Todos iguales, todos diferentes

No sé si te habrás parado a pensar que todos formamos parte de estudios de empresas, gobiernos o instituciones. Si se tiene trabajo o se está en el paro; si se tiene tal o cual electrodoméstico; si has votado a uno u otro partido político, etc. Todos estamos incluidos en números y gráficas que luego aparecen en los medios de comunicación. ¿Quieres saber a qué nos referimos?
¡Adelante!

Bloque IV

Módulo IV

Unidad 5

Índice

1. Concepto

Definir la estadística es relativamente fácil, basta con acudir a cualquier libro de matemáticas o al diccionario y allí la tenemos. Pero lo realmente importante es conocer que esta parte de las matemáticas está presente continuamente en nuestra vida diaria puesto que se encarga del recuento, ordenación y clasificación de los datos obtenidos en las observaciones para poder hacer comparaciones y sacar conclusiones.

Ejemplo: Todos sabemos que hoy en día para sacar una nueva línea de perfumes, por ejemplo, hacen estudios de mercado, de consumo, de poder adquisitivo de los consumidores, etc. Se obtienen datos, se ordenan y clasifican y tras su análisis, se decide sacar a la venta el nuevo perfume o no.

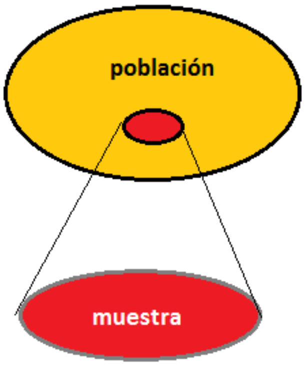
Ampliación: *En España existe un organismo oficial, el INE –Instituto Nacional de Estadística- que se encarga de recopilar datos de muy diversa índole para conocer un poco mejor cómo somos. Seguro que mes a mes has visto o escuchado en los medios de comunicación los datos de la inflación o del paro, por ejemplo..*

1.1 Elementos de un estudio estadístico

Antes de realizar un estudio estadístico debemos conocer los diferentes elementos que lo componen: población, individuo y muestra.

- **Población:** es el conjunto de todos los elementos a los que se somete a un estudio estadístico.

  Ejemplo: si deseamos hacer un estudio sobre la obesidad infantil en un pueblo, la población serán los niños de esa ciudad comprendidos entre las edades que se decida en el estudio.



- **Individuo**: es cada uno de los sujetos a los que se refiere el estudio estadístico.

  Ejemplo: en el caso anterior, cada uno de los niños de la población.

- **Muestra:** es un subconjunto representativo de la población. Como ésta suele ser muy amplia, se toma una muestra representativa de la población. Lógicamente, la muestra siempre es menos que la población.

  Ejemplo: como el número de niños del pueblo será grande no se puede ir uno por uno pesándoles, así que se elige un número representativo de la población. Al emplear la palabra representativo quiere decirse que si en la población el 45% son niñas, en la muestra también.

1.2 Variables estadísticas

La **variable estadística** es cada una de las propiedades o características que podemos estudiar de la población. Cada uno de los posibles resultados que se pueden obtener se llama **valor** y el conjunto de los valores obtenidos son los **datos** de nuestro estudio estadístico.

Ejemplo: veamos un estudio estadístico que trate sobre lo que se obtiene al tirar una moneda al aire. Los **valores** que puede tener nuestro estudio son cara y cruz. Y si la lanzamos cinco veces al aire los **datos** podrán ser: cara, cruz, cruz, cara, cruz.

Las variables, dependiendo de los posibles valores que puedan tomar se clasifican en:

- **Variables cualitativas**. Los valores de la variable no son números sino cualidades como el color, la forma o el sexo.
- **Variables cuantitativas**. Los datos se expresan numéricamente como la edad, el precio de la gasolina o la altura. Y a su vez pueden ser:

  Discretas. Cada una de las variables solo puede tomar valores enteros (1, 2, 3...) como el número de hermanos (puedes tener 1, 2 ó 3 pero nunca 1,5), o el número de ventanas de una casa, etc.

  Continuas. Pueden tomar cualquier valor de un intervalo dado como por ejemplo el peso (una persona puede pesar 74 kg. y otra 75, pero entre medias puede haber muchas con pesos intermedios, 74,5 kg., 74,75 kg., etc).

Recordatorio:

Practica:

1. **Clasifica estas variables en cualitativas o cuantitativas**

a) Marca de coche, temperatura mínima del día, número de hijos, número del zapato, colores preferidos, nota del examen, signo del zodíaco.

b) En las anteriores variables hay cuantitativas, ¿cuáles de ellas son discretas y continuas?

2. Frecuencias y tablas de frecuencias

Una vez que hemos obtenido todos los datos de nuestro estudio estadístico, procedemos a ordenarlos en orden creciente o decreciente según convenga. Si los datos corresponden a una variable cuantitativa discreta, el proceso es muy sencillo (1, 2, 3, 4, etc., por ejemplo). Pero si se tratara de una variable cuantitativa continua, necesitaríamos establecer cuáles con sus respectivas marcas de clase pero eso lo estudiamos más adelante.

Seguidamente elaboraremos una tabla de valores a la que añadiremos las **frecuencias** y se convertirá en una tabla de frecuencias.

Ejemplo: vamos a realizar un estudio estadístico sobre el tamaño de los pies de los trabajadores de una pequeña empresa. Para ello, vemos el número de sus zapatos y obtenemos los siguientes valores: 42, 41, 45, 44, 43, 42, 41, 44, 43, 42, 40, 45, 43, 41, 42.

Hacemos el recuento, ordenamos los valores y confeccionamos la tabla:

Valores	40	41	42	43	44	45
Veces que se repiten	1	3	4	3	2	2

- ***Frecuencia absoluta***

Es el número de veces que se repite un valor determinado de nuestro estudio. Se representa por $\boldsymbol{f_i}$.

- ***Frecuencia relativa***

Es el cociente entre la frecuencia absoluta y el número total de datos obtenidos en nuestro estudio. Se representa por $\boldsymbol{n_i}$.

$$\boldsymbol{n_i = \frac{f_i}{n}}$$

- ***Frecuencia absoluta acumulada***

Es la suma de las frecuencias absolutas de los valores iguales o inferiores al valor tratado. Se representa por $\boldsymbol{F_i}$. Al final, la suma debe ser igual al número de datos.

- ***Frecuencia relativa acumulada***

Es la suma de las frecuencias relativas de los valores iguales o inferiores al valor tratado. Se representa por $\boldsymbol{N_i}$. Al final, la suma debe ser igual a 1.

Hagamos un ejemplo para comprender todo esto mejor.

Ejemplo: sigamos con el estudio estadístico sobre el tamaño de los pies de los trabajadores de una pequeña empresa**.** Hagamos una tabla de frecuencias en las que pongamos las anteriormente descritas.

