

**Universidad Católica del Uruguay**  
**Facultad de Ciencias Humanas**

## **Autotexto**

**Serie Estadística**

# **Análisis univariado. Distribución de frecuencias, medidas descriptivas, construcción de gráficos**



**Universidad  
Católica**

DAMASO A. LARRAÑAGA • URUGUAY

## **Laboratorio Metodológico**

Versión original:

P. Alegre – J. Bogliaccini – F. Rodríguez

Autores revisión 2008:

A. de León – M. Dodel – C. Rafaniello

## **Módulo de Práctica de Análisis**

### **Serie Estadística**

**Tema: Análisis univariado. Distribución de frecuencias, medidas descriptivas, construcción de gráficos.**

### **Descripción**

Buena parte de la tarea de la Estadística descriptiva consiste en la organización y resumen de datos. La distribución de frecuencias permite organizar y mostrar un conjunto de observaciones por medio del agrupamiento de datos, indicando el número de veces que ocurre cada valor.

La distribución puede mostrar los valores absolutos de las observaciones (la cantidad de veces que ocurre la observación) así como los valores relativos de las observaciones. (los porcentajes que presenta cada categoría). Ésta se obtiene de la división de cada frecuencia absoluta entre el número total de frecuencias, por lo que la suma de todos los valores relativos equivale al 100% de las observaciones.

Las frecuencias acumuladas surgen de la suma de cada categoría desde la categoría con menos observaciones hacia la categoría con más observaciones, pudiéndose expresarse tanto en términos absolutos como porcentuales.

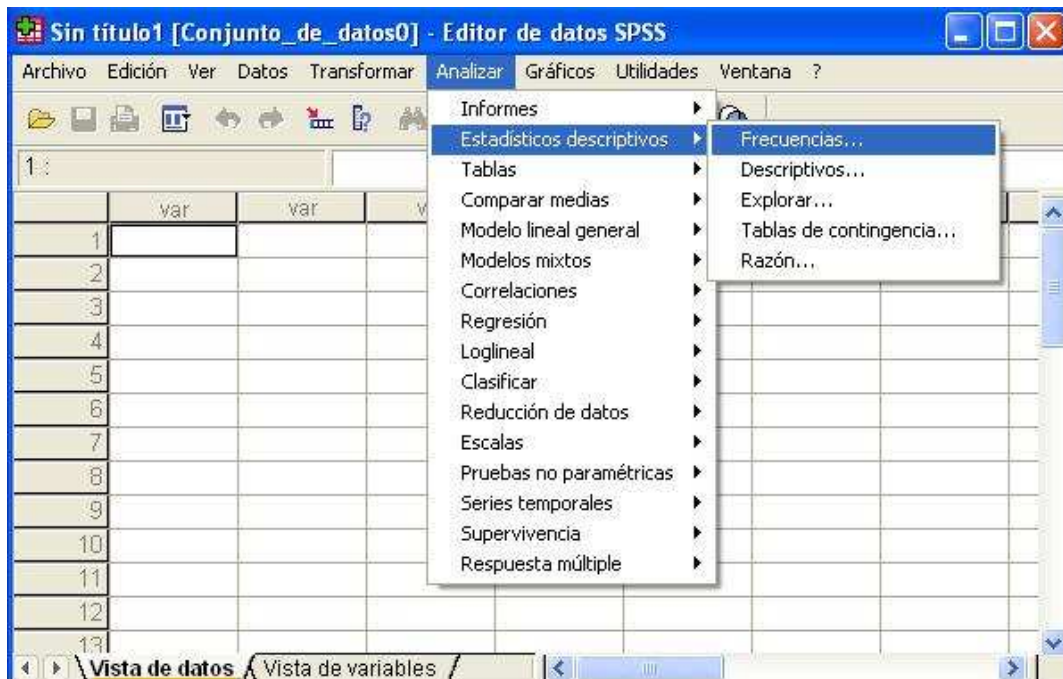
Las distribuciones de frecuencias también pueden visualizarse en forma gráfica. Estas formas de representación de datos permiten visualizar de otro modo la organización de los mismos. Entre algunos de los gráficos más utilizados están los histogramas, las gráficas circulares o de "torta", las gráficas de barras y los polígonos de frecuencia.

