

Universidad Católica del Uruguay
Facultad de Ciencias Humanas

Autotexto

Serie Estadística

Cálculo de Percentiles



Universidad
Católica

DAMASO A. LARRAÑAGA • URUGUAY

Laboratorio Metodológico

Versión original:

P. Alegre – J. Bogliaccini – M. Cardoso – F. Rodríguez

Autores revisión 2008:

A. de León – M. Dodel – C. Rafaniello

Módulo de Práctica de Análisis

Serie Estadística

Tema: Cálculo de cuartiles, quintiles y deciles

Descripción

En esta guía trabajaremos sobre cómo realizar el cálculo de diversos tipos de percentiles (cuartiles, quintiles, deciles) en base a la distribución de una variable continua. Es una técnica que permitirá “ordenar” las unidades de análisis (casos) de acuerdo al valor que adquieran en la distribución de mi variable de interés en la cantidad de grupos que yo le indique, de igual tamaño. Así, si quisiéramos ordenar a los individuos en 10 grupos (deciles) de acuerdo a su estatura, podríamos identificar al 10% más bajo de nuestra población a los efectos de realizar un estudio específico sobre los aspectos nutricionales de ellos.

Lo mismo ocurriría si quisiéramos formar 4 grupos de igual tamaño de estudiantes (cuartiles) de acuerdo a su escolaridad, de forma de poder saber, más allá de los promedios obtenidos por cada uno de ellos, su posición relativa al conjunto de la generación.

a) Ejemplo

Como ejemplo para esta Guía, utilizaremos la Encuesta Continua de Hogares (ECH) que realiza el Instituto Nacional de Estadística (INE). Dicha encuesta se practica en todo el país urbano de forma continua, es decir a lo largo de todo el año, desde hace más de 20 años consecutivos, proporcionando información para el cálculo de indicadores de la actividad laboral, de los ingresos de las personas y los hogares, sobre los años de educación alcanzados por las personas, etc. Para mayor información acerca de la metodología de dicha encuesta puede consultarse:

<http://www.ine.gub.uy/biblioteca/metodologias/ech/metodologiaech.htm>

Procedimientos

Existen diversos mecanismos para lograr obtener grupos de igual tamaño de individuos dependiendo de la distribución de una variable de la muestra:

- 1- Observar dicha distribución y luego transformar la variable en otra con nuevos códigos.
- 2- Solicitarle al SPSS que agrupe automáticamente a los individuos de acuerdo a su posición relativa en la distribución de una variable

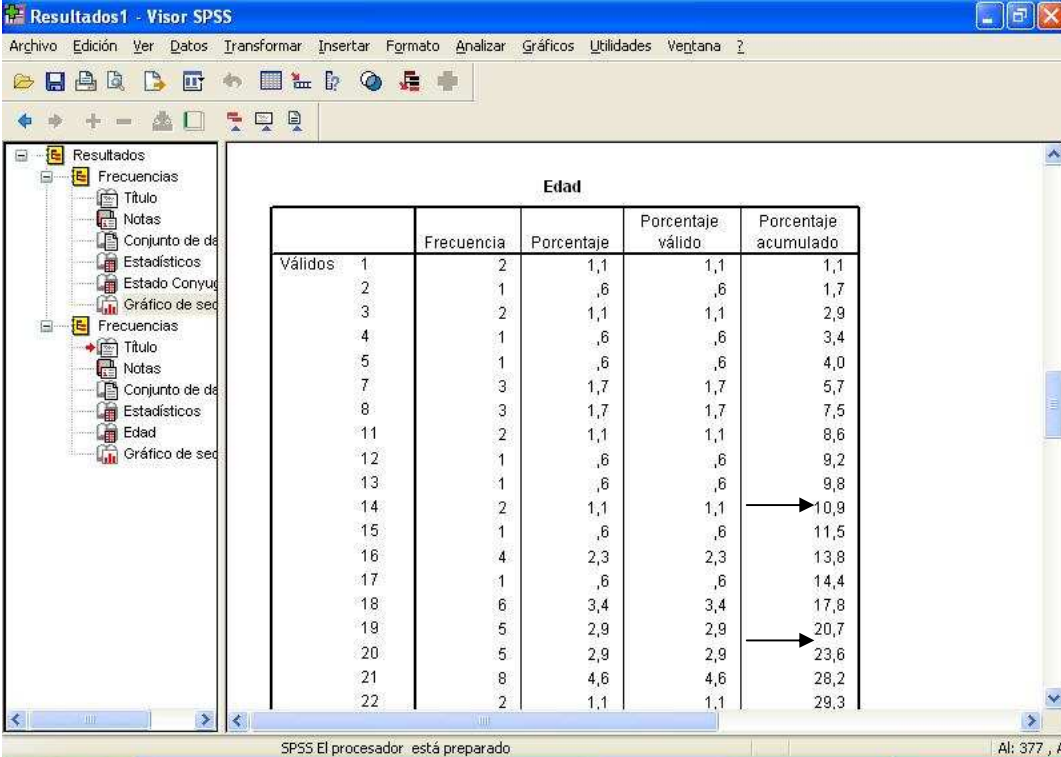
Rutina 1

Existen dos modalidades para realizar el procedimiento de forma manual:

a- *Analizar \ Estadísticos descriptivos \ Frecuencias*¹ y allí selecciono la variable de interés, por ejemplo edad del individuo.