Valores	F. absoluta f_i	F. absoluta acumulada F_i	F. relativa n_i	F. relativa acumulada N_i
40	**1**	**1**	$\frac{1}{15}$	$\frac{1}{15}$
41	**3**	(1+3) = **4**	$\frac{3}{15}$	$\frac{1}{15}+\frac{3}{15}=\frac{4}{15}$
42	**4**	(4+4) = **8**	$\frac{4}{15}$	$\frac{4}{15}+\frac{4}{15}=\frac{8}{15}$
43	**3**	(8+3) = **11**	$\frac{3}{15}$	$\frac{8}{15}+\frac{3}{15}=\frac{11}{15}$
44	**2**	(11+2) = **13**	$\frac{2}{15}$	$\frac{11}{15}+\frac{2}{15}=\frac{13}{15}$
45	**2**	(11+2) = **15**	$\frac{2}{15}$	$\frac{13}{15}+\frac{2}{15}=\frac{15}{15}=1$

La **frecuencia absoluta** es el número de veces que se repite un valor. En la **frecuencia absoluta acumulada** ponemos la frecuencia absoluta del primer valor y los sucesivos serán la suma de la suya más la del anterior. Al final da el número total de datos (15 en nuestro caso).

La **frecuencia relativa** es el cociente entre la frecuencia absoluta y el número total de datos (15). La frecuencia relativa acumulada, al igual que en la absoluta acumulada, ponemos la frecuencia relativa del primer valor y los demás se obtienen de sumar la suya con la del anterior. Al final conseguiremos tener 1.

- ***Frecuencia porcentual***

Es la frecuencia relativa expresada en %. Se representa por P_i **.** Nos indica el porcentaje de cada valor respecto del total.

Ejemplo: añadimos a nuestra tabla de frecuencias la porcentual

Valores	F. absoluta f_i	F. absoluta acumulada F_i	F. relativa n_i	F. relativa acumulada N_i	F. porcentual P_i
40	**1**	**1**	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}\bullet 100=\mathbf{6,66}\%$
41	**3**	**4**	$\frac{3}{15}$	$\frac{4}{15}$	$\frac{3}{15}\bullet 100=\mathbf{20,00}\%$
42	**4**	**8**	$\frac{4}{15}$	$\frac{8}{15}$	$\frac{4}{15}\bullet 100=\mathbf{26,66}\%$
43	**3**	**11**	$\frac{3}{15}$	$\frac{11}{15}$	$\frac{3}{15}\bullet 100=\mathbf{20,00}\%$
44	**2**	**13**	$\frac{2}{15}$	$\frac{13}{15}$	$\frac{2}{15}\bullet 100=\mathbf{13,33}\%$
45	**2**	**15**	$\frac{2}{15}$	$\frac{15}{15}=1$	$\frac{2}{15}\bullet 100=\mathbf{13,33}\%$

Según lo realizado, el 6,66% de los trabajadores de la empresa calza un 40 de zapatos. El 20% calza un 41 al igual que los que llevan el 43, y así con todos los demás.

Practica:

2. **Haz una tabla de frecuencias de este estudio estadístico:**

   Se lanza un dado 25 veces al aire y estos han sido los valores obtenidos:
   4, 1, 5, 6, 3, 2, 1, 5, 3, 6, 4, 3, 5, 2, 1, 6, 4, 3, 5, 2, 5, 3, 1, 2, 5

Hasta ahora hemos realizado una tabla de frecuencias con variable cuantitativa discreta, en donde los valores eran pocos y números enteros. Pero, ¿qué pasaría si tuviéramos una variable cuantitativa continua o que tuviera muchos valores en nuestro estudio estadístico? Tendríamos que agrupar los valores en intervalos de la misma longitud (**clase**) y hallar la marca de clase del intervalo. El resto, lo haríamos como hemos visto anteriormente.
Para entenderlo bien, veamos un ejemplo paso a paso.

Ejemplo: vamos a hacer un estudio estadístico sobre la edad de los trabajadores de una mediana empresa que cuenta con 40 empleados, a los que les preguntamos la edad y anotamos sus respuestas. Como imaginarás, los valores que vamos a obtener son muchos así que lo primero que haremos será agruparlos en intervalos.

Edades:

20	58	26	41	42	22	53	36	25	20	61	39	42	29	33	50	46	24	43	23
29	42	46	31	29	40	37	35	59	45	49	21	51	48	27	36	31	48	55	64

Ordenamos los valores de menor a mayor y vemos que la edad menor es 20 y la mayor, 64. Se pueden establecer intervalos de edad de 10 en 10.

Clase	Marca de clase
[20,30)	25
[30,40)	35
[40,50)	45
[50,60)	55
[60,70)	65

A la hora de escribir el intervalo o clase de 10 en 10, vemos que tiene por un lado tiene forma de **corchete [,** eso quiere decir que el número está <u>incluido</u> en el intervalo. Y por el otro lado tiene un **paréntesis)**, quiere decir que el número <u>no está incluido</u> en el intervalo. Por ejemplo: [20,30) significa que el intervalo va desde el 20 incluido hasta el 30 sin incluir.
La **marca de clase** es el valor central del intervalo.

Y ahora hacemos la tabla de frecuencias:

Clase	Marca de clase	F. absoluta f_i	F. absoluta acumulada F_i	F. relativa n_i	F. relativa acumulada N_i	F. porcentual P_i
[20,30)	25	12	12	$\frac{12}{40}$	$\frac{12}{40}$	30%
[30,40)	35	8	20	$\frac{8}{40}$	$\frac{20}{40}$	20%
[40,50)	45	12	32	$\frac{12}{40}$	$\frac{32}{40}$	30%
[50,60)	55	6	38	$\frac{6}{40}$	$\frac{38}{40}$	15%
[60,70)	65	2	40	$\frac{2}{40}$	$\frac{40}{40}$	5%

Para la frecuencia absoluta contamos los valores que están comprendidos en el intervalo correspondiente y los escribimos. Y para el resto de frecuencias procedemos como ya sabemos. En estos momentos la marca de clase no nos ha hecho falta pero sí para los próximos apartados.

3. Gráficos estadísticos

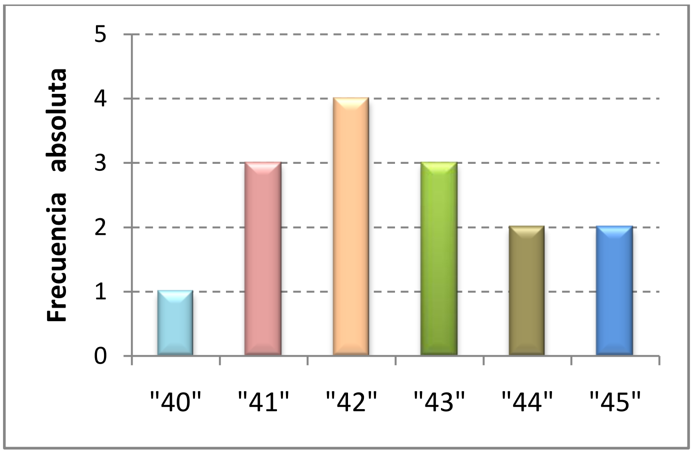
Una vez que tenemos todos los datos y hemos realizado la tabla de frecuencias podemos representar gráficamente estos datos. El hacerlo de forma visual hace que con un simple "golpe de ojo" podamos ver rápidamente información sobre el estudio estadístico.