### **Procedimiento**

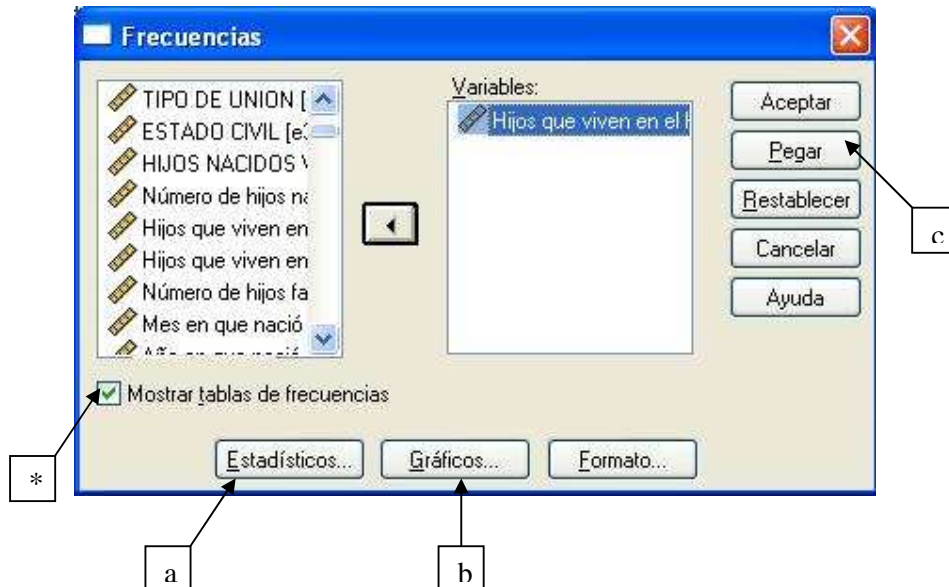
#### **¿Cómo obtener la distribución de frecuencias de una variable?**

En esta guía presentaremos el procedimiento por el cual se describe una variable nominal. Para describir una variable utilizamos el comando *Frecuencias*, en el submenú *Estadísticos descriptivos* del menú *Analizar*.

Este menú, *Analizar*, contiene la totalidad de los procedimientos estadísticos incluidos en el SPSS, ordenados de los más frecuentes a los menos frecuentes. Los procedimientos de estadística univariada y los de estadística bivariada para variables nominales u ordinales están incluidos en el primer submenú *Estadísticos descriptivos*.



Para obtener la distribución de frecuencias de una determinada variable basta con seleccionar el procedimiento **Frecuencias**, seleccionar la variable (o las variables) del menú que aparece a la izquierda, y oprimir el botón **Aceptar/OK**.



Pero en esta oportunidad queremos además:

- a. Calcular medidas descriptivas univariadas (*Estadísticas*)
- b. Obtener una gráfica univariada (*Gráficos*)
- c. Dejar un registro escrito de todos los procedimientos ejecutados (*Pegar* en lugar de *OK*)
- d. Interpretar el Output (archivo de salida) de lo trabajado

## ¿Cómo obtener medidas descriptivas para una variable?

En la caja de diálogo *Estadísticos* del comando *Frecuencias* disponemos de una serie de medidas descriptivas de distinto tipo:

- valores de percentiles (cuartiles; puntos de corte para 10 grupos iguales, o deciles; percentiles)
- medidas de tendencia central (media; mediana; moda; suma)
- medidas de dispersión (desvío estándar; varianza; rango; mínimo; máximo; error estándar de la media)
- medidas de distribución (asimetría; curtosis)



