. Seleccionando uno de sus objetos y pinchando sobre él con el botón derecho del ratón se despliega el menú contextual con las opciones: ¿qué es esto?, información muy resumida de lo que es el índice; Cortar, Copiar Pegar, Pegar objetos con las funciones conocidas; Exportar que permite grabar los resultados en un fichero html (formato que se usa en las páginas web) o en un fichero de texto. Aparecen otras opciones que debido a su complejidad no explicaremos. También puede disponerse de una manera más o menos comprimida pulsando sobre los signos en los que arranca cada rama del árbol de resultados. Si aparece un signo menos estará expandida la rama (con todas sus ramas visibles) y por tanto será susceptible de ser contraída (es decir de esconder todas sus ramas), mientras que si aparece un signo más, es que las ramas están contraídas por lo que serán extensibles.

La parte de la ventana de los resultados es susceptible de ser manejada como ahora veremos. Lo primero que hay que decir de ella, aparte de lo dicho sobre los desplazamientos verticales y horizontales, es que no es sólo una ventana en la que se puedan ver los resultados sino que son editables a todos los efectos. La forma de editar un objeto (una componente) de los resultados es pinchar dos veces sobre el objeto y automáticamente se dispara el Editor de resultados que describiremos de manera breve ahora. La pieza clave del editor de resultados es la Tabla **Pivote**; la tabla pivote es una tabla creada por SPSS y que tiene como características principales el manejo de casilla, filas y columnas con total comodidad (eliminándolas, moviéndolas) y el manejo completo de los textos que aparecen dentro de ellas. De hecho la inclusión de las tablas pivote en SPSS fue uno de los saltos más importantes de este paquete respecto de sus versiones anteriores a la que estamos estudiando. Como hemos indicado la edición de un objeto de los resultados nos permite escribir sobre ellos, cambiando el tipo de letra en toda su extensión, cambiando la orientación de las columnas o de las filas, generando gráficos particulares para el resultado, etc... De hecho El editor de resultados presenta una ventana nueva que permite enlazar con muchas de las funciones de SPSS y que presenta multitud de posibilidades de edición, tan extensas, que aquí no explicaremos pero que invitamos a que sean exploradas por el lector cuando disponga de tiempo. Pinchando una vez fuera de la tabla pivote seleccionada se nos permite liberarla y volver a la ventana general de resultados.

Entendido un poco mejor como presenta SPSS los resultados explicaremos los menús e iconos de la ventana general de resultados, mostrando las funciones que se pueden llevar a cabo.

Menú	Función
Archivo	Todas las funciones que se pueden hacer con archivos: Abrir, cerrar, guardar, importar, exportar, imprimir, etc., pero teniendo en cuenta que en este caso los ficheros a los que nos referimos son esencialmente de resultados (*.spo).
Edición	Realiza todas las funciones típicas de la edición como son: cortar, copiar, eliminar, buscar, reemplazar, etc, teniendo en cuenta que tales acciones se ejecutan sobre los resultados.
Ver	Permite controlar la forma en la que se ve la ventana principal, controlando las barras que aparecen en ella así como la forma en la que se presentan los datos.
Datos	Es el mismo que en el Editor de datos.
Transformar	Es el mismo que en el Editor de datos.
Insertar	Permite insertar en el objeto seleccionado desde títulos o píes de páginas a saltos de página o a objetos cualesquiera, como gráficos e incluso archivos de sonido.
Formato	Alinea los resultados en las tres posiciones, izquierda, centro o derecha.
Analizar	Es el mismo que en el Editor de datos.
Gráficos	Es el mismo que en el Editor de datos.
Utilidades	Es el mismo que en el Editor de datos.
Ventana	Es el mismo que en el Editor de datos.
?	Es el mismo que en el Editor de datos.

Tabla 7.1. Menús de la ventana del visor de resultados.

7.2. Menús e Iconos del Visor de resultados.

Lo primero a lo que nos referiremos será a la línea de menús de la ventana de resultados, que es descrita someramente en la Tabla 7.1, dejando al usuario su estudio más detallado, y que se corresponde en bastante extensión con la del Editor de datos.

Bajo la línea de menús, aparece la barra de herramientas con una serie de iconos que nos hacen más fáciles de acceso algunas de las tareas de los menús. Aparecen dos barras de iconos; la superior (descrita en la Tabla 7.2) es más general, y tiene un gran parecido con la del Editor de datos, mientras que la inferior (descrita en la Tabla 7.3) es mucho más particular y se encarga esencialmente del manejo del índice de los resultados. Con esto acabamos la presentación del Visor de resultados. El lector debe entender que la descripción de tal utilidad podría llevar algo tan extenso como lo que es este manual, por lo que no podemos profundizar más en él.

Icono	Función
<u> i i i i i i i i i i i i i i i i i i i</u>	Abrir archivo de datos para empezar a trabajar con él.
	Grabar el fichero activo.
B	Imprimir el archivo activo.
Q.	Vista previa de los resultados.
Ç.	Exportar resultados.
	Recuperar los cuadros de diálogo empleados.
KO	Deshacer la última acción
	Cambiar a la ventana del Editor de datos
*	Ir a un caso.
[?	Mostrar la información de las variables del fichero activo.
\oslash	Crear conjuntos con las variables.
5	Seleccionar los últimos resultados

Tabla 7.2. Primera barra de iconos de la ventana del Visor de Resultados.

Icono	Función
\$	Ascender el objeto de índice seleccionado un nivel en la jerarquía del mismo.
•	Descender el objeto de índice seleccionado un nivel en la jerarquía del mismo.
÷	Expandir las ramas contenidas en una de las ramas de los resultados.
-	Contraer las ramas contenidas en una de las ramas de los resultados.
	Mostrar en los resultados el objeto seleccionado en el índice.
	Ocultar en los resultados el objeto seleccionado en el índice.
5	Insertar encabezado, en el índice, delante del objeto seleccionado.
	Insertar título, en el objeto de resultados seleccionado.
Ē,	Insertar texto, en el objeto de resultados seleccionado.

Tabla 7.3. Segunda barra de iconos de la ventana del Visor de Resultados.

8. Estadística Descriptiva con SPSS.

A partir de ahora describiremos los procedimientos estadísticos de SPSS. Desde luego que no describiremos todos los que están disponibles en el programa, pues estas notas pretenden ser una iniciación al SPSS, pero sí daremos lo suficiente para que el usuario se enfrente a un volumen importante de problemas de Estadística.

Analizar Gráficos Utilidades	Ver	
Informes	•	
Estadísticos descriptivos	•	Frecuencias
Tablas	•	Descriptivos
Comparar medias	•	Explorar
Modelo lineal general	•	Tablas de cont
Modelos mixtos	•	Razón
Correlaciones	→T	
Regresión	•	
Loglineal	•	
Clasificar	•	
Reducción de datos	•	
Escalas	•	
Pruebas no paramétricas	+	
Series temporales	•	
Supervivencia	•	
Respuesta múltiple	•	
Análisis de valores perdidos		

Figura 8.1 Menú Analizar con todos los procedimientos estadísticos disponibles en SPSS, y con el submenú de Estadísticos descriptivos desplegado.

Todos los procedimientos se invocan como menús dentro del menú *Analizar*, que al desplegarse nos da uno como el de la Figura 8.1. En ella se ve que hay una gran variedad de procedimientos, nosotros empezaremos por el menú *Estadísticos descriptivos*. Este menú tiene un submenú con cinco opciones: *Frecuencias..., Descriptivos..., Explorar..., Tablas de contingencia.... y Razón...*. Por ahora estudiaremos únicamente la primera. A partir de ahora todos los procedimientos estadísticos que se presenten lo harán con un mismo formato; se presentará la ventana general, la ventana de opciones y el conjunto de resultados que se dan; en muchos casos las opciones no se harán en forma de una figura sino que se harán en forma de una tabla de texto en la que aparecerán las distintas alternativas de cada opción.

8.1. Procedimiento Frecuencias.

El procedimiento frecuencias se invoca como *Analizar* \rightarrow *Estadísticos descriptivos* \rightarrow *Frecuencias...*, y permite para cada una de las variables seleccionadas obtener la distribución de frecuencias de ellas, así como algunas medidas básicas de resumen a elegir y algunos gráficos también a elegir. La ventana general es la Figura 8.2. En ella se distinguen dos grandes cajas

R Frecuencias		×
 num peso talla imc tevol tabaco alcohol ingca 	⊻ariables:	Aceptar Pegar Restablecer Cancelar Ayuda
Mostrar tablas de frecuencias	àráficos <u>F</u> ormato	

Figura 8.2 Ventana principal del procedimiento Frecuencias... de SPSS

para las variables: la de la izquierda con todas las variables del fichero activo, la de la derecha con las variables seleccionadas para el análisis; repetiremos una vez más que el paso de una a otra caja se efectúa seleccionando, en una de las cajas, la variable y pulsando el botón correspondiente para pasarla a la otra caja; si es de izquierda a derecha será \blacktriangleright , y si es de derecha a izquierda será \triangleleft .