Existen varios tipos de gráficos y usaremos uno u otro dependiendo del número de datos, del tipo de variable, etc.

- ***Diagrama de barras***

Puede aplicarse a cualquier tipo de variable, aunque se considera el idóneo para variables discretas. Cada valor se corresponde con una barra de longitud proporcional a su frecuencia. En el eje horizontal colocaremos los valores de la variable y en el eje vertical, las frecuencias.

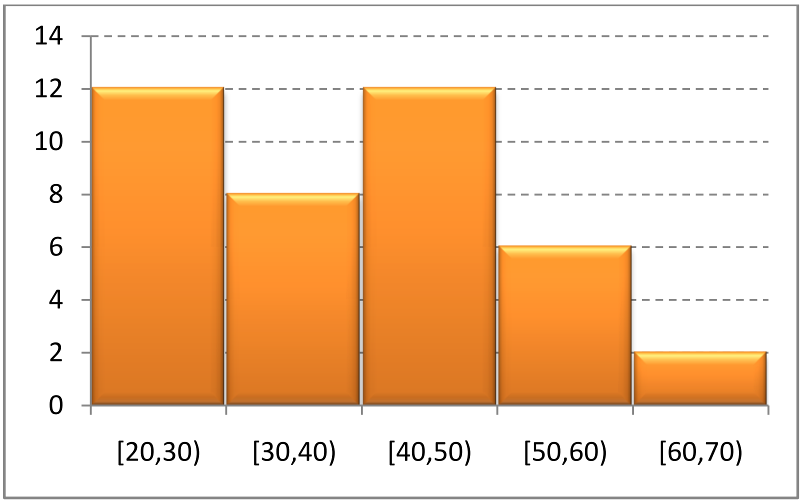
Vamos a hacer un diagrama de barras con el ejemplo del número de calzado de los trabajadores de la pequeña empresa.



- ***Histograma***

Se suele usar cuando los datos vienen agrupados en intervalos. Es muy parecido al diagrama de barras excepto que en este caso no hay separación entre las barras. En el eje horizontal colocaremos los valores de la variable y en el eje vertical, las frecuencias.

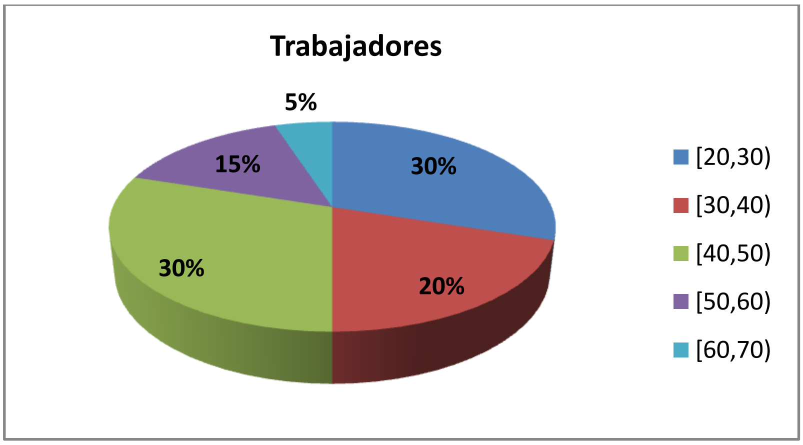
Vamos a hacer un histograma con el ejemplo de edades de los trabajadores de la mediana empresa.



- ***Diagrama de sectores***

Puede aplicarse a cualquier tipo de variable, aunque es el más adecuado cuando la frecuencia viene dada en porcentaje. Es un círculo dividido en sectores cuyo ángulo es proporcional a la frecuencia de cada valor. La amplitud de cada sector se obtiene multiplicando la frecuencia relativa por 360º.

Vamos a hacer un diagrama de sectores con el ejemplo de la edad de los trabajadores de la mediana empresa porque ya tenemos hecha la tabla de frecuencias con las frecuencias relativa y porcentual.



Ampliación: *Para realizar cualquiera de los gráficos propuestos puedes utilizar un programa de hoja de datos como Excel. Con un poco de práctica podrás realizarlos sin problemas.*

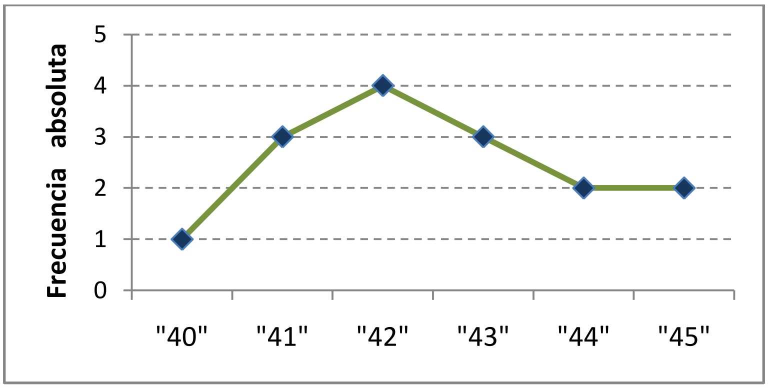
Ejemplo: tomamos la frecuencia relativa del ejemplo de la edad de los trabajadores de la mediana empresa.

Clase	*F. relativa n_i*	*Grados del sector*
[20,30)	$\frac{12}{40}$	$\frac{12}{40} \bullet 360 = 108º$
[30,40)	$\frac{8}{40}$	$\frac{8}{40} \bullet 360 = 72º$
[40,50)	$\frac{12}{40}$	$\frac{12}{40} \bullet 360 = 108º$
[50,60)	$\frac{6}{40}$	$\frac{12}{40} \bullet 360 = 54º$
[60,70)	$\frac{2}{40}$	$\frac{12}{40} \bullet 360 = 18º$

Para calcular el ángulo del sector proporcional a las frecuencias procedemos de la siguiente manera: Multiplicamos la frecuencia relativa por 360º y ese será el ángulo del sector circular.
Después con un transportador dibujaremos cada sector en el papel.

- ***Diagrama lineal o polígono de frecuencia***

Puede aplicarse a cualquier tipo de variables. En el eje horizontal colocaremos los valores de la variable y en el eje vertical, las frecuencias. Tras marcas los puntos correspondientes, procederemos a unirlos con una línea. Hagamos un diagrama lineal con el ejemplo del número de calzado de los trabajadores de la pequeña empresa.



Practica:

3. Resuelve estas dos cuestiones:

a) Las puntuaciones obtenidas por un grupo en una prueba han sido:
15, 20, 15, 18, 17, 13, 13, 16, 15, 19, 18, 15, 16, 20, 16, 15, 18, 16, 14, 13.
Realiza la tabla de frecuencias y un diagrama lineal como gráfico.