Si ejecuto el procedimiento de esta forma, la sintaxis y output/resultado que se obtienen son los siguientes:

```
FRECUENCIAS  
VARIABLES=e2  
/ORDER= ANALYSIS .
```



The screenshot shows the SPSS Results window for 'Resultados1 - Visor SPSS'. The main content is a table titled 'Edad' with columns for 'Frecuencia', 'Porcentaje', 'Porcentaje válido', and 'Porcentaje acumulado'. The data is as follows:

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos 1	2	1,1	1,1	1,1
2	1	,6	,6	1,7
3	2	1,1	1,1	2,9
4	1	,6	,6	3,4
5	1	,6	,6	4,0
7	3	1,7	1,7	5,7
8	3	1,7	1,7	7,5
11	2	1,1	1,1	8,6
12	1	,6	,6	9,2
13	1	,6	,6	9,8
14	2	1,1	1,1	10,9
15	1	,6	,6	11,5
16	4	2,3	2,3	13,8
17	1	,6	,6	14,4
18	6	3,4	3,4	17,8
19	5	2,9	2,9	20,7
20	5	2,9	2,9	23,6
21	8	4,6	4,6	28,2
22	2	1,1	1,1	29,3

A partir de esta salida, como ya fue mencionado, podría conformarse otra variable, mediante el comando recodificar², que asuma el código 1 cuando el individuo tiene menos de 14 años (primer decil, ya que el % acumulado hasta esa edad indica que reúne poco más del 10% de la población). La misma variable deberá asumir el código 2 cuando tiene entre 15 y 19, y así sucesivamente.

¹ Ver autotexto de análisis univariado

² Ver autotexto Recodificar y Calcular

b- Si por otro lado, en el mismo menú de *Frecuencias* entro al sub menú *Estadísticos*, allí puedo solicitarle al SPSS el cálculo de percentiles directamente como estadísticos.



Aquí yo podría solicitar:

- a- El cálculo por defecto de cuartiles (puntos de corte para cuatro grupos iguales)
- b- Seleccionar una cantidad de puntos de corte para n grupos iguales (quintiles, deciles)
- c- Seleccionar distintos percentiles puntuales (puntos de cortes para el 23%, 47%, 78%, etc.)

En este caso, debemos tickear la opción b y solicitar puntos de corte para 10 grupos iguales.

Si ejecuto el procedimiento de esta forma, los agregados en la sintaxis y el output/resultado en relación a la modalidad a son los siguientes:

```
FRECUENCIAS
VARIABLES=e2
/NTILES= 10
/ORDER= ANALYSIS .
```

Estadísticos

Edad		
N	Válidos	174
	Perdidos	0
Percentiles	10	13,50
	20	19,00
	30	23,00
	40	28,00
	50	35,00
	60	43,00
	70	52,50
	80	61,00
	90	74,50

Luego de ello debería pasar a recodificarse³ la variable en estas nuevas categorías.

Pero este procedimiento además de ser más largo y engorroso, puede conducir a errores, por ejemplo en la etapa de recodificación.

Rutina 2

Hay una forma más directa de hacerlo, de manera que sea el propio programa quien se encargue del cálculo. Para ello, seleccionamos en el menú *Transformar* la opción *Asignar rangos a casos*.

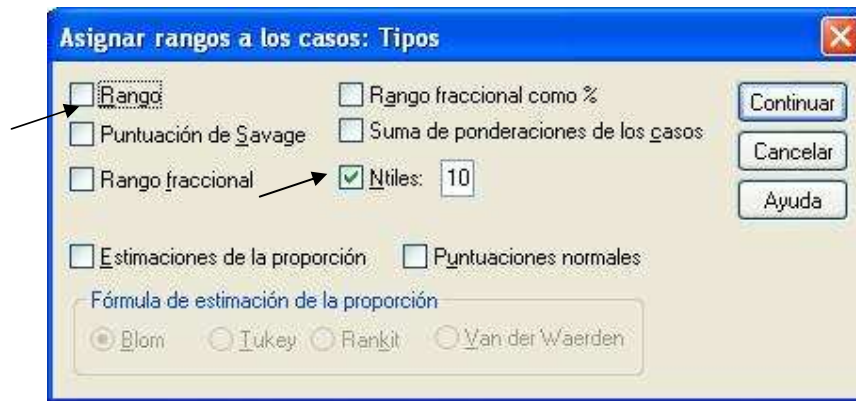


Como primer paso incorporo en el recuadro derecho la variable de interés, en este caso la edad de las personas.