En la ventana principal aparecen a la derecha los botones generales de SPSS y que ya están explicados. Además aparece la posibilidad de incluir la tabla de frecuencias de las variables elegidas sin más que marcar en el recuadro correspondiente a **Mostrar tablas de frecuencias**. En la parte baja de la ventana aparecen tres botones que nos permiten especificar o particularizar la descriptiva que queremos hacer.

El primero es el de **Estadísticos...**, que permite calcular estadísticos básicos de resumen para las variables seleccionadas. Pulsado dicho botón aparecerá una ventana como la de la Figura 8.3. En ella se ve que los estadísticos que se pueden elegir se agrupan en diferentes epígrafes que son: Valores percentiles, Tendencia Central, Dispersión y Distribución (éste último no lo veremos). Dentro de cada epígrafe aparecen las opciones que se corresponden con los estadísticos más básicos, por lo que no se requiere de explicación adicional. En cualquier caso, siempre podremos utilizar la ayuda contextual simplemente pinchando con el botón derecho del ratón sobre la opción. Por último aparece una opción que califica cómo son los datos; la marcaremos cuando tengamos datos agrupados y los valores presentes en la matriz de datos sean puntos medios de grupos (o marcas de clase). Esto afecta al cálculo de los Cuantiles, Percentiles y Mediana.

Frecuencias: Estadísticos	X
Valores percentiles Cuartiles Puntos de corte para 10 grupos iguales Percentiles: Añadir Cambiar Borrar	Tendencia central Continuar ✓ Mgdia Cancelar ✓ Mediana Ayuda ✓ Mgda Suma
Dispersión ✓ Desy. típica ✓ Mínimo ✓ Varianza ✓ Máximo ✓ Amplitud ✓ E.T. media	Distribución Asimetrí <u>a</u> Curtosis

Figura 8.3 Ventana de la opción Estadísticos del procedimiento Frecuencias.

El segundo botón opcional es el de **Gráficos...**; pulsándolo aparece una ventana en la que podemos optar por no presentar ningún gráfico o por presentar uno de los gráficos disponibles: barras, sectores o histograma. A este último se le puede superponer una curva Normal con la media y desviación típica de los datos. Finalmente, en la parte baja de la ventana podemos optar por hacer el gráfico para frecuencias absolutas o para porcentajes.

Tanto en el caso de los estadísticos como en el caso de los gráficos, el usuario debe controlar su pertinencia, pues SPSS no lo hace. Así, si proponemos calcular la media de la variable sexo, SPSS lo hará sin más, aunque obviamente no tenga sentido, lo mismo podemos decir de un diagrama de sectores para la edad que no parece nada apropiado.

El último botón opcional es el de **Formato...** La ventana que aparece al pulsarlo, por una parte, nos permite elegir una de entre las cuatro formas disponibles para ordenar la tabla de frecuencias; por otra parte, nos permite elegir entre mostrar los resultados de todas las variables en una misma tabla (opción **Comparar variables**) y mostrar los resultados en tablas separadas para cada variable (opción **Organizar resultados según variables**) y, finalmente, podemos suprimir las tablas de frecuencias que tengan un número de categorías (filas) superior al especificado por nosotros.

En este procedimiento Frecuencias no nos detendremos en sus resultados, pues parecen bastante obvios; sólo diremos que las tablas de frecuencias aparecen con frecuencias absolutas, porcentajes con respecto al total de datos, porcentaje con respecto a los datos que no son faltantes (datos válidos) y porcentajes acumulados.

8.2. Gráficos con SPSS.

Una parte importante a la hora de exponer los resultados estadísticos es presentar algunos de ellos en forma de gráficos. A ello dedica SPSS todo un menú de la ventana principal, que es el menú **Gráficos**, con las opciones que aparecen en la Figura 8.4. De ese menú no se explicarán todas las opciones, sino sólo las más útiles para los objetivos del manual.

Las dos primeras opciones son: **Galería** e **Interactivos**. La primera ofrece de una manera gráfica y guiada (mediante un tutor) cada uno de los gráficos que se pueden invocar por separado y que se describen a continuación; esta opción puede ser seguida con facilidad por parte del lector pues el tutor que emplea SPSS es bastante amigable y claro. **Interactivos** permite construir los gráficos que veremos a continuación también de una manera interactiva y con un tutor; es el procedimiento gráfico más moderno de SPSS y mejora sensiblemente el manejo de los gráficos del mismo, así como su calidad de presentación (3D, colores, sombreados, etc....); una vez más debido a la facilidad de uso no nos detendremos más en él, por lo que el usuario puede intentarlo por su cuenta.



Figura 8.4 Menú de Gráficos.

De los que aparecen en la Figura 8.4 describiremos con detalle los gráficos de: **Barras**, Líneas, Sectores e Histograma. En secciones posteriores se explicarán **Barras de Error y** Dispersión.

Todos los gráficos tienen una forma de invocarlos y una serie de opciones que matizan su uso, por ello encontraremos que la exposición de los mismos es similar a la de otros procedimientos estadísticos. En el primer gráfico que presentemos haremos mención al editor de gráficos de SPSS que permite la manipulación de los mismos de manera muy avanzada.

8.2.1. Gráficos de Barras.



SPSS denomina

Gráfico de barras a un gráfico de variables categóricas en el que sobre cada modalidad se levanta una barra de altura proporcional a la frecuencia. La llamada del Gráfico de **Barras** da lugar a una primera ventana en la que ha de seleccionarse el tipo de gráfico de barras que se desea representar. La ventana es como la de la Figura 8.5. Lo primero es elegir entre los gráficos: **Simple, Agrupado** y **Apilado**. En el Simple, sólo se representa en el Gráfico una única variable (que es lo más común); en el Agrupado se hace un gráfico simultáneo (compuesto) en el que se presenta la distribución de una variable dentro de cada una de las categorías de otra variable. En el gráfico Apilado, se representan las categorías de una variable apiladas en cada una de las barras en las que se representa la otra variable.

Elegido el tipo de Gráfico de Barras, se debe seleccionar ahora si el Gráfico estará

hecho con casos, a base de resúmenes de variables o a partir de valores individuales de los casos; nosotros nos ceñiremos siempre a la primera opción dejando las otras para situaciones más complejas. Una vez elegida la primera opción pasamos a lo que es la definición del Gráfico de Barras pulsando el botón **Definir**. Hecho esto aparece la ventana de la Figura 8.6.



Como se ve en ella, a la izquierda aparece la típica caja dónde están todas las variables del fichero activo y de entre las que seleccionaremos la variable para la que llevaremos a cabo la representación gráfica. Tal selección se hace llevando a **Eje de categorías:** la variable para la que deseamos el gráfico y que en el ejemplo concreto que hemos puesto aquí ha sido el sexo. El gráfico dibuja barras de altura proporcional a una de las siguientes cantidades: Número de casos (frecuencia absoluta), porcentaje de casos, número acumulado de casos (frecuencia acumulada), porcentaje acumulado de casos y otra función de resumen como puede ser una medida de posición o de dispersión de otra variable. De entre esas cinco posibilidades hay que elegir una. Para la salida que vamos a ver se ha elegido la representación por porcentajes. Por último en la ventana se puede observar la posibilidad de usar una plantilla para aplicar todas las opciones de la misma al gráfico y los botones de **Títulos y Opciones**, que permiten, respectivamente, escribir títulos detallados para el gráfico y determinar cómo representar a los datos faltantes.

Cuando se pulsa el botón **Aceptar** aparece el gráfico en el Visor de resultados, como si fuera un resultado más. En este caso, el resultado tiene algunas particularidades que describiremos a continuación observando la Figura 8.7. Lo primero será decir que cuando se



Figura 8.7 Salida de un Diagrama de Barras para el sexo del

pincha sobre el gráfico en el Visor de resultados, aparece el recuadro en el que está integrado el gráfico y en él una serie de puntos en su contorno que permiten agrandarlo y achicarlo; para ello colocado el ratón sobre uno de los puntos y pulsado el botón izquierdo basta con desplazarlo para conseguir un cambio de tamaño proporcional al tamaño original. Como se ve en las Figuras 8.7 y 8.8, en el gráfico aparece un cuadro o marco donde, dentro y alrededor del cual aparecen los diferentes componentes de dicho gráfico. Esta idea de los componentes del gráfico es básica para entender la edición del mismo; en efecto, cuando decidamos editar el gráfico (en la forma que ahora citaremos) SPSS nos permitirá la edición parcial de cada uno de sus componentes siendo la modificación deseada la suma de las modificaciones parciales. Pero, antes de entrar en la ventana de edición, decir que cuando se selecciona un gráfico y seleccionamos el menú contextual para él (recuérdese con el botón de la derecha) podremos realizar una copia del gráfico al Portapapeles de Windows y de ahí recuperarlo en la aplicación que deseemos como ha ocurrido en el caso de este manual.

En la Figura 8.7 se ven las barras correspondientes a hombres y mujeres (las dos modalidades del sexo del paciente) con altura proporcional al porcentaje de tales modalidades en nuestra muestra. En el eje horizontal aparece la etiqueta de la variable y las etiquetas de los valores de la variable; en el eje vertical, el porcentaje y la escala de porcentajes, obtenida automáticamente según los porcentajes de las modalidades, desde 0 hasta 60. Todas estas características, y otras que no aparecen son susceptibles de ser modificadas con el Editor de Gráficos que comentaremos someramente. La forma de acceder al Editor de Gráficos es a través del menú contextual o, simplemente, haciendo doble clic sobre el gráfico. La ventana del **Editor de Gráficos** aparece en la Figura 8.8.

Lo primero que hay que observar es que el Editor de gráficos es una ventana a todos los efectos, lo mismo que la ventana del Editor de Datos o la ventana del Visor de resultados. Como toda ventana muestra unos menús y una barra de herramientas análogos a los de los otros editores, salvo que ahora están pensados para editar los gráficos.

Debajo de los menús aparecen los iconos que nos sirven para acceder de manera rápida a determinadas opciones, más usuales, de los menús. Si el usuario pasa el ratón por ellos, se despliega una pequeña leyenda con el título de la opción. Por último indiquemos que la forma más rápida de acceder a las opciones de modificación es pulsar (sobre la componente que se desea modificar) dos veces con el ratón y automáticamente se despliega la ventana de modificación oportuna.

Para editar un componente del gráfico (barras, ejes, títulos, ...), en primer lugar hay que seleccionarlo (quedará remarcado dicho componente) y luego hacer los cambios oportunos utilizando los menús desplegables, la barra de herramientas, el menú contextual o la ventana de propiedades (que se despliega haciendo doble clic sobre el componente, o con el icono correspondiente de la barra de herramientas). Cuando se termine la edición del gráfico, cerrar la ventana del editor y volvemos al **Visor SPSS**.



Figura 8.8 Ventana del editor de gráficos

8.2.2. Gráficos de Líneas.

Los gráficos de líneas no son más que polígonos de frecuencias que se pueden representar para una, dos o más variables.

En la primera ventana del Gráfico de líneas se debe elegir entre los gráficos **Simple**, **Múltiple y de Líneas Verticales.** El gráfico **Simple** permite un polígono de frecuencias para una variable. En el **Múltiple** se pueden hacer varios polígonos de frecuencias simultáneamente (gráfico compuesto). En el caso de gráficos de **Líneas verticales** se lleva a cabo un gráfico más complejo que no describiremos aquí. El resto de la ventana es como en los gráficos de barras.

Pulsando el botón **Definir** aparece la ventana en todo análoga a la de la Figura 8.6. Por ello todo lo dicho para el diagrama de barras es válido ahora. Lo dicho para el editor de gráficos sigue siendo válido por lo que pasaremos a un nuevo gráfico.

8.2.3. Gráficos de Sectores.

Como su propio nombre indica, permite hacer un diagrama de sectores para la distribución de una variable. La primera ventana del gráfico de sectores es muy simple, en ella aparecen las opciones que conocemos de gráficos anteriores. Pulsando el botón **Definir** aparece la ventana en todo análoga a la de la Figura 8.6. Por ello, todo lo dicho para el diagrama de barras es válido ahora. Lo dicho para el editor de gráficos sigue siendo válido, salvo que aparecerá algún icono adicional, como el que permite desgajar cada porción del diagrama de sectores.

8.2.4. Histograma.

Con este procedimiento obtendremos un histograma para un carácter cuantitativo continuo. Cuando se selecciona el procedimiento se obtiene una ventana como la de la Figura 8.9. Como siempre, aparece una caja a la izquierda en la que figuran todas las variables del fichero activo, de entre las que se seleccionará la variable para la que se va a hacer el histograma; en nuestro caso particular la variable es el peso. Pinchando sobre **Mostrar curva normal** se representará la curva Normal teórica que se ajustaría al histograma. Por último, lo referente a Plantilla y al botón **Títulos...** ya se ha comentado en un gráfico anterior. Elegidas todas las opciones, pulsando el botón **Aceptar**, se mostrará el histograma como el de la Figura 8.10, que es explicativo por sí sólo y del que sólo cabe destacar que SPSS ha creado los rectángulos centrados en pesos de 5 en 5 unidades.





Figura 8.9 Ventana del Histograma de SPSS

9. Estimación y tests de hipótesis básicos con SPSS.

Los típicos problemas de estimación puntual de una media y de una proporción se resuelven en SPSS con los procedimientos descriptivos que hemos visto hasta ahora. Sin embargo la estimación por intervalos y los tests de hipótesis se resuelven por otros procedimientos. En este apartado veremos algunos de ellos.

9.1. Procedimiento Explorar.

El procedimiento **Explorar** se invoca como Analizar \rightarrow Estadísticos Descriptivos \Rightarrow Explorar.... El procedimiento **Explorar** genera estadísticos de resumen y representaciones gráficas, bien para todos los casos o bien de forma separada para grupos de casos. Existen numerosas razones para utilizar este procedimiento: para inspeccionar los datos, identificar valores atípicos, obtener descripciones, comprobar supuestos y caracterizar diferencias entre subpoblaciones (grupos de casos). La inspección de los datos puede mostrar que existen valores inusuales, valores extremos, discontinuidades en los datos u otras peculiaridades. La exploración de los datos puede ayudar a determinar si son adecuadas las técnicas estadísticas que está teniendo en consideración para el análisis de los datos. La exploración puede indicar que se necesita transformar los datos si la técnica necesita una distribución normal, o bien que el usuario debe utilizar pruebas no paramétricas.

Cuando se invoca aparece la ventana de la Figura 9.1. En esta ventana aparece una caja,

🙀 Explorar		×
 Image: Provide the second seco	Dependientes:	Aceptar Pegar <u>R</u> establecer Cancelar Ayuda
 ♣ acfis ♠ presencia de re ▲ Mostrar ④ Ambos ○ Estadísticos (Etiquetar los <u>c</u> asos mediante: Gráficos Estadísticos	. <u>O</u> pciones

Figura 9.1 Ventana del procedimiento Explorar de SPSS.

a la izquierda, con las variables del fichero activo de entre las que se seleccionarán las variables con las que se llevarán a cabo los análisis. A la derecha aparecen tres cajas distintas: **Dependientes**, en ella se colocarán las variables para las que se calcularán los estadísticos oportunos, en nuestro caso "imc" sólo; **Factores**, en la que se dispondrán las variables según las que se agruparán los estadísticos calculados para las variables dependientes, en nuestro caso el factor elegido es el sexo, por lo que aparecerán las medidas de resumen de "imc" para los diferentes niveles del factor "sexo"; **Etiquetar los casos mediante**, permite indicar una variable que contendrá las etiquetas que deseemos que muestren algunos casos particulares (valores extremos por ejemplo) en algunas de las gráficas que produce **Explorar.**

Debajo de la caja de las variables aparece el epígrafe **Mostrar**, que nos permite especificar si se desea obtener como salida del procedimiento sólo los estadísticos, sólo los gráficos del mismo, o ambas cosas simultáneamente; las tres alternativas son incompatibles. Abajo a la derecha aparecen los tres botones que permiten especificar con detalle los resultados que se desean: **Estadísticos...**, **Gráficos...** y **Opciones...**

Si pulsamos el botón **Estadísticos...** se abre una nueva ventana donde podemos marcar diferentes opciones sobre los estadísticos que queramos calcular (siempre que la variable aluda a una cantidad). Comentemos sólo la opción **Descriptivos**, que nos permite (además del cálculo de la media, mediana, desviación típica y otras medidas) obtener un intervalo de confianza para la media, con una confianza que podemos especificar. Con el botón **Gráficos...** aparece una nueva ventana donde podemos elegir el tipo de gráfico exploratorio que deseamos. Mencionemos aquí sólo la opción **Gráficos con prueba de normalidad**. Si la marcamos, realiza tests de normalidad (el de Kolmogorov-Smirnov y el de Shapiro-Wilk, este último sólo si el tamaño de muestra no excede de 50) y gráficos de normalidad (que no veremos). Finalmente, el botón **Opciones...** hace desplegar una ventana con tres opciones relativas al manejo de los datos faltantes. Éstas se explican en la Tabla 9.1.

La salida de este procedimiento puede ser muy compleja según las opciones que se hayan elegido, aunque la gran mayoría de los resultados aparecerán en forma de tablas de fácil comprensión. El gran volumen de resultados hace que no podamos sacarlos aquí por lo que recomendamos al lector que los obtenga y los revise para lo que no tendrá dificultad.

9.2. Procedimiento Binomial.

El procedimiento **Binomial** permite realizar el test de hipótesis para una proporción, contrastando si la proporción difiere significativamente de un valor p_0 especificado.