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

Peso	**[50, 60)**	**[60, 70)**	**[70, 80)**	**[80,90)**	**[90, 100)**	**[100, 110)**	**[110, 120)**
f_i	8	10	16	14	10	5	2

Realiza la tabla de frecuencias y un histograma como gráfico.

4. Parámetros estadísticos

Los **parámetros estadísticos** son diferentes conceptos que nos van a ayudar a comprender, mejor aún, lo concerniente a nuestro estudio estadístico. El cálculo de parámetros estadísticos se restringe a las variables estadísticas cuantitativas. No podemos calcular parámetros de variables cualitativas, aunque si podemos hacer sus tablas de frecuencias y representarlas gráficamente.

4.1. Medidas de centralización

Las medidas de centralización indican los valores más representativos de un conjunto de datos y son: la moda, la media y la mediana.

- ***Moda***

La moda es el valor de la variable que más se repite. Si nos fijamos en un gráfico de barras o histograma por ejemplo, será el valor que tenga la barra más alta. Se representa por $\boldsymbol{Mo}$

Ejemplo: en nuestro caso de los números de zapatos de los empleados de la pequeña empresa, la moda es: $\boldsymbol{Mo = 42}$.

- ***Media***

La media es el promedio de todos los valores de nuestro estudio. Para realizarla, sencillamente debemos sumar todos los datos y dividirlo por el número de éstos. Se representa por $\boldsymbol{\overline{x}}$**.**

Ejemplo: en nuestro caso de los números de zapatos de los empleados de la pequeña empresa, la media sería:

$$\overline{x} = \frac{40+41+41+41+42+42+42+42+43+43+43+44+44+45+45}{15} = \mathbf{42,5}$$

Por tanto, la media del número de calzado es **42,5**

También podemos hacerla de esta forma.

Valores	F. absoluta f_i	
40	**1**	$40 \bullet 1 = 40$
41	**3**	$41 \bullet 3 = 123$
42	**4**	$42 \bullet 4 = 168$
43	**3**	$43 \bullet 3 = 129$
44	**2**	$44 \bullet 2 = 88$
45	**2**	$45 \bullet 2 = 90$
$\bar{x}$	$\bar{x} = \frac{40+123+168+129+88+90}{15} = \mathbf{42,5}$	

Una vez que tenemos hecha la tabla de frecuencias, multiplicamos cada uno de los valores por el número de veces que se repite, esto es, la frecuencia absoluta (f_i) y dividimos el resultado por el número total de datos.

A continuación, realizamos la media de un estudio estadístico en el que la variable es cuantitativa continua.

Ejemplo: cojamos ahora el caso de las edades de los empleados de la mediana empresa, la media sería:

Clase	Marca de clase	F. absoluta f_i	
[20,30)	25	12	$25 \bullet 12 = 300$
[30,40)	35	8	$35 \bullet 8 = 280$
[40,50)	45	12	$45 \bullet 12 = 540$
[50,60)	55	6	$55 \bullet 6 = 330$
[60,70)	65	2	$65 \bullet 2 = 130$
$\bar{x}$	$\bar{x} = \frac{300+280+540+330+130}{40} = \mathbf{39,5}$		

Nos fijamos en la primera clase. Tenemos 12 trabajadores comprendidos entre los 20 y 30 años. Al estar agrupados no podemos utilizar uno por uno su edad, así que tomamos la marca de clase (25) y la multiplicamos por la frecuencia absoluta ($f_i = \mathbf{12}$).
Repetimos el proceso para cada una de las clases, sumamos y dividimos por el número total de datos (40).

- ***Mediana***

La mediana es el valor de la variable que ocupa la posición central de los datos. Se representa por $\boldsymbol{Me}$. Para realizarla colocaremos todos los datos ordenados y el valor central será la mediana. En caso que hubiera dos datos que ocupen la posición central, se halla la media aritmética de ambos datos y ése será la mediana.

Ejemplo: en nuestro caso de los números de zapatos de los empleados de la pequeña empresa, la mediana sería:

Ordenamos los datos:

$$40, 41, 41, 41, 42, 42, 42, 42, 43, 43, 43, 44, 44, 45, 45$$

Por tanto la mediana es: $\boldsymbol{Me} = \mathbf{42}$ que es el valor central.

Practica:

4. **Calcula la moda, la media y la mediana de este estudio estadístico:**

   Las notas de una determinada clase han sido las siguientes:
   7, 4, 1, 5, 6, 8, 2, 5, 5, 9, 6, 10, 6, 5, 7, 4, 6, 5, 7, 5, 4, 5, 9, 7, 2, 3

4.2. Medidas de dispersión

Las medidas de dispersión nos informan sobre el grado de dispersión de los datos respecto del valor central, esto es, si los datos se acercan entre sí o están muy separados. Como ves, éstas son muy diferentes de las mediadas de centralización que nos informaban sobre los valores centrales de la distribución estadística.

Las medidas de dispersión más usadas son: el recorrido, la desviación media, la varianza y la desviación típica.

- ***Recorrido***

Es la diferencia entre el valor máximo y el mínimo de la variable. Se le llama también rango. Se representa por $\boldsymbol{R}$

Ejemplo: seguimos utilizando el caso de los números de zapatos de los empleados de la pequeña empresa, el recorrido sería:

Valor mayor: 45; valor menor: 40

$$\boldsymbol{R = 45 - 40 = 5}$$

- ***Desviación media***

Es la media aritmética de las desviaciones de todos los datos respecto a la media. Se representa por signo $\boldsymbol{D_{\bar{x}}}$ **.** Para calcularla, primero hacemos la media y después hallamos la diferencia de cada dato con la media en valor absoluto. Por último, sumamos todos los valores absolutos de las diferencias y lo dividimos por el número de datos.

Ejemplo: continuamos utilizando el caso de los números de zapatos de los empleados de la pequeña empresa, la desviación media sería:

Valores	f_i	$\overline{x}$	Desviación entre el valor y la media	Desviación en valor absoluto	$D_{\overline{x}}$
40	1	$42,5$	$40-42{,}5=-2{,}5$	$\lvert -2{,}5\rvert = 2{,}5$	Para el valor 40 hay 1 dato, multiplicamos $2{,}5 \bullet 1 = 2{,}5$
41	3		$41-42{,}5=-1{,}5$	$\lvert -1{,}5\rvert = 1{,}5$	Para el valor 41 hay 3 datos, multiplicamos $1{,}5 \bullet 3 = 4{,}5$.
42	4		$42-42{,}5=-0{,}5$	$\lvert -0{,}5\rvert = 0{,}5$	Para el 42 hay 4 datos, multiplicamos $0{,}5 \bullet 4 = 2{,}0$.
43	3		$43-42{,}5=0{,}5$	$\lvert 0{,}5\rvert = 0{,}5$	Para el 43 hay 3 datos, multiplicamos $0{,}5 \bullet 3 = 1{,}5$.
44	2		$44-42{,}5=1{,}5$	$\lvert 1{,}5\rvert = 1{,}5$	Y así poco a poco con el resto
45	2		$45-42{,}5=2{,}5$	$\lvert 2{,}5\rvert = 2{,}5$	de valores.
	$D_{\overline{x}} = \frac{2{,}5+4{,}5+2{,}0+1{,}5+3{,}0+5{,}0}{15} = \mathbf{1,23}$				

