



Estadística

Contenidos

1. Estadística descriptiva
 - Población y muestra
 - Variables estadísticas
 - Gráficos variables cualitativas
 - Gráficos variables cuantitativas discretas
 - Gráficos variables cuantitativas
2. Medidas de centralización
 - Media, moda y mediana
 - Evolución de la media
 - Evolución de la mediana
 - Media y mediana comparadas
 - Medidas de posición
3. Medidas de dispersión
 - Desviación típica y recorrido
 - Cálculo de las medidas de dispersión
 - La media y la desviación típica
4. Representatividad de las muestras
 - Muestreo estratificado
 - Muestreo aleatorio. Sesgo
5. Estadística bidimensional
 - Distribuciones bidimensionales
 - Correlación lineal
 - Rectas de regresión

Objetivos


- Distinguir los conceptos de población y muestra.
- Diferenciar los tres tipos de variables estadísticas.
- Hacer recuentos y gráficos.
- Calcular e interpretar las medidas estadísticas de centralización más importantes.
- Calcular las principales medidas de dispersión.
- Entender la importancia de la elección de la muestra para que sea representativa.
- Calcular el centro de gravedad, la covarianza, el coeficiente de correlación y la recta de regresión en una distribución bidimensional.

Antes de empezar

Un juego para empezar

Ve pulsando en piezas adosadas al hueco para desplazarlas y así durante un rato para deshacer el puzzle. Reconstrúyelo ahora.



Pulsa  para ir a la página siguiente.

1. Estadística descriptiva

1.a. Población y muestra.

Población es _____ sobre el que se hace un estudio estadístico.


La **muestra** es _____, de ahí que la propiedad más importante de las muestras es su _____.

El proceso seguido en la extracción de la muestra se llama _____.

En la escena adjunta tenemos 625 cuadraditos que representan a los alumnos de un instituto ficticio, si vas pulsando en los cuadraditos, vas seleccionando a parte de los alumnos.

Contesta:

- a. ¿Cuál es la población? _____
- b. ¿Cuál es la muestra? _____
- c. ¿Cómo se llama el proceso en el que se pregunta a toda la población? _____

Pulsa  para ir a la página siguiente.

1.b. Variables estadísticas.

La característica a estudiar en una población es la variable estadística.

Completa la siguiente tabla con las características de los distintos tipos de variables estadísticas:


Tipos de variables estadísticas		
Cualitativas		
Cuantitativas	Discretas	Continuas

En la escena de la derecha tienes ejemplos de cada tipo de variable estadística.

Pulsa en el botón  para hacer un ejercicio.

Completa la tabla con los ejemplos:

Cualitativas	Cuantitativas Discretas	Cuantitativas Continuas

Pulsa  para ir a la página siguiente.


1.c. Gráficos en variables cualitativas

El **diagrama de sectores** es el mas indicado para este tipo de información. El porcentaje de datos de cada valor en una muestra se corresponde con el mismo porcentaje de sector de un círculo.

$$\frac{\text{frecuencia}}{\text{total}} = \frac{\text{ángulo}}{360}$$

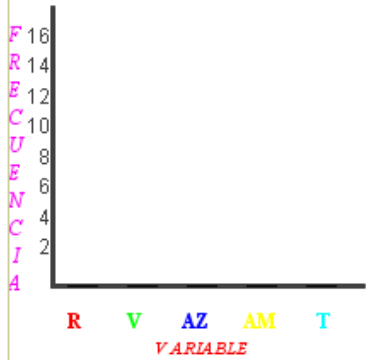
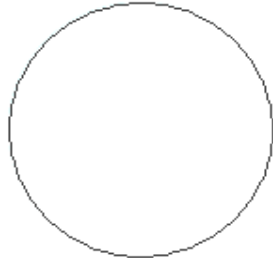
Así por ejemplo, si los datos son A, A, A, A, A, B, B, B, C, C, completa la tabla con los datos correspondientes:


x_i	Frecuencia	Porcentaje	Ángulo
A			
B			
C			

Haz clic en  para ver un vídeo sobre gráficos.

Con la ayuda de la escena de la derecha puedes hacer un ejercicio sobre representación gráfica de variables estadísticas cualitativas.

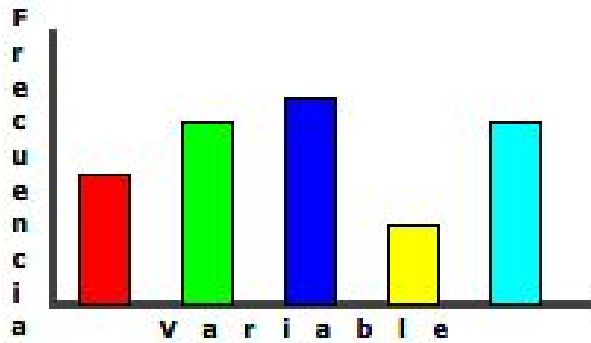
El ejercicio simula que tenemos una población de 30 alumnos y cada uno de ellos elige un color. Pulsando en **Genera** tendremos los 30 colores elegidos aleatoriamente, pulsa **ayuda** y lee como la escena te facilita el recuento y completa la tabla, comprobando que es correcto tú recuento. A continuación pulsa el botón de diagramas para ver los gráficos, y dibújalos en el lugar correspondiente

Color _i	Frecuencia	D. de columnas	D. de sectores
Rojo			<p>D. de sectores</p> 
Verde			
Azul			
Amarillo			
Turquesa			

Pulsa  para ir a la página siguiente.