La forma de invocar el procedimiento es: *Analizar* \rightarrow *Pruebas No Parámetricas* \rightarrow *Binomial...* En la ventana que aparece, como en otras muchas ocasiones, pasaremos a la caja de la derecha las variables (que deben ser numéricas) para las que queremos realizar el test. En la pequeña caja contrastar proporción colocaremos el valor p₀ que deseamos contrastar. SPSS

Introducción al manejo del SPSS 12.0

Epígrafe	Opción	Acción		
	Excluir casos según lista	Excluye los casos que tienen valores perdidos en cualquiera de las variables utilizadas en cualquiera de los análisis.		
Valores	Excluir casos según pareja	Excluye del análisis los casos que tengan valores perdi en cualquiera de las variables de una pareja implicada e cálculo de un estadístico específico.		
perdidos	Mostrar valores	Trata los valores perdidos en las variables de factor como una categoría separada. Se generan todos los resultados para esta categoría adicional. Las tablas de frecuencias incluyen categorías para los valores perdidos.		

Tabla 9.1. Opciones del procedimiento Explorar

realizará un test de la proporción para cada variable introducida, pero todos con el mismo p_0 . En la parte inferior izquierda de la ventana aparece **Definir la dicotomía**, que tiene dos opciones. La primera opción es para variables que tienen sólo dos valores posibles: el valor más pequeño determinará el grupo 1, y el más grande el grupo 2. La segunda opción es para variables con cualquier número de valores distintos, pudiéndose especificar un punto de corte: los valores iguales o más pequeños que el punto de corte determinarán el grupo 1, y los más grandes el grupo 2. Esta definición de la dicotomía tiene que ser idéntica para todas las variables introducidas. En cualquier caso, y para cada variable, SPSS contrastará si la proporción de individuos o casos pertenecientes al grupo 1 difiere significativamente del valor p_0 especificado.

En la ventana del procedimiento aparece el botón de **Opciones...**, donde podemos solicitar cálculos descriptivos de las variables y especificar cómo tratar los valores perdidos. Para esto último existen dos opciones: **Excluir casos según prueba** (excluye los casos con valores perdidos en la variable implicada en el contraste) y **Excluir casos según lista** (excluye los casos con valores perdidos en cualquiera de las variables introducidas en el procedimiento Binomial). Los resultados del procedimiento se dan en una tabla, donde aparecen, para cada test, el grupo 1 y el grupo 2, el valor p_0 de prueba y el valor P del test.

10. Tests de comparación de medias con SPSS.

El experimento comparativo es uno de los más frecuentes en Medicina y dentro de él, el de comparación de dos medias. En este capítulo nos dedicaremos a explicar cómo se puede hacer el test de comparación de dos medias con muestra independientes y con muestras apareadas, suponiendo las variables aleatorias Normales y sin necesidad de suponerlo. Los procedimientos a emplear serán, en algún caso, más generales de lo que se necesita.

10.1. Procedimiento Prueba T para muestras independientes.

🉀 Prueba T para muestra	as independientes	×
 in acfis in presencia de retino in netro in neuro in ca in ca in cr in pthi in pthi in bmdcue in szl24 	Contrastar variables: Image: Provide the second	Aceptar Pegar Restablecer Cancelar Ayuda

Realiza el test de comparación de dos medias con muestras independientes, suponiendo que las varianzas son iguales (test de la t-Student) y suponiendo que las varianzas son distintas (test de Welch). Supone que las variables aleatorias son Normales. Además, para cada muestra calcula: tamaño de la muestra, media, desviación típica y error típico de la media; calculando también la diferencia de medias, el error típico de esa diferencia y el intervalo de confianza para ella (puede especificarse el nivel de confianza).

Definir grupos		×
⊙ <u>U</u> sar valores es	pecificados	Continuar
Grupo <u>1</u> :	1	Cancelar
Grupo <u>2</u> :	2	Ayuda
C Punto de <u>c</u> orte:		

La

forma de ejecutar el procedimiento es Analizar →Comparar medias →Prueba T para muestras independientes. Hecho esto aparece una ventana como la de la Figura 10.1. Como siempre, a la izquierda aparecen las variables del fichero activo, de entre las que se seleccionarán las variables para las que se van a hacer el contraste de hipótesis (eso indica que pueden ser varias). A la derecha está la caja Contrastar variables:, que muestra las variables seleccionadas para llevar a cabo los contrastes de hipótesis; en este caso la variable "hba1c", hemoglobina glicoxilada. Inmediatamente debajo está el recuadro de la Variable de agrupación: que determina los dos grupos que se desean comparar; esta variable puede ser numérica o de cadena. Elegida la variable de agrupación habremos de seleccionar de ella los códigos, o valores, que definen los dos grupos que se desean comparar; ello se hace pinchando el botón Definir

grupos..., lo que despliega una ventana como la de la Figura 10.2. En ella se ve que se pueden definir los dos grupos de dos maneras diferentes; la primera eligiendo Usar valores especificados, y tras ello escribir, en las cajas correspondientes, los códigos que identifican al grupo 1 y al grupo 2; la segunda manera de especificar los dos grupos es pinchar en punto de corte y escribir al lado el mismo, de manera que el primer grupo estará formado por los casos cuvos valores, en la variable de agrupación, sean menores al punto de corte y el segundo por aquellos casos cuyos valores, en la variable de agrupación, sean mayores o iguales que el punto de corte. Si la variable de agrupación es de cadena no se puede usar la opción **Punto de corte**. Identificados los grupos y pulsando Continuar se vuelve a la ventana principal. Por último, en **Opciones** se especifica lo mismo que ya hemos visto en el apartado 9.3. Ejecutando ya el procedimiento con el botón Aceptar, se obtendrá un resultado como el de la Figura 10.3.

Como se ve, lo primero en obtenerse son, por grupos, las medidas básicas de resumen para ellos; en nuestro caso los grupos son el 1 y el 2 del tabaco, que representan a los fumadores y a los no fumadores respectivamente. Inmediatamente después aparece el test

			TTuc		
Estadísticos del grupo					
	TABACO 1=Si; 2=No	N	Media	Desviación típ.	Error típ. de la media
HBA1C	Si	41	8.7439	1.7683	.2762
	No	53	8.4264	1.8265	.2509

Prueba	Т
tadísticos dol	ar

			Prue	ba de	muestra	s indep	endientes	5		
Prueba de Levene para la igualdad de varianzas				Pru	ieba T p	ara la ig	ualdad (de medias	1	
		F	Sig.	Т	gl	Sig. (bilat)	Dif. de medias	Error típ de la dif.	Interv confianz d Inferior	valo de za para la if. Superior
	Se han asumido varianzas iguales	.097	.756	.847	92	.399	.3175	.3747	4267	1.0616
HBA1C	No se han asumido varianzas iguales			.851	87.451	.397	.3175	.3731	4241	1.0590

Figura 10.3. Resultados del procedimiento Prueba T para muestras independientes

propiamente dicho; como se sabe, cuando se hace un test de comparación de medias para variables aleatorias Normales y con muestras independientes, hay que hacer previamente un test de homogeneidad de varianzas; esto es lo que aparece primero en la segunda tabla de los resultados; en este caso el resultado del test de Levene da una F_{exp} =0.097, con P=0.756, que al estar muy lejos de la significación (recuérdese que para este test es P<0.20), nos lleva a no poder rechazar la igualdad de varianzas. Como SPSS no sabe nuestro criterio, ha calculado el test de homogeneidad de medias, suponiendo que las varianzas son iguales, primera fila, y

suponiendo que son distintas, segunda fila; seremos nosotros los que deberemos elegir. En nuestro caso los resultados adecuados son: $t_{exp}=0.847$, 92g.l., P=0.399. Por último, obsérvese que se dan los intervalos de confianza para la diferencia de medias, tanto en el caso de varianzas iguales como en el caso de varianzas distintas. Si hubiera más variables aparecerían sus resultados a continuación.

10.2. Procedimiento Prueba T para muestras relacionadas.

Realiza el test de comparación de dos medias con muestras apareadas (relacionadas les llama SPSS), llevando a cabo determinados cálculos, que veremos, sobre cada pareja de variables implicada. Supone que la diferencia de las variables aleatorias es Normal.



La

forma de ejecutar el procedimiento es Analizar \rightarrow Comparar medias \rightarrow Prueba T para muestras relacionadas. Hecho esto, aparece una ventana como la de la Figura 10.4. A la izquierda está la caja con las variables presentes en el fichero activo para seleccionar de ellas la pareja que nos permitirá realizar el test de comparación de medias con muestras apareadas. Al seleccionar ambas variables, pinchando sobre ellas, aparecerán en la zona de Selecciones actuales; en nuestro caso la pareja es (szl24,sztri). Para acabar de seleccionar la pareja y ejecutar con ella el procedimiento bastará con pulsar el botón ϕ y aparecerá la pareja en la caja de Variables relacionadas. La operación la podemos repetir para cualquier otro par de variables para el que queramos realizar la comparación. En las Opciones aparece el mismo contenido ya visto en el apartado 9.3. Por último, pulsando el botón Aceptar se ejecuta el procedimiento dando lugar a un resultado como el de la Figura 10.5.