Recordatorio: *El valor absoluto de un número es el mismo número sin el signo de + o - .*

- ***Varianza***

Es la media aritmética del cuadrado de las desviaciones respecto a la media de una distribución estadística. La varianza se representa por $\boldsymbol{V}$ **.**

Ejemplo: seguimos con el caso de los números de zapatos de los empleados de la pequeña empresa, la varianza sería:

Valores	f_i	Desviación de cada valor	$Desviación^2$	V
40	1	$\lvert -2{,}5\rvert = 2{,}5$	$2{,}5^2 = 6{,}25$	Como del valor 40 tenemos un dato, multiplicamos $6{,}25 \bullet 1 = 6{,}25$.
41	3	$\lvert -1{,}5\rvert = 1{,}5$	$1{,}5^2 = 2{,}25$	Para el valor 41 hay tres datos, multiplicamos $2{,}25 \bullet 3 = 6{,}75$.
42	4	$\lvert -0{,}5\rvert = 0{,}5$	$0{,}5^2 = 0{,}25$	Para el valor 42 hay cuatro datos, multiplicamos $0{,}25 \bullet 4 = 1{,}0$.
43	3	$\lvert 0{,}5\rvert = 0{,}5$	$0{,}5^2 = 0{,}25$	Para el valor 43 hay tres datos, multiplicamos $0{,}25 \bullet 3 = 0{,}75$.
44	2	$\lvert 1{,}5\rvert = 1{,}5$	$1{,}5^2 = 2{,}25$	Para el valor 44 hay dos datos, multiplicamos $2{,}25 \bullet 2 = 4{,}50$.
45	2	$\lvert 2{,}5\rvert = 2{,}5$	$2{,}5^2 = 6{,}25$	Para el valor 45 hay dos datos, multiplicamos $6{,}25 \bullet 2 = 12{,}50$.
	$V = \frac{6{,}25+6{,}75+1{,}0+0{,}75+4{,}50+12{,}50}{15} = \mathbf{2,11}$			

- ***Desviación típica***

La desviación típica es la raíz cuadrada de la varianza. Se representa por la letra $\boldsymbol{\sigma}$ **.**

Ejemplo: seguimos con el caso de los números de zapatos de los empleados de la pequeña empresa, la desviación típica sería:

$$\sigma = \sqrt{V} = \sqrt{2,11} = 1,45$$

Con la varianza y la desviación típica podremos saber si un estudio estadístico es más uniforme o disperso que otro, teniendo ambos la misma media aritmética.

Practica:

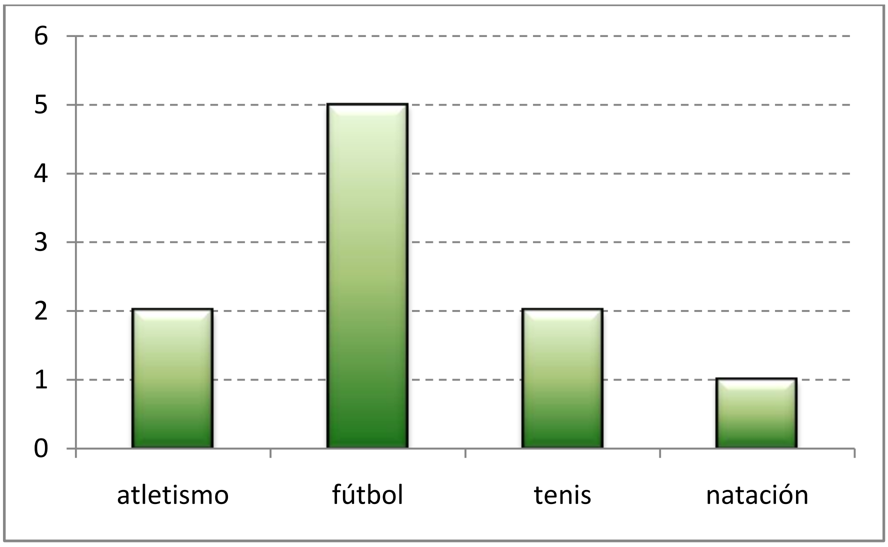
5. **Calcula el recorrido, la desviación media, la varianza y la desviación típica de este estudio estadístico:**

Las notas de una determinada clase han sido las siguientes:
7, 4, 1, 5, 6, 8, 2, 5, 5, 9, 6, 10, 6, 5, 7, 4, 6, 5, 7, 5, 4, 5, 9, 7, 2, 3

5. Interpretación de gráficos estadísticos

A la vista de un gráfico estadístico debemos de ser capaces de construir su tabla de frecuencias y sobre todo entenderlo e interpretarlo. Generalmente, en el eje horizontal estarán las variables y en el vertical, las frecuencias. El tamaño de las barras nos informará de la frecuencia con la que aparece cada valor. Es un proceso sencillo, así que vamos a ver un ejemplo.

Ejemplo: Tenemos un gráfico estadístico sobre los deportes que practican una serie de personas de un grupo.



Vamos a calcular la *tabla de frecuencias* fijándonos en la gráfica.

Valores	F. absoluta f_i	F. absoluta acumulada F_i	F. relativa h_i	F. relativa acumulada F_i	F. porcentual P_i
atletismo	2	2	$\frac{2}{10} = 0,2$	$0,2$	$0,2 \bullet 100 = 20\%$
fútbol	5	7	$\frac{5}{10} = 0,5$	$0,7$	$0,5 \bullet 100 = 50\%$
tenis	2	9	$\frac{2}{10} = 0,2$	$0,9$	$0,2 \bullet 100 = 20\%$
natación	1	10	$\frac{1}{10} = 0,1$	1	$0,1 \bullet 100 = 10\%$

¿A qué porcentaje de las personas no le gusta el ciclismo? Mirando la tabla de frecuencias encontramos respuesta a esta pregunta: **el 50%.**
Podrían hacernos más preguntas pero observando la gráfica y la tabla de frecuencias encontraríamos respuesta a todas ellas.

