

# Práctica 2

## Estadística Descriptiva

---

### Contenido

Procedimiento básico para un análisis estadístico . . . . .	1
Tablas de frecuencias . . . . .	2
Medidas de centralización, dispersión y forma . . . . .	7
Representaciones gráficas . . . . .	8
Bibliografía . . . . .	10

---

### Procedimiento básico para un análisis estadístico

Una vez que se ha aprendido a introducir y manipular datos en el editor, comienza un proceso de extracción de información que se inicia con el análisis descriptivo de la muestra (objetivo de los ejercicios que siguen) y, mediante distintas técnicas estadísticas, culminará con conclusiones sobre la población (inferencia estadística, que se llevará a cabo en prácticas posteriores).

Los pasos que se siguen en cualquier tipo de análisis con SPSS son los siguientes:

1. **Elegir un archivo de datos.** Los datos para el análisis se introducirán manualmente o bien, se abrirá un archivo previamente guardado según se ha descrito en las páginas anteriores.
2. **Elegir el procedimiento estadístico**, para transformar el archivo de datos, calcular estadísticos u obtener un gráfico.
3. **Seleccionar las variables y opciones para el análisis.** Las variables disponibles se muestran en el cuadro de diálogo del procedimiento elegido.
4. **Examinar los resultados** que se muestran en el Visor de resultados.

SPSS recoge distintos procedimientos que hacen posible realizar análisis de tipo descriptivo de un conjunto de datos. Estos procedimientos permiten, en líneas generales:

- La **tabulación** de distribuciones de variables de tipo continuo y discreto, así como distribuciones de frecuencias para atributos.

- El cálculo de **medidas de centralización, dispersión y forma**, que resuman la información contenida en la distribución de frecuencias.
- La elaboración de **representaciones gráficas**, entre las que se puede destacar los diagramas de barras, los histogramas de frecuencia o los diagramas de sectores entre otros.

En esta Práctica, vamos a llevar a cabo a través de algunos ejercicios, estos tres procedimientos.

## Tablas de frecuencias

Tras superar la etapa de recogida de datos del análisis estadístico, se debe proceder a la ordenación y presentación de la información recogida. Una buena forma de presentar la información es mediante la elaboración de tablas.

Una tabla de frecuencias recoge los distintos valores de la variable o modalidades del atributo, junto con sus frecuencias absolutas y relativas. Distinguiremos si estamos trabajando con datos de tipo continuo (agrupados) o con datos de tipo discreto (sin agrupar). Para los ejercicios que siguen usaremos el fichero de datos **Mundo 95.sav** que se encuentra en el directorio raíz de SPSS.

### Tablas de frecuencias con datos sin agrupar

Un procedimiento posible (más adelante veremos otro) para la elaboración de una tabla de frecuencias consiste en pulsar en la barra de menú



En el siguiente ejercicio vamos a seguir este procedimiento para crear una tabla de frecuencias.

---

#### Ejercicio 1

---

En el Editor de Datos, edita el fichero **Mundo 95.sav**, incluido en el directorio raíz de SPSS y elabora una tabla de frecuencias. Usa para ello, las variables

*poblac* (Población  $\times$  1000);

*pib\_cap* (Producto interior bruto per-cápita);

*nac\_def* (Tasa de Nacimientos/Defunciones).

Practica también con otras variables de tu elección.

---

**SOLUCIÓN:** Como consecuencia de haber pulsado en los comandos anteriores, aparece el cuadro de diálogo **Frecuencias**, de donde hay que seleccionar, del campo de la izquierda, la variable o variables de las cuales se desee elaborar la tabla de frecuencias y desplazarla al recuadro de **Variables**, haciendo un sólo *click* en la flecha  situada entre estos dos

recuadros. Haz esta operación con las tres variables del enunciado del ejercicio. Seguidamente, activa la opción **Mostrar tablas de frecuencias** mediante un sólo *clic* en el cuadrado que se encuentra situado a su izquierda. De los tres botones de la parte inferior, sólo usaremos **Formato...**, que abre el cuadro de diálogo del mismo nombre en el que elegiremos una de las cuatro opciones (la que quieras) del grupo **Ordenar por**. En el grupo **Múltiples variables** marca cualquiera de las opciones, ya que al no haber pulsado el botón **Estadísticos**, el efecto es prácticamente el mismo. Para regresar al cuadro anterior, pulsa el botón **Continuar**. Finalmente, pulsando el botón **Aceptar**, se abre la pantalla del *Visor de Resultados de SPSS* que mostrará las tablas de frecuencias que se han elaborado.