Una vez que seleccionamos la edad, debemos de clicar donde dice *Tipos de rango* para marcar el tipo de percentil que queremos calcular. En el ejemplo,

³ Ver autotexto remodificación y calcular

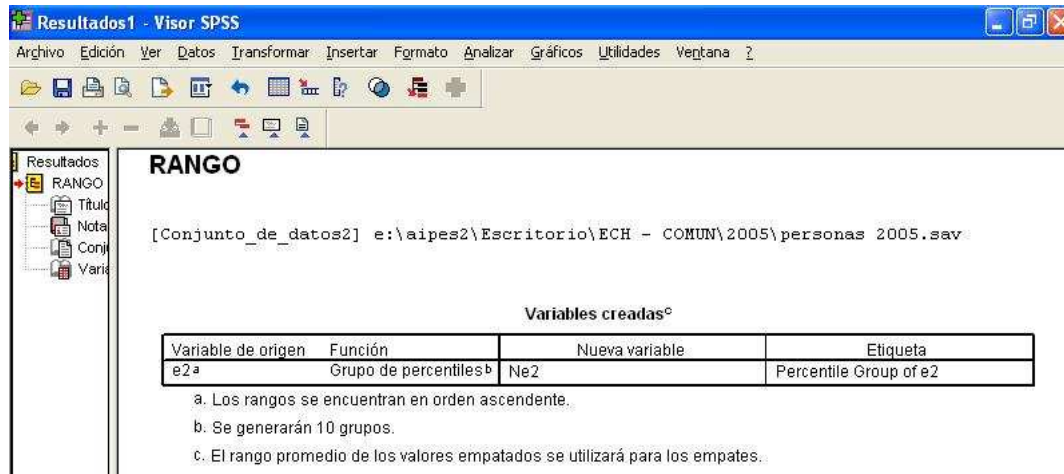
marcaremos 10 para el cálculo de deciles. Esta opción se encuentra en el icono *Ntiles*. Luego deseccionamos la opción *Rango* y presionamos *Continuar* y *Pegar* en la nueva pantalla.



La sintaxis que nos aparecerá será la siguiente:

```
RANK
VARIABLES=e2 (A) /NTILES (10) /PRINT=YES
/TIES=MEAN .
```

Además, en la pantalla de resultados deberá indicar que generó 1 variable nueva: "Ne2" que nos indica el valor del decil para cada caso.



Para corroborar y ver la distribución de las personas de acuerdo a cada decil basta con obtener una frecuencia de esta segunda variable:

The screenshot shows the SPSS Results Viewer window titled 'Resultados1 - Visor SPSS'. The main content area displays the following statistics for 'Percentile Group of e2':

Estadísticos	
Percentile Group of e2	
N Válidos	174
Perdidos	0

Percentile Group of e2					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	1	17	9,8	9,8	9,8
	2	19	10,9	10,9	20,7
	3	15	8,6	8,6	29,3
	4	20	11,5	11,5	40,8
	5	14	8,0	8,0	48,9
	6	20	11,5	11,5	60,3
	7	17	9,8	9,8	70,1
	8	17	9,8	9,8	79,9
	9	18	10,3	10,3	90,2
	10	17	9,8	9,8	100,0
	Total	174	100,0	100,0	

At the bottom of the window, it states: '4 elementos seleccionados (1 ocultos/contiene SPSS El procesador está preparado'.

Observación

Obsérvese que la distribución de los deciles indica que en cada uno de ellos hay aproximadamente un 10% de los casos. ¿Por qué no marca un 10% absoluto que es lo que esperaríamos? Sucede que no siempre la distribución de la variable original permite “cortes” exactos. Si seguimos con el ejemplo de la edad y tomamos el primer decil (9.8%) mediante una Tabla de doble entrada podríamos observar que contempla a las personas cuyas edades van de 0 a 13 años. Ahora bien, si al adicionar al primer decil a los niños de 14 años de edad superamos el 10% previsto, ¿qué hacemos con ellos? El programa busca la “mejor” solución, intentando equilibrar la composición de todos los deciles de manera que cada uno contenga un valor aproximado al 10%. Si observamos toda la distribución obtenida en el output puede corroborarse que los porcentajes no se alejan mucho de ese valor.

Ejercicios

1. Calcular los quintiles del ingreso total de los hogares
Obtener una distribución simple de dicha variable y los valores mínimo y máximo del ingreso que contiene cada quintil
2. Calcular los cuartiles de edad
Obtener una distribución simple de dicha variable y los valores mínimo y máximo de los años de educación que contiene cada cuartil
3. ¿Qué diferencias encuentra en la distribución de ambas variables generadas?