En la primera tabla aparecen las medidas de resumen que calcula para las variables implicadas en la comparación. En la siguiente tabla aparece el coeficiente de correlación de las dos variables. En la última tabla aparecen los resultados propios de la comparación; primero aparece la media de las diferencias, seguida de su desviación típica, el error estándar y un intervalo de confianza para la diferencia entre las dos medias, y a continuación aparece el resultado del test: t_{exp} =14.378, 93 g.l., P<0.001.

Г

Estadísticos de muestras relacionadas								
Media N Desviación típ. Error típ. de la med								
SZL24	.8073	94	.1400	1.444E-02				
SZTRI	8956	94	1.2111	.1249				
5	ZL24 ZTRI	Media ZL24 .8073 ZTRI 8956	Media N ZL24 .8073 94 ZTRI 8956 94	Media N Desviación típ. ZL24 .8073 94 .1400 ZTRI 8956 94 1.2111				

		N	Correlación	Sig.
Par 1	SZL24 v SZTRI	94	.494	.000

	Prueba de muestras relacionadas								
Diferencias relacionadas									
	MediaDesviaciónError típ. de laIntervalo de confianza para la diferencia		t	gl	Sig. (bilat.)				
			up.	media	Inferior	Superior			
Par 1	SZL24 - SZTRI	1.7030	1.1484	.1184	1.4678	1.9382	14.378	93	.000

Figura 10.5. Resultados del procedimiento Prueba T de muestra relacionadas

10.3. Gráfico Barras de error.



Con frecuencia, para acompañar al test de comparación de medias, se da una gráfica en la que se representa, para cada grupo, su media y una línea o barra (con centro en la media) que

representa una medida de error asociado a dicha media. SPSS tiene un procedimiento que permite hacer tales representaciones gráficas.

La forma de invocar el procedimiento es: *Gráficos* →*Barras de error...*. Entonces se despliega una ventana en la que tendremos que elegir entre un gráfico simple (el que aquí veremos) o un gráfico agrupado. El gráfico simple contiene una única barra de error para cada valor o categoría de la variable del eje de categorías. Los datos del gráfico pueden ser resúmenes para grupos de casos o para distintas variables; la primera situación es la ideal para el test de comparación de medias con muestras independientes, y será la que veamos aquí. Pulsando el botón Definir se pasa a la ventana en la que realmente se define el gráfico, tal ventana aparece en la Figura 10.6 para el gráfico simple.

A la izquierda aparece la típica caja con todas las variables del fichero activo de entre las que seleccionaremos, tanto la variable numérica, para la que se calcularán y dibujarán las medias (se colocará en la caja **Variable:**), como aquella que establecerá las categorías que se muestran en el gráfico (se colocará en la caja **Eje de categorías:**). El gráfico representará la



Consumo de tabaco (1=Sí, 2=No)

Figura 10.7. Resultados del procedimiento Gráficos de Barras de Error

media y barras para cada valor de esta variable. Las barras de error que parten de la media de cada categoría o grupo pueden representar tres entidades distintas: Intervalo de confianza para la media, Error típico de la media y Desviación típica; en la caja de **Las barras representan** se escogerá la opción, pudiendo elegirse, en el caso del intervalo de confianza para la media, la confianza que se desea; en el caso del error estándar de la media y de la desviación típica, puede elegirse el multiplicador para ellas. Como en otros gráficos, se puede simplificar la tarea de configurarlos sin más que tomar la plantilla de otro gráfico existente. Asimismo, los títulos del gráfico y las opciones cumplen la función que se ha citado en otros casos. En nuestro ejemplo concreto hemos pedido el gráfico de barras de error para la variable hemoglobina glicoxilada (hba1c) en las categorías 1 y 2 de la variable "tabaco"; las barras representan el intervalo de confianza para la media al nivel del 95%. La salida está en la Figura 10.7, que se comenta por sí sola.

10.4. Procedimiento Pruebas no-paramétricas para dos muestras independientes.

Los procedimientos que hemos explicado hasta ahora, exigían que las variables fueran Normales; por eso, presentamos ahora las pruebas que no exigen tal condición, que, como se sabe, son pruebas no-paramétricas. La forma de invocar la prueba para comparar la distribución de dos variables medidas en sendas muestras independientes es: *Analizar* \rightarrow *Pruebas noparamétricas* \rightarrow *2 muestras independientes.* Hecho esto, se despliega una ventana como la de la Figura 10.8. La primera mitad de esta ventana es en todo similar a la de su correspondiente prueba T (apartado 10.1). En la segunda mitad de la ventana habrá que elegir el test noparamétrico de entre los cuatro tests que se nos ofertan; el preferido es el test de la U de Mann-Whitney (que es el mismo test de Wilcoxon para muestras independientes) puesto que los otros tres ofrecidos son menos potentes que éste. La significación del test empleado será exacta siempre que el número total de casos, entre las dos muestras, no exceda de 40, y aproximada en caso contrario. El botón **Opciones** nos permite calcular medidas descriptivas para ambas muestras y elegir la forma de tratamiento que queremos dar a los datos faltantes, igual que se vio en el apartado 9.3.



Figura 10.8. Ventana principal de Pruebas no-paramétricas para dos muestras independientes.

En la Figura 10.9 exponemos los resultados obtenidos para el análisis propuesto en la Figura 10.8, en la que se han pedido como opción los estadísticos descriptivos. La primera tabla que aparece es la de los estadísticos descriptivos básicos de las dos variables implicadas en el análisis, "hba1c" y "tabaco". Debajo, a la izquierda, aparece la suma de rangos y los rangos medios en los dos grupos. Por último aparecen, a la derecha en la disposición que hemos hecho aquí, los resultados del test de Mann-Whitney y del de Wilcoxon (que son equivalentes, con su nivel de significación: en este caso asintótico puesto que había 94 casos en total).

	Pruebas no paramétricas							
Estadísticos descriptivos								
	Ν	Media	Desviación típica	Mínimo	Máximo			
HBA1C	94	8.5649	1.7987	4.60	13.80			
TABACO	94	1.56	.50	1	2			

Rangos						
	Tabaco	N	Rango promedio	Suma de rangos		
HBA1C	1	41	50.30	2062.50		
	2	53	45.33	2402.50		
	Total	94				

Prueba de Mann-Witney

Estadísticos de contraste (a)					
	HBA1C				
U de Mann-Witney	971.500				
W de Wilcoxon	2402.500				
Z	877				
Sig. Asintót. (bilat) .380					
a Variable de agrupación: TABACO					

Figura 10.9. Resultados del procedimiento Pruebas no-paramétricas con dos muestras independientes.

10.5. Pruebas no-paramétricas para dos muestras relacionadas.

Igual que en el caso de variables aleatorias Normales vimos el test de comparación de dos medias para muestras apareadas, veremos ahora la alternativa no-paramétrica para dicho test. La forma de invocar la prueba para comparar la distribución de dos variables medidas en sendas muestras apareadas es: *Analizar* \rightarrow *Pruebas no-paramétricas* \rightarrow *2 muestras relacionadas.* Hecho esto, se despliega una ventana como la de la Figura 10.10, en todo análoga a la que vimos para a prueba T con muestras relacionadas. Su funcionamiento es el mismo que se vio. La única diferencia está en que ahora hay que elegir la prueba a llevar cabo, que en términos generales será el test de Wilcoxon (puesto que es más potente que el de los signos y el de McNemar el cual es un test válido para el caso de que la variable dependiente sea dicotómica, con lo que es realmente un test de comparación de proporciones). Por último, en las Opciones aparece el mismo contenido que en el caso de dos muestras independientes (apartado 10.4).



Pulsando el botón **Aceptar** se ejecuta el procedimiento, dando lugar a un resultado como el de la Figura 10.11. En la primera tabla aparecen el número, rango medio y suma de

Pruebas no paramétricas

Prueba de los rangos con signo de Wilcoxon

Rangos						
		Ν	Rango promedio	Suma de rangos		
	Rangos negativos	85(a)	50.28	4273.50		
	Rangos positivos	8(b)	12.19	97.50		
5211(1-52L24	Empates	1(c)				
	Total	94				
a SZTRI < SZL2	4					
b SZTRI > SZL2	4					
c SZL24 = SZTR	RI					

Estadísticos de contraste(b)				
	SZTRI - SZL24			
Z	-8.001(a)			
Sig. asintót. (bilateral) .000				
a Basado en los rangos positivos.				

b Prueba de los rangos con signo de Wilcoxon

rangos tanto de las diferencias negativas como de las positivas. El resultado del test, con su probabilidad de significación, figura en la última tabla; en este caso se da el resultado asintótico, puesto que el tamaño de la muestra desborda a 30.

11. El test Chi-cuadrado con SPSS.

La comparación de proporciones, ya sea con muestras independientes o con muestras apareadas, y la asociación de caracteres cualitativos son problemas estadísticos que se resuelven con el uso del estadístico chi-cuadrado; al procedimiento de SPSS dedicado a tal estadístico es al que dedicaremos este apartado.