Observa que como consecuencia del procedimiento llevado a cabo en el último ejercicio, se abre la ventana Visor de Resultados mostrando una pequeña tabla con el título **Estadísticos** conteniendo el número de datos válidos y perdidos para cada una de las variables que hemos elegido. A continuación aparecen una tabla por cada variable conteniendo

- Valores o categorías de la variable.
- Frecuencia: frecuencia absoluta de cada valor.
- Porcentaje: frecuencia relativa, incluyendo los valores perdidos.
- Porcentaje válido: frecuencia relativa, eliminando de la muestra los valores perdidos.
- Porcentaje acumulado: frecuencia relativa acumulada, eliminando de la muestra los valores perdidos.

Para cambiar de la pantalla Visor de resultados a la pantalla Visor de datos y viceversa, pulsa el comando **Ventana** de la barra de menú y elige la opción adecuada en el menú desplegable. También puedes usar la barra de estado de Windows al pie de la pantalla.

Hay otro procedimiento para elaborar tablas de frecuencias que permite mostrar en una sola tabla varias variables. Es conveniente, para que la tabla no resulte confusa, que las distintas variables se refieran a la misma magnitud. Por ejemplo, en el fichero **Mundo 95.sav**, las variables *espvidaf* y *espvidam* son las esperanzas de vida femenina y masculina respectivamente, de modo que tiene sentido presentarlas en la misma tabla. Naturalmente, también puede usarse para mostrar una sola variable. El procedimiento se inicia pulsando en la barra de menú



---

### Ejercicio 2

Elabora una tabla de frecuencias con los datos del fichero **Mundo 95.sav**. Usa para ello las variables

*espvidaf* (Esperanzas de vida femenina);

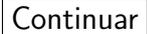
*espvidam* (Esperanzas de vida masculina).

Ensayá las diferentes características: estadísticos, diseño, formato y títulos, probando en cada una de ellas las distintas opciones que aparecen en los respectivos cuadros de diálogo.

---

**SOLUCIÓN:** Al pulsar en los comandos anteriores, aparece el cuadro de diálogo **Tablas de frecuencias**, de donde hay que seleccionar, del campo de la izquierda, la variable o variables de las cuales se desee elaborar la tabla de frecuencias y desplazarla al recuadro de **Frecuencias para**, haciendo un sólo *clic* en el botón . Seguidamente, se procederá a especificar los estadísticos, el diseño, formato y títulos de la tabla.

- **Estadísticos:** pulsando el botón  del cuadro de diálogo **Tablas de frecuencias**, aparece un nuevo cuadro de diálogo llamado **Tablas de frecuencias: Estadísticos**. En este cuadro selecciona:
  - **Recuento:** Determina la frecuencia absoluta de la variable o atributo, pudiéndose especificar la etiqueta de este estadístico y su formato de visualización.
  - **Porcentajes:** Determina la frecuencia relativa de la variable o atributo en porcentaje, pudiéndose especificar la etiqueta de este estadístico y su formato de visualización.
  - **Totales:** Muestra los totales de cada uno de los estadísticos anteriores, pudiéndose especificar la etiqueta de este estadístico.

A continuación pulsa el botón .

- **Diseño:** Para especificar el diseño de la tabla, pulsamos el botón  del cuadro de diálogo **Tablas de frecuencias**. Aparece entonces un nuevo cuadro de diálogo llamado **Tablas de frecuencias: Diseño**, donde podremos establecer:
  - La posición de las etiquetas de variable. Bien en la parte superior de la tabla o bien en el lateral.
  - La posición de las etiquetas de estadísticos que se utilizan como encabezados. Bien en la parte superior (cada estadístico aparece en una columna distinta) o bien en el lateral (cada estadístico aparece en una fila distinta) o en distintas capas (cada estadístico aparece en una capa distinta).

A continuación pulsa el botón .