Glosario

- **Población:** es el conjunto de todos los elementos a los que se somete a un estudio estadístico.
- **Individuo:** es cada uno de los sujetos a los que se refiere el estudio estadístico.
- **Muestra:** es un subconjunto representativo de la población. Como ésta suele ser muy amplia, se toma una muestra representativa de la población.
- **Variable estadística:** es cada una de las propiedades o características que podemos estudiar de la población.
- **Frecuencia absoluta:** es el número de veces que se repite un valor determinado de un estudio estadístico. Se representa por f_i **.**
- **Frecuencia relativa:** es el cociente entre la frecuencia absoluta y el número total de datos obtenidos en nuestro estudio. Se representa por n_i .
- **Frecuencia porcentual:** es la frecuencia relativa expresada en %. Se representa por P_i **.** Nos indica el porcentaje de cada valor respecto del total..
- **Moda:** es el valor de la variable que más se repite. Se representa por Mo.
- **Media:** es el promedio de todos los valores de nuestro estudio. Para realizarla, sencillamente debemos sumar todos los datos y dividirlo por el número de éstos. Se representa por $\bar{x}$**.**
- **Mediana:** es el valor de la variable que ocupa la posición central de los datos. Se representa por Me **.**
- **Recorrido:** es la diferencia entre el valor máximo y el mínimo de la variable. Se le llama también rango. Se representa por R.
- **Desviación media:** es la media aritmética de las desviaciones de todos los datos respecto a la media. Se representa por signo $D_{\bar{x}}$ **.**
- **Varianza:** La varianza es la media aritmética del cuadrado de las desviaciones respecto a la media de una distribución estadística. La varianza se representa por V **.**
- **Desviación típica:** es la raíz cuadrada de la varianza. Se representa por la letra σ **.**

Actividades

Actividad 1:

El número de hermanos de los alumnos de una clase es el siguiente:

0 1 0 0 3 2 1 4 0 0 1 1 2 0 1 1 2 0 1 1 2 1 3 0 0 2 1 2 3 5

Elabora una tabla de frecuencias en las que se incluyan: frecuencia absoluta, absoluta acumulada, relativa, relativa acumulada y porcentual.

Actividad 2:

El número de goles metidos por partido por un equipo de fútbol por partido es el siguiente:

0 1 0 2 3 2 1 3 0 0 1 0 3 0 1 1 0 0 1 1 2 1 2 0 1 2 1 5 3 5

a) Haz una tabla de frecuencias en la que incluirás la absoluta, la relativa y la porcentual.
b) Elabora un gráfico de barras.

Actividad 3:

En una encuesta sobre vivienda se pregunta cuántas personas viven en la casa, obteniéndose las siguientes respuestas:

4 4 8 1 3 2 1 3 4 2 2 7 0 3 8 0 1 5 6 4 3 3 4 5 6 8 6 2 5 3 3 5 4 6 2 0 4 3 6 1

a) Elabora una tabla en la que se recojan todas las frecuencias.
b) ¿Cuántas viviendas fueron objeto de estudio? ¿En cuántas de ellas no vive nadie?
c) ¿Qué porcentaje de viviendas está ocupado por más de cinco personas?
d) Dibuja un diagrama de barras con las frecuencias absolutas acumuladas y un polígono de frecuencias absolutas.

Actividad 4:

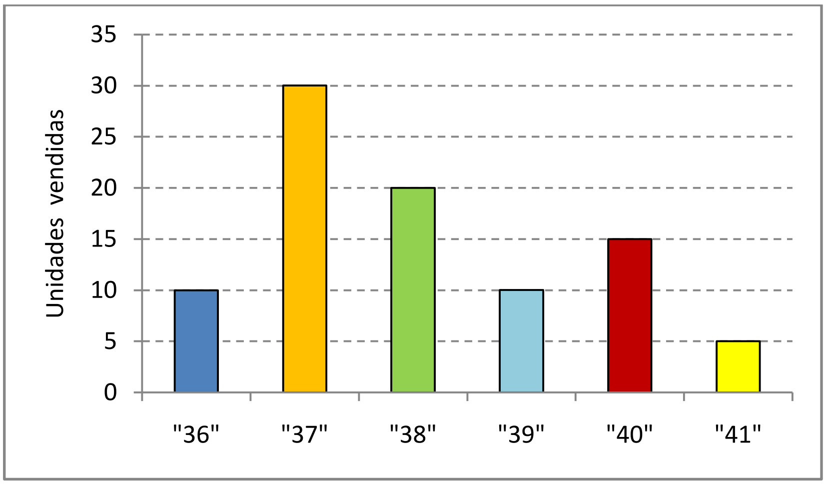
La siguiente tabla refleja las calificaciones de 30 alumnos en un examen de Matemáticas:

nota	2	4	5	6	7	8	9	10
Nº alumnos	2	5	8	7	2	3	2	1

a) ¿Cuántos alumnos aprobaron?
b) ¿Cuántos alumnos sacaron como máximo un 7?
c) ¿Cuántos sacaron como mínimo un 6?
d) Calcular la nota media, la moda y la mediana

Actividad 5:

La siguiente gráfica recoge la cantidad de parejas de zapatos de mujer vendidas en una tienda a lo largo del día:



a) ¿Cuántas parejas de zapatos del número 37 se han vendido?
b) Haz una tabla de frecuencias absoluta y absoluta acumulada.
c) ¿Cómo se llama la gráfica que nos han dado?
d) ¿Qué porcentaje de zapatos vendidos eran números del 39 ó 40?
e) Dibuja un polígono de frecuencias absolutas acumuladas.

Actividad 6:

En una gasolinera durante la mañana han repostado varios conductores y estos han sido los importes facturados:

23, 54, 8, 21, 42, 19, 8, 45, 36, 25, 11, 31, 30, 40, 20, 16, 5, 45, 55, 28, 15, 50, 28, 46, 37, 35, 45, 5, 19, 26, 36, 52, 46, 20, 35, 25, 45, 15, 10, 24, 38, 41, 40, 35, 26, 8, 52, 21, 37, 20.

a) ¿Cuál es la población del estudio?
b) ¿Cuál es la variable estadística? ¿Qué tipo de variable es?
c) Haz una tabla de frecuencias. NOTA: son muchos datos, así que conviene que los agrupes en intervalos.
d) Realiza un histograma.

Actividad 7:

Estos son los resultados de notas de los alumnos de una clase durante la primera y segunda evaluación:

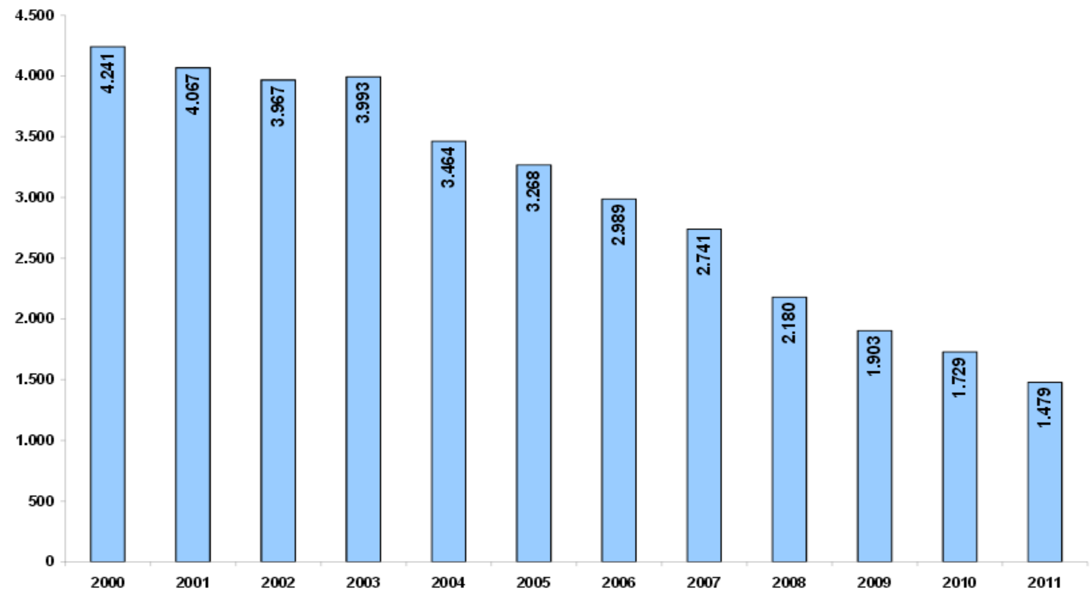
1ª evaluación	2ª evaluación
2, 5, 8, 5, 6, 4, 3, 8, 9, 5, 6, 5, 7, 7, 8	8, 1, 10, 5, 2, 8, 9, 3, 1, 5, 9, 6, 7, 9, 2

En ambos casos:

a) Realiza una tabla de frecuencias con la absoluta, la relativa y la porcentual.
b) Calcula la media, la desviación media, la varianza y la desviación típica.
c) Haz un gráfico de barras.

Actividad 8:

Grafica de fallecidos .2000-2011.



En este gráfico de la Dirección General de Tráfico sobre el número de personas fallecías entre al año 2000 y 2011,

a) ¿Cuál es la media de fallecidos por año?
b) ¿Qué año se ha desviado más de la media en resultados positivos?

Actividad 9:

En un grupo de 30 mujeres hemos medido la estatura, en centímetros, de cada una de ellas, obteniendo los siguientes resultados:

160, 163, 165, 164, 162, 161, 164, 167, 168, 154, 166, 168, 169, 168, 175, 167, 159, 160, 163, 164, 167, 164, 165, 164, 150, 166, 147, 170, 171, 169

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

Soluciones a los practica

Practica 1

a) *Cualitativas: marca de coche, colores preferidos y signo del zodíaco.*
Cuantitativas: temperatura mínima del día, número de hijos, número del zapato y nota del examen.

b) *Cuantitativas discretas: número de hijos y número del zapato.*
Cuantitativas continuas: temperatura mínima del día y nota del examen.

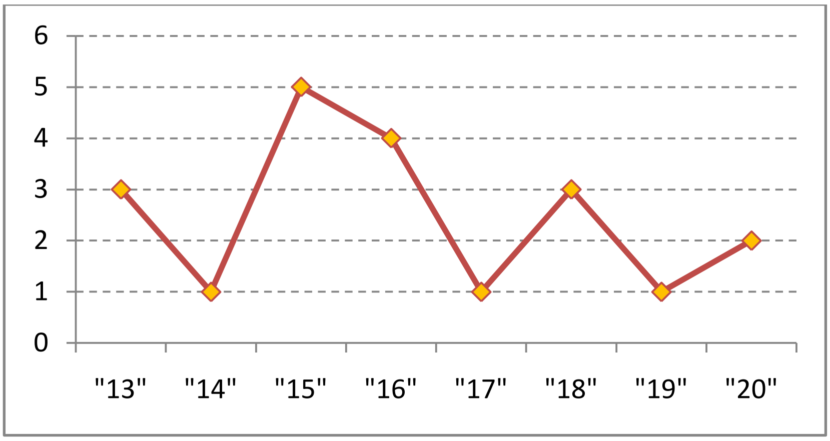
Practica 2

Valores	F. absoluta f_i	F. absoluta acumulada F_i	F. relativa n_i	F. relativa acumulada N_i	F. porcentual P_i
1	**4**	**4**	$\mathbf{\frac{4}{25}}$	$\mathbf{\frac{4}{25}}$	$\frac{4}{25} \bullet 100 = \mathbf{16}\%$
2	**4**	**8**	$\mathbf{\frac{4}{25}}$	$\mathbf{\frac{8}{25}}$	$\frac{4}{25} \bullet 100 = \mathbf{16}\%$
3	**5**	**13**	$\mathbf{\frac{5}{25}}$	$\mathbf{\frac{13}{25}}$	$\frac{5}{25} \bullet 100 = \mathbf{20}\%$
4	**3**	**16**	$\mathbf{\frac{3}{25}}$	$\mathbf{\frac{16}{25}}$	$\frac{3}{25} \bullet 100 = \mathbf{12}\%$
5	**6**	**22**	$\mathbf{\frac{6}{25}}$	$\mathbf{\frac{22}{25}}$	$\frac{6}{25} \bullet 100 = \mathbf{24}\%$
6	**3**	**25**	$\mathbf{\frac{3}{25}}$	$\mathbf{\frac{25}{25} = 1}$	$\frac{3}{25} \bullet 100 = \mathbf{12}\%$

Practica 3

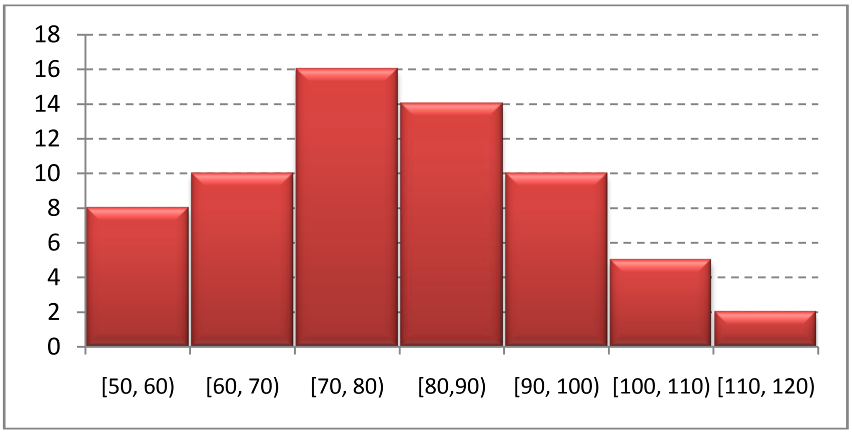
a)

Valores	F. absoluta f_i
13	**3**
14	**1**
15	**5**
16	**4**
17	**1**
18	**3**
19	**1**
20	**2**



b)

Peso	F. absoluta f_i
[50, 60)	**8**
[60, 70)	**10**
[70, 80)	**16**
[80,90)	**14**
[90, 100)	**10**
[100, 110)	**5**
[110, 120)	**2**



