

## Temas de Estadística Práctica

Antonio Roldán Martínez

Proyecto <http://www.hojamat.es/>

### Tema 2: Medidas de tipo paramétrico

#### Documento

#### Datos continuos (o agrupados en intervalos)

Los datos que pueden tener tantos decimales como se desee y que entre cada dos de ellos siempre puede haber otro, se llaman **continuos**. Al poder estar muy cerca unos de otros, no se pueden estudiar de uno en uno y se agrupan en **intervalos**.

Son continuos el **peso**, la **estatura**, el **nivel de glucosa** en la sangre, etc.

Estos datos se suelen estudiar clasificados en intervalos: De 20 a 30 kg. , de 30 g a 40 g., de 120 cm. a 140 cm. , etc.

La primera operación que hay que aprender es la de **agrupar** datos. Por ejemplo, estudiaremos la siguiente serie de pesos (en gramos) de frutas tomadas de un montón de ellas en una frutería:

230	235,5	240	221	223	245	245,6	202
198	210,7	220,3	262	256	244	241	188,6
275,4	244,2	270	268,4	238,7	236	233,7	225,4
243,4	243,5	221,4	255,3	221	232	228	243

235 235 226 229 239 237 230,2 235,4

Aquí es muy difícil asignar una frecuencia a cada caso, pues nos saldrían tablas muy grandes y frecuencias muy pequeñas. Por ejemplo ¿qué frecuencia tiene el peso 244? ¿y el peso 268,4?: Siempre nos daría 1 ó 2 y no sería manejable. Es preferible preguntar: ¿Cuántos pesos hay entre 235 y 240 gramos?. De esta forma las frecuencias son más grandes y se pueden estudiar mejor.

La lista anterior se puede agrupar en **intervalos** de pesos de **diez en diez gramos**: Las que pesan entre 200 y 210, entre 210 y 220, etc. La fruta más pequeña pesa 188,6 gramos, luego podíamos poner como **inicio** de la tabla 190 g., y como final 280 g.

El número de intervalos será el resultado de dividir 300-180 entre 10, es decir, 12. La tabla final podría tener este aspecto:

Intervalos	Frecuencia
180-190	
190-200	
200-210	
210-220	
220-230	
230-240	
240-250	
250-260	
260-270	
270-280	

Ahora habría que ir contando las frutas que entran en cada intervalo. Hazlo con cuidado para ver qué frecuencias te salen. Si un peso está en el límite de dos intervalos, como

230 g, lo situaremos en el intervalo siguiente 230-240 y no en el anterior 220-230. Esto es por convenio.

Puedes corregir tu trabajo con el programa **LibreOffice**. Inicia el programa y abre un archivo nuevo. Escribe los pesos de las frutas en la parte izquierda de la hoja. Para agruparlos de diez en diez, escribe a su lado, en columna, los intervalos que queremos formar.

Para ello escribe 180 en una celda, selecciona hacia abajo y después pide **Editar Rellenar serie** y la concretas como **lineal de intervalo 10 y límite 300**. Te deberá dar una columna de números desde el 180 al 300.

Para agrupar datos usa el modelo **recoge3.ods** (carpeta tema2\open). Para trabajar con rapidez, puedes tener abiertos a la vez *recoge3.ods* y *datoscont.ods*. Como en este último los datos de las frutas están escritos en columna, puedes ir seleccionando grupos de 10 e ir copiando y pegando en *recoge3.sxc*. Como mínimo de la tabla escribe 180, como máximo 280 y número de intervalos 10.

Escribe el resultado con bolígrafo en la tabla de arriba.

Los datos continuos se dibujan mediante un **Histograma**, que es como un diagrama de barras que están adosadas unas a otras formando un área continua. En **LibreOffice** se sustituyen por los **diagramas de áreas**, que equivalen a los **polígonos de frecuencias**, que son muy parecidos, pero con aspecto de una línea curva o quebrada continua y cuya área representa las frecuencias. Para dibujarlo en este caso selecciona ese tipo de gráfico **De áreas**. Te resultará el gráfico continuo correspondiente. Debes añadirle todos los rótulos y títulos convenientes y pegarlo en la Hoja.

Obtendrás un gráfico parecido al de la figura.



Se parece a una campana que, como recordarás, sugiere que la distribución es parecida a la **Normal**, que es la distribución que siguen muchos fenómenos naturales, como el peso, la inteligencia, el volumen del cráneo, etc.

Para saber si el ajuste a la campana es bueno, copia los datos con sus frecuencias, abre la hoja **normali.ods** (**carpeta tema2/open**), que tendrás en tu directorio, pega los datos donde se te indique y leerás el grado de ajuste, que aquí es de un 5%, lo que es aceptable, aunque no óptimo. Busca el gráfico comparativo para ver bien el grado de semejanza entre la teoría y la práctica. Observarás que existe un cierto parecido entre ambas gráficas.

### Otro ejemplo de datos continuos

Supongamos que la tabla siguiente describe la distribución de pesos de los alumnos y alumnas de un curso:

Peso	Frecuencia
40-50	2
50-60	4
50-70	15

70-80	8
80-90	1
100-110	1

Para estudiarlos con **LibreOffice** abre de nuevo una hoja y escribe los datos sustituyendo los intervalos 40-50, 50-60, etc. por sus **puntos medios**, que también se llaman **marcas de clase**, porque son los representantes de todo el intervalo. Para calcular la media o la desviación típica usaremos esos puntos medios como valor de **x**.

Calcula así el coeficiente de variación de esta tabla. Te debe resultar un 13,3%.

Dibuja su gráfica de áreas o con una línea en 3D. ¿Se parece más o menos a la campana de Gauss que el ejemplo anterior? Compruébalo con la hoja **normalidad** y escribe tu opinión:

### Otro ejemplo

La variable **Inteligencia** funciona como continua. Los datos que ves debajo corresponden a las puntuaciones de un test de inteligencia.

120 100 110 102 98 142 121 88 89 93 95 95  
101 106 93 112 93 77 115 102 101 94 118 86

Datos	Frec.


Forma una tabla de frecuencias con esos datos contenidos en Datoscont.ods, a partir de los

intervalos 80-90, 90-100, 100-110, etc. Cópiala a la derecha:

Sustituye cada intervalo por su punto medio y calcula la media, la desviación típica y escríbelas aquí.

Media:

Desviación típica:

Construye un histograma continuo y estudia su grado de ajuste a la curva normal.

Comenta todo lo que observes:

## Ejercicio final

Abre la hoja **normali.ods** (carpeta tema2/open) e inventa unos intervalos y frecuencias que tengan el mejor ajuste posible a la curva normal (basta que llegues a un 60%, más o menos)

Rellena esta tabla con tu ejemplo:

