

## TEMA 11: ESTADÍSTICA

### TABLAS Y GRÁFICOS

El objetivo de la estadística es la recopilación, organización, proceso, análisis e interpretación de datos. Para ello es necesario seguir los siguientes pasos:

- 1º. Decidir lo que se quiere estudiar y para qué se necesita.
- 2º. Enumerar los posibles valores que puede tomar la variable.
- 3º. Recoger los datos.
- 4º. Tabularlos y representarlos gráficamente.
- 5º. Analizar los resultados y extraer conclusiones.

La tabla de frecuencias es una forma de ordenación de los datos estadísticos. En ella podemos encontrar:

- 1.- La frecuencia absoluta: número de veces que aparece un determinado valor. ( $f_i$ )
- 2.- La frecuencia absoluta acumulada: suma de las frecuencias absolutas de todos los valores inferiores o iguales al valor considerado. ( $F_i$ )
- 3.- La frecuencia relativa: cociente entre la frecuencia absoluta de un valor y el número total de datos. ( $n_i$ )
- 4.- La frecuencia relativa acumulada: suma de las frecuencias relativas de todos los valores inferiores o iguales al valor considerado. ( $N_i$ )
- 5.- La frecuencia porcentual: la frecuencia relativa en tanto por ciento. ( $p_i$ )

Sobre los gráficos estadísticos señalar que existen una gran cantidad de ellos. Destacar:

- 1.- Diagrama de barras: Indicado para variables discretas. Formado por barras rectangulares de igual ancho que sirven para mostrar y comparar frecuencias.
- 2.- Histograma: Indicado para variables continuas que toman valores en un intervalo. Similares a los diagramas de barras con la diferencia que en este caso las barras rectangulares están unidas.
- 3.- Diagrama de sectores: Su forma es una circunferencia dividida en sectores, por medio de radios que dan la sensación de un pastel cortado en porciones.

#### Ejemplo 1:

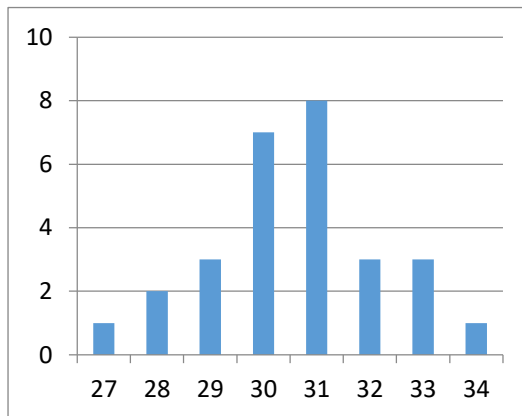
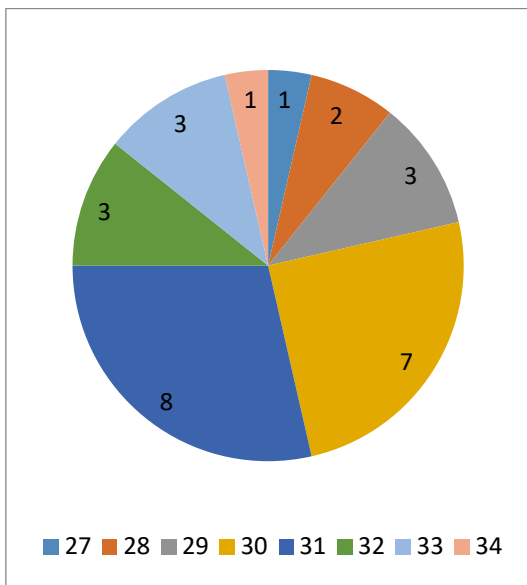
*Durante el mes de julio, en una ciudad se han registrado las siguientes temperaturas máximas:*

32, 31, 28, 29, 33, 32, 31, 30, 31, 31, 27, 28, 29, 30, 32, 31, 31,  
30, 30, 29, 29, 30, 30, 31, 30, 31, 34, 33, 33, 29, 29.

- a) Realizar un recuento y confeccionar una tabla de frecuencias.
- b) Realizar una representación adecuada de los resultados.