- **Formato:** Para especificar el formato de la tabla, pulsamos el botón  del cuadro de diálogo **Tablas de frecuencias**. Aparece entonces un nuevo cuadro de diálogo llamado **Tablas de frecuencias: Formato**, donde podremos establecer:
  - El aspecto de las casillas vacías. Éstas podrán aparecer vacías o con el valor cero.
  - En el recuadro de **los estadísticos perdidos aparecen como**, se puede indicar un carácter, palabra o frase que se mostrará en los resultados para los estadísticos perdidos.

A continuación pulsamos el botón .

- **Títulos:** Para especificar el título, el pie y el texto de la esquina de la tabla, pulsamos el botón  del cuadro de diálogo **Tablas de frecuencias**. A continuación aparecerá el cuadro de diálogo **Tablas de frecuencias: Títulos**, donde podremos establecer:
  - **Título.** Permite definir el título que aparecerá en el encabezado de los resultados.
  - **Pie:** Permite definir el título que aparecerá en el pie de los resultados.
  - **Esquina:** Este campo se usa para escribir un texto que aparecerá en la esquina superior izquierda de la tabla. Puede especificar hasta 10 líneas.

A continuación pulsamos el botón .

Finalmente para elaborar la tabla de frecuencias pulsamos el botón . Observa el resultado de todas las acciones llevadas a cabo en la pantalla del Visor de resultados.

## Tablas de frecuencias con datos agrupados

Como se indica en los apuntes del Tema 1 de la asignatura, dedicado a la Estadística Descriptiva, la agrupación de los datos en *clases* es un procedimiento para reducir la variabilidad de los datos cuando éstos toman muchos valores (numéricos) distintos. SPSS dispone de un procedimiento para la construcción de intervalos de clase, que se activa pulsando en la barra de menú

→ .

El nombre del comando se debe a que SPSS llama *categorizar* al proceso de agrupar los datos en clases. En el siguiente ejercicio se lleva a cabo tal agrupación.

### Ejercicio 3

- En el fichero **Datos de empleados.sav**, agrupa manualmente en cuatro clases, los valores de la variable de nombre **tiempemp** (que tiene por etiqueta **Meses desde el contrato**), creando para ello una nueva variable de nombre **tiempclas** y etiqueta **Meses desde el contrato categorizada**.
- En el mismo fichero agrupa automáticamente en nueve clases los valores de la variable **salario** (cuya etiqueta es **Salario actual**). Para ello crea una nueva variable **salclas** y etiqueta **Salario actual categorizado**.
- Por último elabora, por alguno de los procedimientos explicados en la sección anterior, tablas de frecuencias para las nuevas variables.

**SOLUCIÓN:** a) Después de pulsar los anteriores comandos de la barra de menú, se abre el cuadro de diálogo **Categorizador visual** en el que elegimos las variables cuyos valores deseamos agrupar en clases, de la lista que aparece en el campo de la izquierda, pasándolas al campo de derecha con la flecha ►, en nuestro caso **tiempemp** y **salario**. En este mismo

cuadro podemos elegir, si marcamos la opción **Limitar el número de casos explorados a:**, y escribimos en el campo correspondiente un número, cuántos de los casos (empezando por el primero) se van a categorizar (no uses esta opción por ahora). Con el botón **Continuar**, pasamos al siguiente cuadro de diálogo que tiene el mismo nombre **Categorizador visual**.

En la ventana de la izquierda marcamos la variable **Meses desde el contrato**, cuyos valores vamos a agrupar, apareciendo en los campos de la derecha (arriba) el nombre y la etiqueta de esta variable, un campo en blanco para escribir el nombre de la nueva variable que resulta al categorizar ésta (aquí debes escribir **tiempclas**), y la etiqueta, que es la misma que la de la original, pero con la palabra (**Categorizada**) entre paréntesis. Modifica esta última etiqueta escribiendo en su lugar **Meses desde el contrato categorizada**. Todos estos nombres y etiquetas se pueden cambiar más tarde, si se desea, en la *Vista de variables*.

Bajo estos campos figuran los valores mínimo y máximo de los datos y un pequeño histograma que nos da una idea previa de la distribución de frecuencia de los datos.

En la *rejilla* de más abajo introduciremos manualmente los valores de los extremos superiores de las clases, y si lo deseamos, unas etiquetas para nombrar dichas clases. Como el dato más pequeño es 63, y el más grande 98, tomamos el intervalo  $[60, 100]$  y lo dividimos en cuatro clases, para lo que escribimos en **Valor** los números 60, 70, 80, 90 y 100.

En el grupo de la derecha **Límites superiores** marcamos la primera opción **Incluidos ( $\leq$ )**, pulsamos el botón **Crear etiquetas** (observa lo que ocurre, y recuerda que has pulsado la opción **Incluidos ( $\leq$ )**), y por último no cierras todavía el cuadro de diálogo con el botón **Aceptar** porque vamos a proceder con la otra variable.

b) Con la variable **Salario actual** llevamos a cabo las mismas actuaciones que acabamos de realizar en el apartado a), con la diferencia que ahora, como los valores mínimo y máximo de los datos son 15750 y 135000, consideramos el intervalo  $[15750, 135000]$ . Vamos a clasificar los datos en nueve clases, de modo que la amplitud de cada una es  $\frac{135000 - 15750}{9} = 13250$ . En el grupo **Límites superiores** marcamos ahora la segunda

opción **Excluidos ( $<$ )**, y a continuación pulsamos el botón **Crear puntos de corte...** para proceder a la categorización automática de los datos. En el cuadro de diálogo **Crear puntos de corte** usa únicamente la primera opción **Intervalos de igual amplitud**, y en el grupo **Intervalos:** rellene al menos dos campos escribe en el campo **Posición del primer punto de corte** 15750, en el campo **Número de puntos de corte** escribe 10 (ten en cuenta que nueve intervalos están delimitados por diez puntos de separación). En el campo **Amplitud** aparecerá un número, que debes sobrescribir con el valor 13250 que hemos calculado. A la derecha de **Posición del último punto de corte:** aparecerá el número 135000, que es precisamente el mayor de los datos. Pulsando el botón **Aplicar** regresamos al cuadro de diálogo anterior. Ahora creamos las etiquetas pulsando de nuevo el botón **Crear etiquetas**. Observa el resultado de esta acción en la columna **Etiquetas** de *rejilla*.

Finalmente, cerramos el cuadro de diálogo con el botón **Aceptar**. En la *Vista de datos* habrá dos nuevas variables con los valores de los límites superiores de las clases que acabamos de crear.

c) Una vez creadas las variables categorizadas, éstas pueden tabularse. Hazlo usando el procedimiento que prefieras de los descritos en la sección anterior.

## Medidas de centralización, dispersión y forma

Para el cálculo de estas medidas SPSS dispone de cuatro procedimientos diferentes. En esta práctica, usaremos dos: el procedimiento **Frecuencias**, que proporciona la media, mediana, moda, suma, desviación típica, varianza, amplitud, valores mínimo y máximo, error típico de la media, asimetría y curtosis (ambos con sus errores típicos), cuartiles y percentiles, y el procedimiento **Descriptivos**, que proporciona la media, suma, desviación típica, varianza, amplitud, valores mínimo y máximo, error típico de la media, asimetría y curtosis.

Para ejecutar el procedimiento Frecuencias, pulsamos en la barra de menú



En el siguiente ejercicio se detalla el uso de este procedimiento.

### Ejercicio 4

En el fichero de datos **Mundo 95.sav**, usa la variable *densidad* (*Habitantes por Km<sup>2</sup>*) para calcular estadísticos descriptivos usando el procedimiento **Frecuencias**.

**Solución:** Ya hemos usado este procedimiento en un ejercicio anterior para construir tablas. Ahora lo vamos a emplear para calcular estadísticos descriptivos.

Como tenemos que volver a usar el fichero **Mundo 95.sav**, para cambiar a ese fichero (si no lo has cerrado) usa el comando **Ventana** de la barra de menú. Una vez situado en este fichero y después de pulsar los comandos anteriores en la barra de menú, seleccionamos la variable *densidad* (*Habitantes por Km<sup>2</sup>*) en el campo de la izquierda del cuadro de diálogo que aparece, tal como se ha descrito en secciones anteriores. Aunque podemos obtener una tabla de la variable, como eso ya lo hemos hecho en ejercicios anteriores, ahora lo omitiremos por brevedad, para lo que debemos **desmarcar** la opción **Mostrar tablas de frecuencias**. Seguidamente, pulsaremos el botón **Estadísticos** y marcaremos en el nuevo cuadro de diálogo los estadísticos que deseamos obtener: En el grupo **Valores percentiles**, marca **Cuartiles** y **Percentiles**; y en el campo correspondiente introduce los valores 15, 35, 64 y 87 (por ejemplo) pulsando el botón **Añadir** entre cada valor y el siguiente.