11.1. Procedimiento Tablas de contingencia.

El procedimiento para hacer los análisis que se han citado es el procedimiento **Tablas de** contingencia. Tal procedimiento se invoca como: *Analizar →Estadísticos Descriptivos →Tablas de Contingencia*. Cuando se lleva a cabo tal acción, se abre una ventana como la de la Figura 11.1.

Como se ve, a la izquierda aparece la típica caja con el conjunto de las variables presentes en el fichero activo. De entre ellas se elegirán las que irán por filas (se colocarán en la caja **Filas**), las que irán por columnas (se colocarán en la caja **Columnas**), y, si acaso, las que determinarán las categorías para las que se obtendrán las tablas de contingencia (se colocarán en la caja **Capa 1 de 1** y constituirán las variables de la capa 1). El procedimiento obtendrá una tabla de contingencia para cada combinación de dos variables, una de **filas** y otra de **columnas**, y si existen variables en la **capa 1**, dichas tablas de contingencia las obtendrá para cada categoría de cada una de estas variables; por ejemplo, si "tabaco" es la única variable de la **capa 1**, se obtendrán las tablas de contingencia para fumadores y no fumadores. En general se pueden definir diferentes capas de variables, lo que da más posibilidades al procedimiento, pero lo complica (no lo veremos). En el caso de la Figura 11.1 deseamos comprobar si la distribución del grado de retinopatía que sufren los pacientes es la misma en hombres que en mujeres; estando por tanto en la situación de un test de comparación de proporciones con dos muestras independientes. También sería éste, de forma equivalente, el problema de estudiar la asociación entre el sexo y la presencia de retinopatía.



Inmediatamente debajo de la caja de la izquierda aparecen dos opciones que pueden o no estar

seleccionadas: la opción **Mostrar los gráficos de barras agrupadas** (que muestra un gráfico de barras agrupadas para cada combinación de variables de **Filas** y **Columnas**, dentro de las categorías de cada variable de la capa 1) y la opción **Suprimir tablas** (que muestra el análisis de las tablas de contingencia, sin llegar a mostrar las propias tablas). Inmediatamente debajo aparecen los botones: **Estadísticos..., Casillas...** y **Formato...**, que hacen referencia a los estadísticos que deseamos calcular para la tabla, las cantidades que deseamos que aparezcan en las casillas y el formato en el que aparecerán las tablas. A continuación los describiremos con detalle.

Cuando se pincha sobre el botón **Estadísticos....** se despliega una ventana con muchas alternativas que, marcándolas, nos permitirán el cálculo de estadísticos o medidas aplicables a diferentes situaciones, y que, la mayoría, van más allá de un curso básico de estadística al que pretende dar cobertura este manual. Por ello, aquí sólo comentaremos algunas de ellas. La primera, **Chi-cuadrado**, nos calcula estadísticos del tipo chi-cuadrado (entre ellos, el chi-cuadrado de Pearson para una tabla de contingencia) y hace el test de Fisher en el caso de tablas 2×2 . La opción **Coeficiente de Contingencia** nos calcula dicho coeficiente en cualquier tabla de contingencia. La opción **Riesgo** nos permite obtener medidas de asociación en tablas 2×2 : razón del producto cruzado y riesgo relativo. Finalmente la opción **McNemar** lleva a cabo el test de McNemar para comparar dos proporciones con muestras apareadas (caso de una tabla 2×2).

Al pulsar el botón **Casillas...** nos aparece una ventana donde, marcando las opciones correspondientes, podemos calcular y escribir en cada casilla de la tabla: las frecuencias observadas, las frecuencias esperadas, los porcentajes por filas, los porcentajes por columnas, los porcentajes respecto al total, los residuos no tipificados (diferencia entre frecuencias observadas y esperadas) y otros residuos que no veremos aquí. Por último, la opción **Formato...** permite presentar las filas de la tabla en orden ascendente o descendente.

Tabla de contingencia sexo del paciente * presencia de retinopatía							
			presei	Total			
		no	Moderada	grave			
	hamhaa	Recuento	33	8	4	45	
sexo del	nomore	% de sexo	73.3%	17.8%	8.9%	100.0%	
paciente	maniar	Recuento	30	10	9	49	
	mujer		61.2%	20.4%	18.4%	100.0%	
Total		Recuento	63	18	13	94	
		% de sexo	67.0%	19.1%	13.8%	100.0%	

Elegidas todas las opciones y pulsando la tecla **Aceptar** se llevará a cabo el análisis solicitado. A continuación mostraremos los resultados obtenidos para varios análisis distintos.

Pruebas de chi-cuadrado								
	Valor	gl	Sig. asint. (bilateral)					
Chi-cuadrado de Pearson	2.122(a)	2	.346					
Razón de verosimilitud	2.169	2	.338					
Asociación lineal por lineal	2.058	1	.151					
N de casos válidos	94							
a 0 casillas (.0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 6.22.								

Figura 11.2. Resultados del procedimiento Tablas de contingencia.

En la Figura 11.2 aparece lo fundamental de los resultados del análisis de asociación entre el sexo del paciente y el nivel de retinopatía que ha alcanzado. En la primera tabla, por filas aparece el sexo del paciente y por columnas figura el nivel de la retinopatía. En las casillas se presentan las frecuencias observadas así como los porcentajes por filas. La segunda tabla contiene las pruebas de chi-cuadrado. En la primera fila se muestra el clásico test Chi-cuadrado de Pearson para tablas de

contingencia (los otros no los veremos aquí); también al pie de esta tabla observamos el número de casillas (de la tabla de contingencia) que tienen frecuencia esperada inferior a 5, de utilidad para comprobar la validez del test anterior.

En la Figura 11.3 aparece lo fundamental de los resultados del test de McNemar para la comparación de dos proporciones con muestras apareadas. Se dice que un individuo tiene

Tablas de contingencia

			SZCU		
			Sí	No	Total
SZTRI_R	SÍ	Recuento	8	5	13
		% del total	8,5%	5,3%	13,8%
	No	Recuento	16	65	81
		% del total	17,0%	69,1%	86,2%
Total		Recuento	24	70	94
		% del total	25,5%	74,5%	100,0%

Tabla de contingencia SZTRI_R * SZCUE_R

Pruebas de chi-cuadrado

	Valor	Sig. exacta (bilateral)
Prueba de McNemar	Valor	,027 ^a
N de casos válidos	94	

a. Distribución binomial utilizada

Figura 11.3. Resultados del test de McNemar de comparación de dos proporciones con muestras apareadas (procedimiento Tablas de contingencia)

osteoporosis en una zona cuando el "z-score" en esa zona es menor que -2.0; así se han determinado los que son osteoporóticos en el triángulo de Ward (variable SZTRI_R) y los que lo son en el cuello del fémur (variable SZCUE_R). La pregunta que se pretende responder es si el porcentaje de osteoporóticos es el mismo en ambas zonas. Para ello se lleva a cabo el test de McNemar. En la primera tabla de los resultados se dan los valores observados y los porcentajes con respecto al total. Debajo aparece el resultado del test, con la probabilidad de significación exacta calculada a partir de la distribución Binomial.

			SZCUE_R		
			Sí	No	Total
Consumo de tabaco	Sí	Recuento	15	26	41
(1=Sí, 2=No)		%	36,6%	63,4%	100,0%
	No	Recuento	9	44	53
		%	17,0%	83,0%	100,0%
Total		Recuento	24	70	94
		%	25,5%	74,5%	100,0%

Tabla de contingencia Consumo de tabaco (1=Sí, 2=No) * SZCUE_R

Pruebas de chi-cuadrado

	Valor	gl	Sig. asint. (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	4,673 ^b	1	,031		
Corrección de continuidad	3,699	1	,054		
Razón de verosimilitud	4,661	1	,031		
Estadístico exacto de Fish				,035	,027
Asociación lineal por linea	4,623	1	,032		
N de casos válidos	94				

a. Calculado sólo para una tabla de 2x2.

b. 0 casillas (,0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima espera

Estimación de riesgo

		Intervalo de confianza al 95%	
	Valor	Inferior	Superior
Razón de las ventajas para Consumo de tabaco (Sí / I	2,821	1,082	7,352
Para la cohorte SZCUE_R = 1,00	2,154	1,050	4,421
Para la cohorte SZCUE_R = 2,00	,764	,588	,993
N de casos válidos	94		

Figura 11.4. Resultados del procedimiento Tablas de Contingencia.