Solución:

| Temperatura | fi | Fi | ni    | Ni    | pi     |
|-------------|----|----|-------|-------|--------|
| 27          | 1  | 1  | 0,032 | 0,032 | 3,20%  |
| 28          | 2  | 3  | 0,065 | 0,097 | 6,50%  |
| 29          | 3  | 9  | 0,194 | 0,29  | 19,40% |
| 30          | 7  | 16 | 0,226 | 0,516 | 22,60% |
| 31          | 8  | 24 | 0,258 | 0,774 | 25,80% |
| 32          | 3  | 27 | 0,097 | 0,871 | 9,70%  |
| 33          | 3  | 30 | 0,097 | 0,968 | 9,70%  |
| 34          | 1  | 31 | 0,032 | 1     | 3,20%  |
| Total       | 31 |    | 1     |       |        |



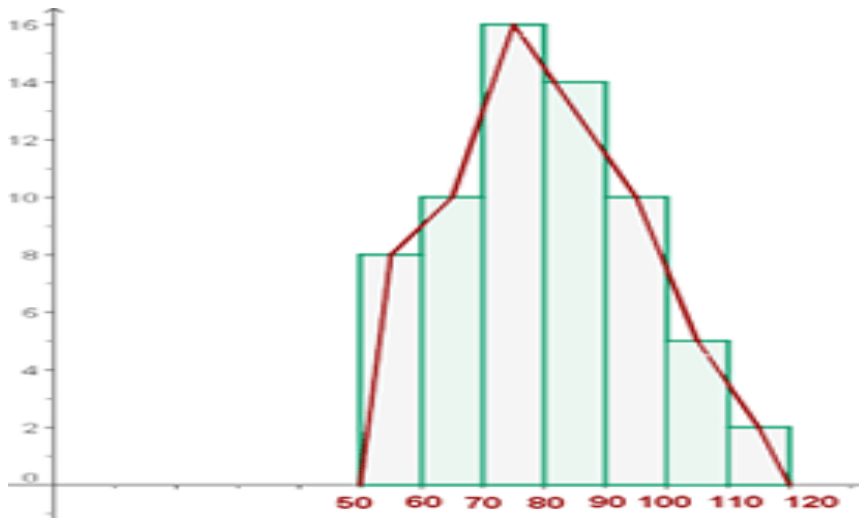
Ejemplo 2:

Los pesos de los 65 empleados de una fábrica vienen dados por la siguiente tabla:

|       |         |         |         |         |          |           |           |
|-------|---------|---------|---------|---------|----------|-----------|-----------|
| $X_i$ | [50,60) | [60,70) | [70,80) | [80,90) | [90,100) | [100,110) | [110,120) |
| $f_i$ | 8       | 10      | 16      | 14      | 10       | 5         | 2         |

Realiza un gráfico adecuado para analizar los resultados.

Solución:



**EJERCICIOS**

**EJERCICIO 1 :** En un grupo de personas hemos preguntado por el número medio de días que practican deporte a la semana. Las respuestas han sido las siguientes:

4    2    3    1    3    7    1    0    3    2    6    2  
 3    3    4    6    3    4    3    6    0    2    3    2

- a) Haz una tabla de frecuencias completa.
- b) Representa gráficamente la distribución.

**EJERCICIO 2 :** En una clase de 4º ESO hemos preguntado a las alumnas y a los alumnos por las horas de estudio que dedican a la semana. Estas han sido las respuestas:

16    11    17    12    10    5    1    8    10    14    15    20  
 3    2    5    12    7    6    3    9    10    8    10    6  
 16    16    10    3    4    12

- a) Ordena los datos en una tabla de frecuencias, agrupándolos en intervalos de la forma que creas más conveniente.
- b) Representa gráficamente la distribución.

**EJERCICIO3** : El número de ordenadores que hay en los hogares de un grupo de personas, A, viene dado en la siguiente tabla:

|                |    |    |    |   |   |
|----------------|----|----|----|---|---|
| Nº Ordenadores | 0  | 1  | 2  | 3 | 4 |
| Nº Personas    | 15 | 22 | 10 | 2 | 1 |

- Completa la tabla de frecuencias.
- Realiza una representación gráfica adecuada.