En el grupo **Dispersión** marca **Varianza**, **Mínimo** y **Máximo**. En el grupo **Tendencia central** marca **Media**, **Mediana** y **Moda**. El resto de las opciones del cuadro de diálogo no las usaremos.

Pulsamos **Continuar** para regresar al cuadro de diálogo anterior, en el que pulsaremos **Aceptar** para que SPSS lleve a cabo el análisis cuyos resultados aparecerán en el Visor de resultados en un cuadro con el título *Estadísticos*. Obsérvalo para que reconozcas cada uno de los estadísticos descriptivos que has seleccionado.

Como queda dicho más arriba, otro método para calcular medidas de centralización, dispersión y forma es mediante el procedimiento **Descriptivos**. Para llevarlo a cabo, elegimos los menús desplegables



Los detalles se describen en el siguiente ejercicio.

---

**Ejercicio 5**

---

En el fichero de datos **Mundo 95.sav**, usa las variables

- *pib\_cap* (*Producto interior bruto per-capita*)
- *inc\_pob* (*Aumento de la población (% anual)*)
- *poblac* (*Población en miles*)
- *urbana* (*Habitantes en ciudades (%)*)

para calcular diversos estadísticos descriptivos usando el procedimiento **Descriptivos** según se acaba de explicar.

---

**SOLUCIÓN:** En el cuadro de diálogo **Descriptivos**, seleccionamos en la lista de variables de la izquierda aquellas para las que deseamos efectuar el análisis, y pulsando el botón , las trasladamos a la lista de la derecha. Ahora pulsamos el botón **Opciones** y en el cuadro de diálogo que se abre, marca las opciones **Media**, **Desv. típica**, **Amplitud**, **Máximo** y **Mínimo**. El grupo **Orden de visualización** indica el orden en que aparecerán los resultados del análisis en el Visor de resultados. Elige el que quieras.

Pulsando **Continuar** se cerrará el cuadro, regresando al anterior, en el que pulsaremos **Aceptar** para que los resultados aparezcan en el Visor de resultados en un cuadro cuyo título es *Estadísticos descriptivos*. Observa dicho cuadro para verificar que las medidas que contiene son las que habías elegido.

## Representaciones gráficas

Del mismo modo que en la sección anterior, SPSS dispone de varios procedimientos para crear representaciones gráficas. El más simple de ellos consiste también en el procedimiento **Frecuencias**, que proporciona gráficos de barras, gráficos de sectores e histogramas.

Para ejecutar el procedimiento, pulsamos en la barra de menú



---

**Ejercicio 6**

---

Usando una vez más el fichero de datos **Mundo 95.sav**, realiza un histograma con el procedimiento **Frecuencias**. Para ello, usa la variable etiquetada como *Aumento de la población (% anual)*.

---

**SOLUCIÓN:** En el cuadro de diálogo **Frecuencias** seleccionamos la variable de etiqueta *Aumento de la población (% anual)*. A continuación, desmarcamos la opción **Mostrar tablas de frecuencias** para que los resultados sean más sencillos, pulsamos el botón **Gráficos**,

y podremos establecer, en el nuevo cuadro de diálogo, las representaciones gráficas deseadas. En este caso marcamos la opción **Histogramas:**. Pulsando  , y  , aparecerá el histograma en el Visor de resultados.

Pero el procedimiento general para realizar gráficos en SPSS, se encuentra en el comando **Gráficos** de la barra de menú, que despliega diversas modalidades de representaciones gráficas. Sólo vamos a considerar por brevedad los gráficos de *líneas*, *sectores* y *Pareto*.<sup>1</sup>

---

### Ejercicio 7

---

Traza un diagrama de líneas para representar una frente a la otra, las variables de etiquetas *Esperanza de vida masculina* y *Esperanza de vida femenina*, del fichero de datos **Mundo 95.sav**.