Por último, en la Figura 11.4 aparecen los resultados de un test de asociación consistente en saber si el tabaco es un factor de riesgo para la osteoporosis en el cuello del fémur. En primer lugar aparece la tabla de contingencia para ambos caracteres, donde por filas se ha colocado el factor de riesgo (el consumo de tabaco) y por columnas la enfermedad (la osteoporosis en el cuello del fémur). Esta disposición no es azarosa, sino que corresponde a la disposición que utiliza SPSS, obligándonos a ella, para que tenga sentido el cálculo e interpretación de los riesgos que después veremos. Pero la disposición obliga también a que la primera categoría del factor de riesgo, sea la presencia de éste, y que la presencia de la enfermedad sea la primera categoría de esta variable. En segundo lugar se muestra la tabla con los resultados de los tests de hipótesis que podemos utilizar en este tipo de problemas. Respecto a la correspondiente tabla de la figura 11.2, se dan aquí dos nuevos resultados por tratarse de una tabla 2×2: el test chi-cuadrado con corrección de continuidad (con la corrección de Yates) y el test exacto de Fisher.

Hecho el test para ver si existe asociación, éste da, como vemos, significativo, P=0.035 según el test exacto de Fisher; por ello tiene sentido el cálculo de medidas de la fuerza de la asociación entre el factor de riesgo y la enfermedad, las cuales fueron solicitadas marcando la opción Riesgo del botón **Estadísticos...** Estas medidas aparecen en la última tabla de los resultados, (figura

11.4). En la primera línea de dicha tabla se da razón del producto cruzado (también llamada razón de ventajas). Justo debajo aparecen las nuevas medidas de estimación de riesgo con nombres que no son del todo correctos. La primera de éstas, etiquetada como "para la cohorte zscue_r = si", es, simplemente, el riesgo relativo de padecer la enfermedad cuando se está expuesto al factor de riesgo (ésta es la medida que comúnmente llamamos riesgo relativo). La segunda es el riesgo relativo de no padecer la enfermedad cuando se está expuesto al factor. Como se ven en la tabla, para las tres medidas se dan sus correspondientes intervalos de confianza. Finalmente hemos de hacer dos precisiones: en primer lugar, la palabra cohorte que aparece en las dos últimas medidas de riesgo, está claramente mal usada pues está referida a la enfermedad y no al factor de riesgo, y todo ello suponiendo que estemos en presencia de un seguimiento; en segundo lugar, si bien el lector no debe preocuparse con la razón del producto cruzado, puesto que es válida en todo tipo de estudios, no ocurre lo mismo con los riesgos relativos de los que el lector debe controlar su aplicación rigurosa, cosa que no hace SPSS.

12. Regresión y correlación lineal simple con SPSS

Pasemos ahora a estudiar cómo se puede resolver, con SPSS, el problema de si dos caracteres cuantitativos están asociados o no. El problema tiene, al menos, dos facetas: saber si están relacionados y conocer la forma de la relación, lo que nos permitiría predecir una variable en función de la otra. Nosotros estudiaremos aquí la regresión lineal simple y la correlación lineal simple. Empezaremos, como debe empezarse cualquier problema de regresión lineal simple y por último el procedimiento de correlación. En todo lo que vayamos presentando, la regresión lineal múltiple aparecerá al lado, con un conjunto importante de comandos, no obstante no entraremos en ellos pues no es el objeto de este manual.

12.1. Gráficos de Dispersión (nubes de puntos).

En todo problema de regresión y correlación, el primer paso consiste en dibujar la nube de puntos, con objeto de detectar posibles no-linealidades en la relación o alguna otra falla en las hipótesis del modelo de regresión lineal. Por ello, en primer lugar, veremos el procedimiento que permite dibujar nubes de puntos, que SPSS denomina Diagrama de Dispersión.



La forma de invocar el procedimiento es: *Gráficos* \rightarrow *Dispersión...*. Al hacer esto, aparece una ventana como la de la Figura 12.1. Como se ve, se puede elegir entre cuatro opciones que son muy distintas. La primera opción, **Simple**, dibuja un gráfico que representa dos variables en dos ejes de escala, una variable define el eje horizontal y la otra define el eje vertical. Las otras opciones dan lugar a gráficos más complejos, que no veremos en este manual.

Cuando se selecciona la opción **Simple** y se pulsa el botón **Definir** aparece una ventana como la de la Figura 12.2. Como en casi todas las ventanas, aparece a la izquierda, una caja con todas las variables del fichero activo. A la derecha de ella aparecen dos primeras cajas en las que caerán las variables dependiente (**Eje Y**) e independiente (**Eje X**), siendo, en nuestro caso, la variable independiente la "imc", mientras que la dependiente será la "sztri". Debajo de ellas aparece la caja, **Establecer marcas por**, donde se coloca la variable que nos permite establecer una representación gráfica en la que los casos se presentan de manera particular, según el grupo al que pertenecen, no siendo obligatoria la selección de esta variable; debajo de ella aparece la

otra caja, **Etiquetar los casos mediante...**, donde se coloca la variable que nos permite establecer una etiqueta individual para cada caso, no siendo obligatoria la selección de esta variable; la etiqueta de cada caso no aparecerá en el gráfico a menos que se active la opción de etiquetar caso en el Editor de Gráficos. Como se ve en la Figura 12.2, también aparecen opciones para usar una plantilla gráfica procedente de otro gráfico, guardado previamente en un archivo, colocar Títulos particulares al gráfico y establecer las opciones que ya hemos visto en

🃆 Diagrama de dispersión	n simple		×
 		Eje ⊻: I 🛞 sztri	Aceptar
() ← ca () ← p () ← ca		Eje <u>X</u> :	P <u>eg</u> ar <u>R</u> establecer
the second		Establecer marcae per:	Cancelar
	\rightarrow	E <u>s</u> tablecer marcas por.	Ayuda
szcue retinn	\rightarrow	Etiguetar los casos mediante:	
Plantilla Usar las especificacione <u>Archivo</u>	es gráficas (de:	
		<u>I</u> ítulos <u>O</u> pciones	

otros procedimientos.

Cuando se ha establecido el conjunto de características que deseemos que tenga el gráfico, se pulsará **Aceptar** y nos aparecerá una gráfica como la de la Figura 12.3. En ella se puede ver que la relación entre las dos variables no muestra ningún síntoma de curvilinealidad, en todo caso, muestra un ligero decrecimiento de la variabilidad de la variable dependiente conforme aumentan los valores de la independiente; en resumen, la relación entre ambas variables parece clara, cosa que no se podrá afirmar hasta que hagamos el análisis de regresión y/o de correlación.

12.2. El procedimiento Regresión lineal.

Si bien las nubes de puntos representan un gráfico esencial para los estudios de la relación entre caracteres cuantitativos, los estudios que son concluyentes para determinar si existe relación entre las variables, son los de regresión y/o correlación que son los que veremos a partir de ahora.



Figura 12.3. Nube de puntos producida por el procedimiento Diagrama de Dispersión Simple

El procedimiento que realiza los cálculos para una regresión lineal es el de **Regresión Lineal**, que se invoca como Analizar - Regresión - Lineal. Una vez que se ha invocado el procedimiento se mostrará una ventana como la de la Figura 12.4. En ella, lo primero que se ve, a la izquierda, es la caja de variables del fichero activo, de entre las que se seleccionarán las variables dependiente e independiente de la regresión lineal simple. A la derecha aparece la caja que albergará la variable dependiente (la Y), que en nuestro caso es "sztri". Inmediatamente debajo aparece el espacio bloque, que sólo se empleará en el caso de regresión múltiple. Inmediatamente debajo aparece la caja de las variables **independientes** (las X), pues el programa está pensado para la regresión múltiple, pero como nosotros sólo nos ocuparemos de la simple, ahí colocaremos una única variable; en el ejemplo que nos ocupa la variable ha sido "imc", con lo que nosotros pretendemos saber si ambas variables están asociadas y además estimar la ecuación que presenta la "sztri" en función del "imc". Justo debajo aparece una caja que hace referencia al Método, utilizado sólo en el caso de la regresión lineal múltiple. Debajo aparece la caja de Variable de selección, que nos permite elegir una variable de selección que limite el análisis a un subconjunto de casos que tengan un valor particular para esta variable; cuando en esa caja hava una variable se iluminará el botón de **Regla** que, pulsado, nos permitirá establecer la condición que seleccionará los casos que serán incluidos en el análisis. La caja Etiquetas de caso designa una variable para identificar los casos o puntos en los gráficos. Debajo aparecen los botones de diferentes opciones que veremos a continuación.

🚜 Regresión lineal		×
 Icohol Ingca Acfis 	Dependiente:	Aceptar P <u>e</u> gar
presencia de retino presencia de retino presencia de retino	Anterior Bloque 1 de 1 Siguiente	<u>R</u> establecer Cancelar
	€ imc	Ayuda
ter p ter pthm	Método: Introducir	
 ♥ pthi ♥ bmdcue ● szl24 	⊻ariable de selección: ▶ Reg[a	
 Image: A state of the state of	Etiquetas de <u>c</u> aso:	
MC <u>P</u> >>	Estadísticos <u>G</u> ráficos <u>Gu</u> ardar <u>G</u>	<u>)</u> pciones
	Po <u>n</u> deración MCP:	

El botón **Estadísticos...** sirve para solicitar resultados estadísticos opcionales. Al pulsarlo se nos muestra una tabla con muchas opciones que podemos marcar, entre ellas: **Estimaciones** (calcula los coeficientes de la recta de regresión), **Intervalos de confianza** (para los coeficientes de la recta, al 95% de confianza) y **Descriptivos** (calcula media y desviación típica para las variables y el coeficiente de correlación lineal entre ambas con su prueba de significación).