**EJERCICIO 4** : Midiendo el peso, en kilogramos, de los niños y las niñas de un determinado grupo, todos ellos de la misma edad, hemos obtenido los siguientes resultados:

|             |         |         |         |         |         |
|-------------|---------|---------|---------|---------|---------|
| Peso (kg)   | [10,13) | [13,16) | [16,19) | [19,22) | [22,25) |
| Nº niños/as | 6       | 50      | 32      | 9       | 3       |

- Realiza una tabla de frecuencias completa.
- Representa de forma adecuada la distribución.

**EJERCICIO 5** : El peso de unos paquetes de café, en gramos, es:

108, 98, 92, 102, 100, 101, 99, 100, 93, 109, 90, 97, 99, 100, 105, 104, 98, 102, 101, 97, 95, 100, 102, 103, 104, 97, 98, 101, 100, 102, 95, 96, 97, 99, 100.

Forma la tabla de frecuencias absolutas, relativas y acumuladas agrupando los datos de 5 en 5 gramos.

## PARAMETROS DE CENTRALIZACIÓN Y DISPERSIÓN

Parámetros centrales: Valores alrededor de los cuáles se distribuyen los datos.

1-Media aritmética: se calcula sumando todos los valores obtenidos de la variable estudiada y dividiéndolos por el número de datos que haya.

2-Mediana: valor que ocupa el lugar del medio, el central, después de haber ordenado todos los datos de menor a mayor. Si el número de datos es impar, sólo habrá uno en el medio. Si es par, habrá dos, y la moda será la media aritmética de ambos.

3- Moda: valor que más se repite, el que mayor frecuencia absoluta tiene. Puede haber más de una moda.

La media y la mediana en forma conjunta, sirven para valorar la simetría de los datos. Si sus valores coinciden, la distribución es completamente simétrica.

Parámetros de dispersión: Valores que tratan de analizar como de juntos o separados se encuentran los datos.

1-Recorrido: Diferencia entre el valor máximo y mínimo de la muestra.

2-Desviación media: Promedio de las distancias de los datos a la media.

$$DM = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \dots + |x_n - \bar{x}|}{n} = \frac{\sum |x_i - \bar{x}|}{n}$$

3-Varianza: Promedio del cuadrado de las distancias de los valores a la media.

$$V = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

4-Desviación típica: Raíz cuadrada de la varianza.

Parámetros de posición: Valores que dividen una muestra de datos en cuatro partes iguales.

1-Primer cuartil: Q1. Valor que deja el 25% de los datos a la izquierda y el 75% restante a la derecha.

2-Segundo cuartil: Q2. Coincide con el valor de la mediana.

3-Tercer cuartil: Q3. Valor que deja el 75% de los datos a la izquierda y el 25% restante a la derecha.

La representación de dichos valores se realiza con los diagramas de caja.

Ejemplo 1:

Calcula el valor de todos los parámetros estadísticos vistos en el siguiente grupo de datos:

1, 1, 2, 2, 2, 2, 3, 4, 4, 5, 5, 6, 7

Solución:

$$\text{Media} = \frac{1+1+2+2+2+2+3+4+4+5+5+6+7}{13} = 3.38$$

Mediana = 3    Cuartil1 = 2    Cuartil3 = 5

Moda = 2

Rango = 7 - 1 = 6

$$\text{Desviación Media} = \frac{|1-3,38|+|1-3,38|+|2-3,38|+|2-3,38|+\dots+|7-3,38|}{13} = 1.64$$

$$\text{Varianza} = \frac{(1-3,38)^2+(1-3,38)^2+(2-3,38)^2+(2-3,38)^2+\dots+(7-3,38)^2}{13} = 3.46$$

$$\text{Desviación típica} = \sqrt{\text{Varianza}} = \sqrt{3.46} = 1.86$$

Ejemplo 2:

Sea una distribución estadística que viene dada por la siguiente tabla:

|           |    |    |    |    |    |
|-----------|----|----|----|----|----|
| <i>xi</i> | 61 | 64 | 67 | 70 | 73 |
| <i>fi</i> | 5  | 18 | 42 | 27 | 8  |

Calcular:

- a) La moda, mediana y media.
- b) El rango, desviación media, varianza y desviación típica.