**Tabla 1:** Medidas descriptivas que pueden aplicarse a los distintos niveles de medición

	Nominal	Ordinal	Interval
Valores de percentiles		Todos	Todos
Medidas de tendencia central	Moda	<ul style="list-style-type: none"> <li>• Moda</li> <li>• Mediana</li> </ul>	<ul style="list-style-type: none"> <li>• Moda</li> <li>• Mediana</li> <li>• Media</li> </ul>
Medidas de dispersión			<ul style="list-style-type: none"> <li>• Mínimo</li> <li>• Máximo</li> <li>• Desvío estándar</li> <li>• Varianza</li> <li>• Rango</li> <li>• Error estándar de la media</li> </ul>
Medidas de distribución			<ul style="list-style-type: none"> <li>• Asimetría</li> <li>• Curtosis</li> </ul>

## ¿Cómo obtener gráficas en un análisis univariado?

En la caja de diálogo *Gráficos* del comando *Frecuencias* disponemos de distintos tipos de gráficas:

- barras
- pastel, torta o sectores
- histograma (con opción de curva normal)



**Tabla 2:** Gráficas univariadas que pueden aplicarse a los distintos niveles de medición

	Nominal	Ordinal	Interval
Gráficas	<ul style="list-style-type: none"><li>• Barras</li><li>• Torta/Pastel</li></ul>	<ul style="list-style-type: none"><li>• Barras</li><li>• Torta/Pastel</li></ul>	<ul style="list-style-type: none"><li>• Barras</li><li>• Torta/sectores</li><li>• Histograma</li></ul>

## Sintaxis

Ejemplo:           FREQUENCIES  
                      VARIABLES=e4  
                      /STATISTICS=MODE  
                      /PIECHART PERCENT  
                      /ORDER= ANALYSIS .

Estructura de la sintaxis:

“*Frecuencias*”: Es el nombre que indica el comando, frecuencias en Ingles. Puede abreviarse como *freq*

“*Variables=*”: Como el propio nombre lo explicita, indica sobre qué variables se realizará el comando.

Se puede realizar el mismo procedimiento sobre más de una variable a la vez. El SPSS tomará en cuenta todas las variables que se encuentren en esta línea de comandos debiendo separarlas simplemente por un espacio. Esta sentencia del comando también se puede abreviar como “*var=*”

Advertencia: Se debe tener cuidado con el nivel de medición de las distintas variables para no aplicar estadísticos o gráficas inadecuadas para el nivel de la misma.

“*/Statistics=*” Estadísticos en Ingles. Bajo este sub-comando se le indica al programa qué tipo de estadísticos aplicar en el análisis de la/s variable/s. Tal como los comandos, los estadísticos se digitan según sus nombres en ingles, a modo de ejemplo, *mode* es moda y *mean* media<sup>1</sup>. Al igual que en el caso de las variables, se puede solicitar más de estadístico a la vez con sólo separarlos por un espacio.

Este sub-comando también se puede abreviar como “*/stat=*” y debe tenerse el mismo tipo de cuidado que con las gráficas en relación a los niveles de medición de las variables.

#### **Otros sub-comandos:**

Gráficas: Cada gráfico univariado tiene un comando particular, aunque como todo sub-comando siempre comienza con la barra diagonal: /

- De barras: “*/barchart*”; De torta/sectores: “*/piechart*”. En ambos casos se hace necesario indicar también si se quiere el gráfico con las frecuencias o los porcentajes. Simplemente se le debe de agregar “*freq*” o “*percent*” al final de la línea.
- Histograma: “*/histogram*” Si se desea que en el gráfico también aparezca la curva normal simplemente se debe agregar “*normal*” en la línea.

Percentiles:<sup>2</sup> Este sub-comando se digita “*/ntiles=*” y luego de la igualdad la cantidad de puntos de corte que se desean. A modo de ejemplo, si se desean obtener cuartiles, deberíamos de digitar: “*/ntiles= 4*”. De manera similar que en los estadísticos y las frecuencias, se pueden ordenar cuarteles, quintiles, etc. a la vez, separando los puntos de corte con un espacio.

“*/order= análisis.*” Con esto se finaliza el comando de frecuencias.

Advertencia: Como en todo comando, es fundamental no olvidarse de digitar el punto **al final** del comando para indicarle al programa que terminamos de escribir.

---

<sup>1</sup> Por más información ver autotextos de tendencia central y dispersión

<sup>2</sup> Por mayor información dirigirse al autotexto de percentiles

## ¿Cómo interpretar el archivo de salida de frecuencias?

Los resultados propiamente dichos, en este caso son: dos tablas y una gráfica.