Practica 4

Valores	F. absoluta f_i	**Moda**	**Media**	
1	**1**		$1 \bullet 1 = 1$	Sumamos todas las cantidades y dividimos por el total de datos (26)
2	**2**	$\boldsymbol{Mo = 5}$	$2 \bullet 1 = 2$	
3	**1**	Porque es el valor que más se repite	$3 \bullet 1 = 3$	
4	**3**		$4 \bullet 3 = 12$	
5	**7**		$5 \bullet 7 = 35$	$\boldsymbol{\bar{x} = \frac{141}{26} = 5,42}$
6	**4**		$6 \bullet 4 = 24$	
7	**4**		$7 \bullet 4 = 28$	
8	**1**		$8 \bullet 1 = 8$	
9	**2**		$9 \bullet 2 = 18$	
10	**1**		$10 \bullet 1 = 10$	

Mediana

Ordenamos los valores:

1, 2, 2, 3, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 7, 8, 9, 9, 10

5

En este caso los valores centrales son 2. Hacemos la media aritmética y nos da ***5***

$$\boldsymbol{Me = 5}$$

Practica 5

Recorrido $\boldsymbol{R = 10 - 1 = 9}$

Hallamos la ***desviación media***

Valores	$\boldsymbol{f_i}$	$\boldsymbol{\bar{x}}$	Desviación entre el valor y la media	Desviación en valor absoluto	$\boldsymbol{D_{\bar{x}}}$
1	**1**		$1 - 5{,}42 = -4{,}42$	$\lvert -4{,}42 \rvert = 4{,}42$	Para el valor 1 hay 1 dato, $4{,}42 \bullet 1 = 4{,}42$
2	**2**		$2 - 5{,}42 = -3{,}42$	$\lvert -3{,}42 \rvert = 3{,}42$	Para el valor 2 hay 2 datos, $3{,}42 \bullet 2 = 12{,}84$.
3	**1**	$\boldsymbol{5,42}$	$3 - 5{,}42 = -2{,}42$	$\lvert -2{,}42 \rvert = 2{,}42$	Para el 3 hay 1 dato $2{,}42 \bullet 1 = 2{,}42$
4	**3**		$4 - 5{,}42 = -1{,}42$	$\lvert -1{,}42 \rvert = 1{,}42$	Para el 4 hay 3 datos, $1{,}42 \bullet 3 = 4{,}26$
5	**7**		$5 - 5{,}42 = -0{,}42$	$\lvert -0{,}42 \rvert = 0{,}42$	Para el 5 hay 7 datos, $0{,}42 \bullet 7 = 2{,}94$

6	***4***	$6-5{,}42=0{,}58$	$\lvert 0{,}58\rvert=0{,}58$	*Para el 6 hay 4 datos,* $0{,}58 \bullet 6=3{,}48$
7	***4***	$7-5{,}42=1{,}58$	$\lvert 1{,}58\rvert=1{,}58$	*Para el 7 hay 4 datos,* $1{,}58 \bullet 4=6{,}32$
8	***1***	$8-5{,}42=2{,}58$	$\lvert 2{,}58\rvert=2{,}58$	*Para el 8 hay 1 dato,* $2{,}58 \bullet 1=2{,}58$
9	***2***	$9-5{,}42=3{,}58$	$\lvert 3{,}58\rvert=3{,}58$	*Para el 9 hay 2 datos,* $3{,}58 \bullet 2=7{,}16$
10	***1***	$10-5{,}42=4{,}58$	$\lvert 4{,}58\rvert=4{,}58$	*Para el 10 hay 1 dato,* $4{,}58 \bullet 1=4{,}58$

$$\boldsymbol{D_{\overline{x}}}=\frac{4{,}42+12{,}84+2{,}42+4{,}26+2{,}94+3{,}48+6{,}32+2{,}58+7{,}16+4{,}58}{26}=\mathbf{1,96}$$

Ahora calcularemos la ***varianza***

Valores	$\boldsymbol{f_i}$	*Desviación de cada valor*	$Desviación^2$	***V***
1	***1***	$\lvert -4{,}42\rvert=4{,}42$	$4{,}42^2=19{,}53$	*Como del valor 1 tenemos un dato,* $19{,}53 \bullet 1=19{,}53$.
2	***2***	$\lvert -3{,}42\rvert=3{,}42$	$3{,}42^2=11{,}69$	*Para el valor 2 hay dos datos,* $11{,}69 \bullet 2=23{,}38.$
3	***1***	$\lvert -2{,}42\rvert=2{,}42$	$2{,}42^2=5{,}85$	*Para el valor 3 hay un dato,* $5{,}85 \bullet 1=5{,}85.$
4	***3***	$\lvert -1{,}42\rvert=1{,}42$	$1{,}42^2=2{,}01$	*Para el valor 4 hay tres datos,* $2{,}01 \bullet 3=6{,}03$
5	***7***	$\lvert -0{,}42\rvert=0{,}42$	$0{,}42^2=0{,}17$	*Para el valor 5 hay siete datos,* $0,17 \bullet 7=1{,}19.$
6	***4***	$\lvert 0{,}58\rvert=0{,}58$	$0{,}58^2=0{,}33$	*Para el valor 6 hay cuatro datos,* $0{,}33 \bullet 4=1{,}32.$
7	***4***	$\lvert 1{,}58\rvert=1{,}58$	$1{,}58^2=2{,}49$	*Para el valor 7 hay cuatro datos,* $2{,}49 \bullet 4=9{,}96.$
8	***1***	$\lvert 2{,}58\rvert=2{,}58$	$2{,}58^2=6{,}65$	*Para el valor 8 hay un dato,* $6{,}65 \bullet 1=6{,}65.$
9	***2***	$\lvert 3{,}58\rvert=3{,}58$	$3{,}58^2=12{,}81$	*Para el valor 9 hay dos datos,* $12{,}81 \bullet 2=25{,}62.$
10	***1***	$\lvert 4{,}58\rvert=4{,}58$	$4{,}58^2=20{,}97$	*Para el valor 10 hay un dato,* $20{,}97 \bullet 1=20{,}97.$

$$\boldsymbol{V}=\frac{19{,}53+23{,}38+5{,}85+6{,}03+1{,}19+1{,}32+9{,}96+6{,}65+25{,}62+20{,}97}{26}=4,\mathbf{63}$$

Y por último, la ***desviación típica***

$$\boldsymbol{\sigma}=\sqrt{\boldsymbol{V}}=\sqrt{\mathbf{4,63}}=\mathbf{2},\mathbf{15}$$

Bibliografía recomendada

- Gobierno de Aragón. *Matemáticas y Tecnología, módulo 4. Educación Secundaria para Personas Adultas*. España. Gobierno de Aragón. 2011. 126 p.

- Web: http://recursostic.educacion.es/bancoimagenes/web/ INTEF (Instituto Nacional de Tecnologías Educativas y Formación del Profesorado).

- Web: http://matematicasesomj.blogspot.com.es/p/segunda-evaluacion.html
- Web: http://www.vitutor.com
- Web: http://recursostic.educacion.es/secundaria/edad/index_mat.htm. INTEF-CICEAD.