El botón **Gráficos** solicita gráficos opcionales, que no veremos en este manual. El botón **Guardar** permite guardar las predicciones, los residuos y medidas relacionadas, como nuevas variables que se añaden al archivo de datos de trabajo. Cuando se pulsa aparece una ventana en la que figuran múltiples opciones, algunas de las cuales se describen en la tabla 12.1.

Epígrafe	Opción	Acción
Valores pronosticados	No tipificados	Calcula el valor que predice el modelo para la variable dependiente en cada valor de la independiente.
Residuos	No tipificados	Calcula las diferencias entre valores observados y sus correspondientes valores pronosticados por el modelo.
Intervalos de pronóstico	Medias	Calcula el intervalo de confianza para la <u>media</u> de la variable dependiente en cada uno de los valores de la variable independiente.
	Individuos	Calcula el intervalo de confianza para el <u>valor individual</u> de la variable dependiente en cada uno de los valores de la variable independiente.

Tabla 12.1. Algunas opciones del botón Guardar... del procedimiento Regresión Lineal.

Por último, si pulsamos el botón **Opciones...** aparece una ventana con diferentes opciones. Las primeras nos permiten controlar criterios en regresión múltiple (no lo veremos aquí). Justo debajo, si el usuario marca la opción **Incluir constante en la ecuación**, el programa calcula un modelo con el término independiente, mientras que si no se marca esa opción, supondrá que se desea un modelo sin el término independiente, es decir, un modelo por el origen. Finalmente, existen tres opciones sobre el manejo de los valores perdidos: **Excluir casos según lista** (que excluye los casos que tienen valores perdidos en cualquiera de las variables utilizadas en cualquiera de los análisis), **Excluir casos según pareja** (que excluye del análisis los casos que tengan valores perdidos en alguna de las variables de la pareja implicada en dicho análisis) y **Reemplazar por la media** (que sustituye los valores perdidos por la media de la variable).

Estadísticos descriptivos

	Media	Desviación típ.	Ν
Valor tipificado de la densidad de masa ósea en el triángulo	-,8956	1,2111	94
Índice de masa corporal	23,984	3,854	94

		Valor tipificado de la densidad de masa ósea en el triángulo	Índice de masa corporal
Correlación de Pearson	Valor tipificado de la densidad de masa ósea en el triángulo	1,000	,354
	Índice de masa corporal	,354	1,000
Sig. (unilateral)	Valor tipificado de la densidad de masa ósea en el triángulo	,	,000
	Índice de masa corporal	,000	,
N	Valor tipificado de la densidad de masa ósea en el triángulo	94	94
	Índice de masa corporal	94	94

Correlaciones

Coeficientes^a

		Coeficie estanda	entes no arizados	Coeficientes estandarizados			Intervalo de para B	e confianza al 95%
Modelo		В	Error típ.	Beta	t	Sig.	Límite inferior	Límite superior
1	(Constante)	-3,562	,744		-4,785	,000	-5,040	-2,083
	imc	,111	,031	,354	3,627	,000	,050	,172

a. Variable dependiente: Valor tipificado de la densidad de masa ósea en el triángulo (imc)

Figura 12.5. Resultados del procedimiento Regresión Lineal.

Los resultados lógicamente dependerán de las opciones que hayamos marcado. Para el ejemplo que arrastramos y marcando las opciones básicas de **Estimaciones** e **Intervalos de confianza** para los coeficientes de regresión y **Descriptivos** (botón **Estadísticos...**) junto con la opción **Incluir constantes en la ecuación** (botón **Opciones...**) obtenemos unos resultados como los de la figura 12.5. En ella los primeros resultados que aparecen son la media y la desviación típica para las dos variables. Inmediatamente debajo aparece el coeficiente de correlación entre

las dos variables y la significación del test para dicho coeficiente. En la última tabla aparecen los coeficientes de la recta de regresión estimada, denominándose "constante" a la altura en el origen (-3.562 en nuestro caso) y apareciendo debajo la pendiente de la recta de regresión (0.111 en nuestro ejemplo). Para cada uno de estos coeficientes aparecen (a su derecha): el error estándar, la cantidad experimental para el contraste de que el coeficiente sea significativamente distinto cero, la probabilidad de significación del test y el correspondiente intervalo de confianza (la columna de coeficientes estandarizados no la comentaremos). Si se marcan otras opciones, como el cálculo de valores pronosticados o residuos no tipificados, éstos aparecerán en la matriz de datos, y en el visor de resultados aparecerá una tabla con medidas de resumen sobre dichos valores y otra con diagnósticos por caso (que nos permitirá detectar casos extremos).

12.3. El procedimiento Correlaciones bivariadas.

El procedimiento que nos realiza el cálculo del coeficiente de correlación entre dos variables es **Correlaciones bivariadas**, que se invoca mediante: **Analizar** \rightarrow **Correlaciones** \rightarrow **Bivariadas...**. Cuando se ejecuta esa secuencia se obtiene una ventana como la de la Figura 12.6. En ella, a la izquierda, hay una caja con las variables a seleccionar por el procedimiento conocido, y de las que algunas pasaremos a la caja **Variables:**, indicándole eso al programa que deseamos las correlaciones por parejas de todas las variables que estén en la caja. Los coeficientes de correlación que el programa ofrece son: el de **Pearson**, el de la **Tau-b de Kendall** y el de la **rho de Spearman**; los tres pueden ser seleccionados simultáneamente, aunque el de la Tau-b de Kendall no es muy usado. Bajo el epígrafe **Prueba de significación**, el usuario puede seleccionar la unilateralidad o la bilateralidad del contraste al que serán sometidos los coeficientes de correlación; lo común es que el usuario no tenga razones lo suficientemente fuertes como para que pida contrastes de una cola, por lo que SPSS ofrece, de entrada, el bilateral. El programa da también la posibilidad de marcar automáticamente los coeficientes de correlación que resulten significativos; de manera que si se selecciona **Marcar correlaciones significativas**, el programa marca con un asterisco los coeficientes de correlación

Correlaciones bivariadas	×
Variables: ✓ edad ↔ sexo ↔ peso ↔ talla ↔ imc ↔ tevol ↔ tabaco ↔ alcohol	Aceptar Pggar Bestablecer Cancelar Ayuda
Prueba de significación © <u>B</u> ilateral © <u>B</u> ilateral	
Marcar las correlaciones significativas	<u>O</u> pciones

significativos al nivel 0,05 y, con dos asteriscos, los significativos al nivel 0,01.

El botón de **Opciones...** nos da la posibilidad de pedir algunos estadísticos de resumen para las variables implicas en el análisis y elegir entre dos opciones sobre el manejo de valores perdidos: **Excluir casos según pareja** (excluye los casos con valores parecidos en alguna de las variables de la pareja que se analiza) y **Excluir casos según lista** (excluye los casos con valores perdidos en cualquiera de las variables seleccionadas en el procedimiento).

Cuando se hayan seleccionado las variables con las que se desea trabajar y las características particulares del análisis, obtendremos unos resultados donde lo fundamental aparece en forma de tabla cuadrada (como la de la figura 12.7) en la que, tanto por filas como por columnas, aparecen las variables seleccionadas; en nuestro caso: imc, szl24, sztri y szcue. Donde se cruzan una fila y una columna aparecen: el coeficiente de correlación entre esas dos variables, la significación del contraste de hipótesis para él (marcados con asteriscos los significativos), y el número de casos sobre los que se ha calculado ese coeficiente de correlación. Por último, hemos de llamar la atención sobre el hecho de que la tabla es simétrica con respecto a su diagonal principal, como puede comprobar el lector.

Correlaciones					
		IMC	SZL24	SZTRI	SZCUE
IMC	Correlación de Pearson	1.000	.101	.354(**)	.229(*)
	Sig. (bilateral)		.335	.000	.026
	Ν	94	94	94	94
SZL24	Correlación de Pearson	.101	1.000	.494(**)	.823(**)
	Sig. (bilateral)	.335	•	.000	.000
	Ν	94	94	94	94
SZTRI	Correlación de Pearson	.354(**)	.494(**)	1.000	.584(**)
	Sig. (bilateral)	.000	.000	•	.000
	Ν	94	94	94	94
SZCUE	Correlación de Pearson	.229(*)	.823(**)	.584(**)	1.000
	Sig. (bilateral)	.026	.000	.000	
	Ν	94	94	94	94
** La correlación es significativa al nivel 0,01 (bilateral).					
* La correlación es significante al nivel 0,05 (bilateral).					

Figura 12.7. Resultados del procedimiento Correlaciones Bivariadas