**SOLUCIÓN:** Para que el gráfico resulte fácil de interpretar, vamos a ordenar previamente los casos con respecto a una de las dos variables. Para ello, en la barra de menú pulsa

→

En el cuadro de diálogo **Ordenar casos** traslada la variable *Esperanza de vida masculina* al campo **Ordenar por**, elige el orden de clasificación **Ascendente** y pulsa el botón  .

Con ello, todos los casos se han reordenado en orden creciente con respecto a los valores de la variable elegida.

A continuación, en la barra de menú pulsa

→

y en el cuadro de diálogo **Gráficos de líneas**, elige la modalidad de gráfico **Simple**, la opción **Valores individuales de los casos** y pulsa  . Se abre con ello otro cuadro de diálogo, en el que debes seleccionar de la lista de la izquierda, la variable *Esperanza de vida masculina*, y mediante el botón  pasarla al campo **La línea representa:** y ahora en el campo **Etiquetas de las categorías** marca la opción **Variable**, lo que te permitirá seleccionar en la lista de variables de la izquierda *Esperanza de vida femenina* y pasarla al campo de la derecha. De este modo, hemos situado en el eje de abcisas los datos de esta última variable, y en el de ordenadas los de *Esperanza de vida masculina*. Dada la índole de estas variables, es indiferente en que eje situemos cada una de ellas, pero en otros casos puede ser importante la elección de uno u otro para determinadas variables. Pulsa el botón  , observa el diagrama aparecerá en el Visor de resultados e intenta interpretar la forma de la curva que resulta.

---

### Ejercicio 8

---

Usando la variable cuya etiqueta es *Región económica* del fichero de datos **Mundo 95.sav**, construye un diagrama de sectores del número de países que pertenecen a cada una de las regiones económicas del mundo.

**SOLUCIÓN:** En la barra de menú pulsa

---

<sup>1</sup>Wilfredo Pareto (1848-1923), economista y sociólogo.

Gráficos → Sectores...

En el cuadro de diálogo **Gráficos de sectores**, elige la opción **Resúmenes para grupos de casos** y pulsa **Definir**. Se abre otro cuadro de diálogo, en el que debes seleccionar de la lista de la izquierda, la variable *Región económica* y mediante el botón  pasarla al campo **Definir sectores por:**. En el grupo **Los sectores representan** marca la opción **% de casos**. Pulsa el botón **Aceptar**, observa el diagrama de sectores que aparecerá en el Visor de resultados e interprétalo.

---

### Ejercicio 9

Usando la variable cuya etiqueta es *Religión mayoritaria* del fichero de datos **Mundo 95.sav**, construye un diagrama de Pareto para esta variable cualitativa.

---

SOLUCIÓN: En la barra de menú pulsa

Gráficos → Pareto...

En el cuadro de diálogo **Gráficos de Pareto**, elige la modalidad de gráfico **Simple**, la opción **Recuento o suma para grupos de casos** y pulsa **Definir**. Se abre otro cuadro de diálogo, en el que debes seleccionar de la lista de la izquierda, la variable *Religión mayoritaria* y mediante el botón  pasarla al campo **Eje de categorías:**. En el grupo **Las barras representan** marca la opción **Recuentos**. Marca la opción **Mostrar línea acumulativa** y por fin pulsa el botón **Aceptar**, observa el diagrama de Pareto que aparecerá en el Visor de resultados con sus barras y polígono de frecuencias acumuladas.

## Bibliografía

*Manual de SPSS de la Universidad de Cádiz.*

<http://www2.uca.es/serv/ai/formacion/spss/Inicio.pdf>

*Cuaderno de prácticas de SPSS de la asignatura Análisis de datos en Psicología I. Universidad Autónoma de Madrid.*

[http://www.uam.es/personal\\_pdi/psicologia/carmenx/MaterialID.html](http://www.uam.es/personal_pdi/psicologia/carmenx/MaterialID.html)

Manzano, V. et al. *SPSS para Windows*. Madrid, Ra-Ma,.

Pérez, César, *Técnicas Estadísticas con SPSS*. Prentice Hall.

Portilla, M. et al. *Manual práctico del paquete estadístico SPSS 9 para Windows*. Universidad Pública de Navarra.

*Tema 1. Estadística Descriptiva*. Publicación interna del Departamento de Matemática Aplicada II de la Escuela Universitaria Politécnica de Sevilla.