| <i>Xi</i> | <i>fi</i> | <i>Fi</i> | <i>Xi · fi</i> | $ Xi - \bar{x} $ | $ Xi - \bar{x}  \cdot fi$ | $Xi^2 \cdot fi$ |
|-----------|-----------|-----------|----------------|------------------|---------------------------|-----------------|
| 61        | 5         | 5         | 305            | 6.45             | 32.25                     | 18605           |
| 64        | 18        | 23        | 1152           | 3.45             | 62.10                     | 73728           |

|    |     |     |      |      |        |        |
|----|-----|-----|------|------|--------|--------|
| 67 | 42  | 65  | 2814 | 0.45 | 18.90  | 188538 |
| 71 | 27  | 92  | 1890 | 2.55 | 68.85  | 132300 |
| 73 | 8   | 100 | 584  | 5.55 | 44.40  | 42632  |
|    | 100 |     | 6745 |      | 226.50 | 455803 |

Moda = 67

Mediana= (100/2=50) 67

Media=6745/100=67.45

Desviación media= 226.5/100=2.265

Rango=73-61=12

Varianza= 455803/100 - 67.45<sup>2</sup> =8.53

Desviación típica =  $\sqrt{8.53} = 2.92$

Ejemplo 3:

Preguntando la edad a un conjunto de 20 personas hemos obtenido los siguientes resultados:

36 25 37 24 39 20 36 45 31 31  
39 24 29 23 41 40 33 24 34 40

Realiza una representación de un diagrama de caja y bigote con ellos.

Para el cálculo de los parámetros estadísticos necesarios para realizar el diagrama, comenzamos ordenando los datos:

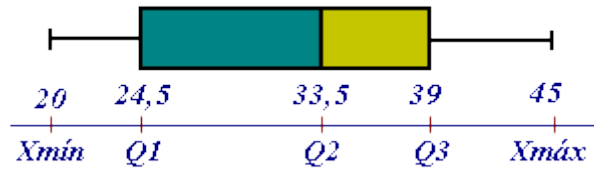
20 23 24 24 24 25 29 31 31 33 34 36 36 37 39 39 40 40 41 45

Calculamos el valor de Q1 teniendo en cuenta que es el valor mayor que deja el 25% de los valores de la distribución. Como n = 20, resulta que n/4=5 por lo que Q1= (24+25)/2=24,5.

Calculamos Q2, la mediana de la distribución, valor de la variable que ocupa el valor central. Como n=20, n/2=10 por lo que Q2=(33+34)/2=33,5.

Por último, Q3 es el valor que sobrepasa al 75% de los valores de la distribución. En nuestro caso 3n/4=15 por lo que Q3=(39+39)/2=39.

El diagrama de caja y bigote resultante es el siguiente:



## EJERCICIOS

**EJERCICIO 1 :** La dirección de tráfico ha recogido la siguiente información sobre las multas diarias a los coches que circulan por una carretera:

| Numero de multas | Dias |
|------------------|------|
| (0,5]            | 6    |
| (5,10]           | 14   |
| (10,15]          | 20   |
| (15,20]          | 10   |

- Obtén el número medio de multas diarias impuestas.
- Calcula e interpreta la mediana de las multas.
- Calcula la moda de esta distribución.

**EJERCICIO 2 :** Dibuja el polígono de frecuencias acumuladas, y calcula media, moda, mediana y cuartiles para la tabla:

| Intervalo | Frecuencia |
|-----------|------------|
| (0,10]    | 6          |
| (10,20]   | 7          |
| (20,30]   | 10         |
| (30,40]   | 9          |
| (40,50]   | 5          |

**EJERCICIO 3 :** La siguiente tabla relaciona el número de goles marcados en partidos de fútbol:

| Goles | Partidos |
|-------|----------|
| 0     | 12       |
| 1     | 16       |
| 2     | 22       |
| 3     | 20       |
| 4     | 21       |
| 5     | 4        |
| 6     | 2        |



- a) Calcula la media, mediana y la moda.
- b) Calcula la varianza y la desviación típica.

**EJERCICIO 4 :** Las puntuaciones obtenidas en un test de inteligencia por 25 personas han sido:

2 4 4 3 5 1 4 3 4 1 3 4 5 1 1 2 1 3 4 5 3 4 3 2 1

- a) Calcula la puntuación media
- b) ¿Cuál es la moda de la distribución?
- c) Calcula la mediana
- d) Calcula la desviación típica y el rango.
- e) Realiza la representación de diagrama caja y bigote con ellos.

**EJERCICIO 5 :** Se ha preguntado a un grupo de 70 personas por el número de zapato. Los resultados vienen dados por:

| Número de zapato | Número de personas |
|------------------|--------------------|
| 35               | 4                  |
| 36               | 15                 |
| 37               | 17                 |
| 38               | 20                 |
| 40               | 10                 |
| 42               | 4                  |

- a) Realiza un diagrama de barras
- b) Halla la media, moda y mediana.
- c) Halla los cuartiles.
- d) Halla el rango.
- e) Halla la desviación típica y la varianza.

## TABLAS DE DOBLE ENTRADA

Una tabla de doble entrada es una tabla de datos que hace referencia a dos variables. En la cabecera de las filas, establecemos los valores de una de ellas, y en la cabecera de las columnas, los valores de la otra. Es una forma de resumir la información sobre dos variables en la misma población.

### Ejemplo 1:

En una clase con 32 estudiantes se pregunta que prefieren jugar, un partido de fútbol o uno de baloncesto. En la siguiente tabla se muestran los resultados:

| Deporte/Estudiantes | Chicos | Chicas | TOTAL |
|---------------------|--------|--------|-------|
| Fútbol              | 15     | 3      | 18    |
| Baloncesto          | 4      | 10     | 14    |
| TOTAL               | 19     | 13     | 32    |

- ¿Qué significa el 14?
- ¿Qué significa el 13?
- ¿Qué significa el 4?

### Solución:

- Que catorce estudiantes prefieren jugar al baloncesto que al fútbol.
- Que hay un total de trece chicas entre los treinta y dos estudiantes.
- Que 4 chicos prefieren jugar un partido de baloncesto a uno de fútbol.

### Ejemplo 2:

En una clase con 36 estudiantes se pregunta por su color favorito, azul o rosa. De un total de 19 chicas, 16 prefieren el rosa, los chicos hay 15 que prefieren el azul. Construye una tabla de doble entrada.

### Solución:

| Color\Estudiantes | Chicos | Chicas | TOTAL |
|-------------------|--------|--------|-------|
| Azul              | 15     | 3      | 18    |
| Rosa              | 2      | 16     | 18    |
| TOTAL             | 17     | 19     | 36    |

## EJERCICIOS

**EJERCICIO 1 :** Los datos obtenidos al estudiar las variables X = “número de goles marcados” e Y = “número de goles recibidos”, en 40 partidos jugados por el equipo campeón de la liga de fútbol sala, son:

(5, 4), (4, 2), (6, 3), (4, 4), (3, 2), (6, 4), (3, 1), (4, 2), (4, 2), (6, 4), (4, 2), (5, 3), (3, 1), (2, 2), (4, 3), (3, 1), (4, 2), (5, 3), (5, 3), (4, 2), (3, 3), (1, 1), (4, 2), (5, 3), (3, 2), (5, 3), (6, 4), (4, 2), (5, 3), (2, 1), (3, 2), (6, 4), (5, 3), (4, 2), (4, 2), (3, 3), (3, 1), (2, 2), (6, 4), (5, 3)

Elabora la tabla de doble entrada.

**EJERCICIO 2 :** En una clase compuesta por 30 alumnos, se ha hecho un estudio sobre el número de horas diarias de estudio X y el número de suspensos Y, obteniéndose los siguientes resultados:

(2, 0), (2, 2), (0, 5), (2, 1), (1, 2), (2, 1), (3, 1) (4, 0), (0, 4), (2, 2) (2, 1), (2, 1), (4, 0), (3, 1), (2, 4), (2, 1), (1, 2), (2, 1), (2, 0), (3, 0) (3, 2), (2, 2), (2, 2), (2, 1), (0, 5), (1, 3), (2, 2), (2, 1), (1, 3), (1, 4)

Construye la tabla estadística bidimensional de doble entrada.

**EJERCICIO 3 :** En una clase hay doce chicas y 18 chicos. De las chicas, 3 usan gafas y hay 13 chicos que no usan gafas. Completa la siguiente tabla:

|              | Chicos | Chicas | <b>TOTAL</b> |
|--------------|--------|--------|--------------|
| Con gafas    |        |        |              |
| Sin gafas    |        |        |              |
| <b>TOTAL</b> |        |        |              |