La primera tabla muestra el total de casos válidos y perdidos para la variable, y las medidas descriptivas solicitadas (en este caso, la moda, valor más frecuente de una variable, que resulta ser Soltero).

### Estadísticos

Estado Conyugal		
N	Válidos	174
	Perdidos	0
Moda		6

La segunda tabla contiene la distribución de frecuencias.

Siempre que la solicitemos, obtendremos una tabla con un formato de 5 columnas ([ver \\*](#)).

La primera contiene las categorías/valores de la variable. La segunda la frecuencia o cantidad de casos para cada una de las categorías. Mientras que la tercera detalla el porcentaje o la frecuencia relativa.

La cuarta columna, el porcentaje válido, presenta las frecuencias relativas pero solamente para los valores seleccionados como válidos o no perdidos<sup>3</sup>.

La última columna presenta el porcentaje acumulado, y es útil a la hora de analizar variables no nominales (por ejemplo podemos obtener la mediana y percentiles a través de la misma<sup>4</sup>).

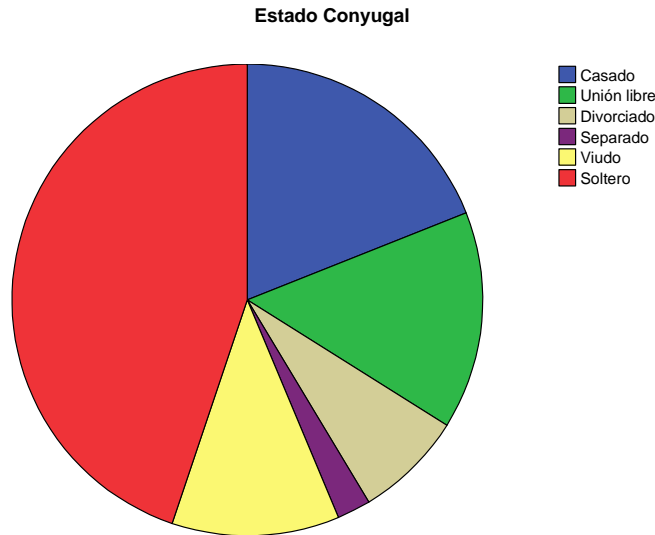
### Estado Conyugal

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Casado	33	19,0	19,0	19,0
	Unión libre	26	14,9	14,9	33,9
	Divorciado	13	7,5	7,5	41,4
	Separado	4	2,3	2,3	43,7
	Viudo	20	11,5	11,5	55,2
	Soltero	78	44,8	44,8	100,0
	Total	174	100,0	100,0	

Finalmente, la Gráfica 1 es un equivalente visual de la distribución de frecuencias de la segunda tabla. Esta gráfica se denomina gráfico de tortas/sectores o “pie chart” en inglés, y resulta muy útil cuando se representan las frecuencias relativas.

<sup>3</sup> Ver autotexto general

<sup>4</sup> Ver autotexto percentiles



**Ejercicios con Encuesta Continua de Hogares 2005, Personas.**

A)

- Abrir la base.
- Generar la distribución de frecuencias de la variable estado civil.
- Calcular las medidas descriptivas aplicables a su nivel de medición.
- Graficar la distribución
- ¿Qué características tiene la distribución?. ¿A que atribuye los valores porcentuales existentes?.

B)

- Realice una distribución de la frecuencia de la variable edad
- A partir de las distintas medidas descriptivas disponibles (tanto de tendencia central, como de dispersión o distribución) caracterice el comportamiento de la distribución en Montevideo e Interior. Puede apoyarse en gráficos para visualizar los patrones de similitud o diferencia.

---

**Bibliografía de referencia**

- Blalock, H. (1966) Estadística Social. FCE. México.
  - Peña, D. Romo, J - Introducción a la Estadística para las Ciencias Sociales. Mc Graw Hill 1997.
  - Mason y Lind - Estadística para administración y economía. Alfaomega 1998. México, D.F. 8ª edición.
-