1.d. Gráficos en variables cuantitativas discretas

Diagrama de barras. Bastará que observes ejemplos hechos de la escena de la derecha para comprender como se hacen y su significado. Este es el gráfico mas indicado para las variables cuantitativas discretas.



Puedes leer un **artículo** del Instituto Nacional de Estadística, sobre el comportamiento o actuaciones de nuestro país con el medio ambiente y la energías renovables, en él se muestran diversos tipos de diagramas.

Con la ayuda de la escena de la derecha puedes hacer unos ejercicios sobre representación gráfica de variables estadísticas cuantitativas discretas. En la tabla siguiente copia uno de ellos

El ejercicio simula que tenemos una población de 30 alumnos y cada uno de ellos nos dice el número de hermanos que tiene. Pulsando en **Genera** tendremos los 30 datos generados aleatoriamente, pulsa **ayuda** y lee como la escena te facilita el recuento y completa la tabla, comprobando que es correcto tú recuento. A continuación pulsa el botón de diagramas para ver los gráficos, y dibújalos en el lugar correspondiente

Variable	Frecuencia	D. de columnas	D. de sectores
0		<p>D. de columnas</p>	<p>D. de sectores</p>
1			
2			
3			
4			

Pulsa para ir a la página siguiente.

1.e. Gráficos en variables cuantitativas continuas

Histograma.

Lee la explicación de este tipo de gráfico estadístico.

Contesta.	RESPUESTA
¿Qué figura se utiliza para representar los datos?	
Si todos los intervalos son de la misma amplitud, ¿qué nos indica la altura?	
Si todos los intervalos <u>no</u> son de la misma amplitud, ¿qué magnitud es proporcional a la frecuencia?	

Pulsa en el enlace: **Ejemplo**. Fíjate en el ejemplo resuelto que aparece.

Polígono de frecuencias. Uniremos los centros de la parte superior de todos los rectángulos para obtenerlo.

También se suele dibujar el histograma de las **frecuencias acumuladas**, en cada dato se acumula la frecuencia de los datos anteriores.

Con la ayuda de la escena de la derecha puedes hacer unos ejercicios sobre representación gráfica de variables estadísticas cuantitativas continuas.

En la tabla siguiente copia uno de ellos:

El ejercicio simula que tenemos una población de 30 alumnos y medimos la altura de cada uno de ellos. Pulsando en **Pulsa para empezar** tendremos los 30 datos generados aleatoriamente, pulsa **ayuda** y lee como la escena te facilita el recuento y completa la tabla, comprobando que es correcto tú recuento. A continuación pulsa el botón de diagramas para ver los gráficos, y dibújalos en el lugar correspondiente

Intervalo	Frecuencia	Histograma	D. de frec. acumuladas
[150, 160)			
[160, 170)			
[170, 180)			
[180, 190)			
[190, 200)			

Pulsa para ir a la página siguiente.

EJERCICIOS

1. Clasifica los siguientes ejemplos de variables estadísticas: Longitud de un camión, Carga máxima, nº de ruedas, nº de ejes, tipo de camión, marcas de neumáticos, tipo de tapicería, nº de puertas, altura máxima.

Cualitativas:

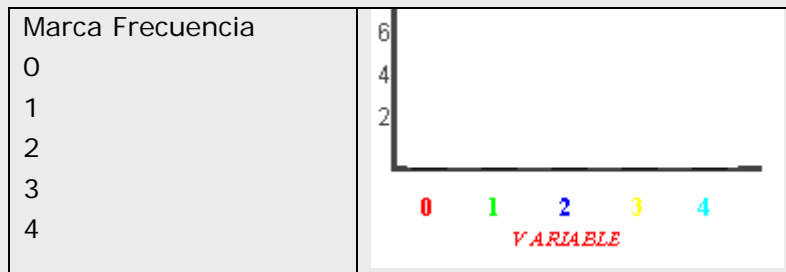
C. discretas:

C. continuas:

2. Calcula los grados que corresponden a cada valor en un gráfico de sectores hecho a partir de los datos: R, R, V, V, V, V, V, V, A, A, A

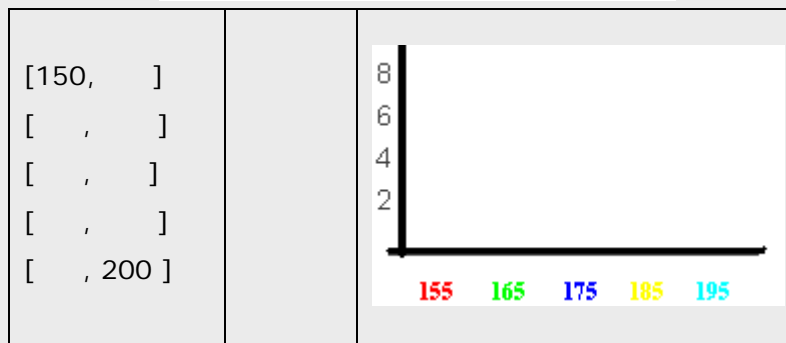
3. Agrupa los datos siguientes y haz un diagrama de barras adecuado.

Datos = { 0 1 0 2 3 4 1 2 2 1 2 2 3 4 3 2 1 3 }



4. Clasifica los datos en intervalos y dibuja un histograma adecuado.

180 197 154 181 189 162 152 162 167 190
 189 160 166 197 187 194 152 181 173 154
 177 184 186 174 177 159 158 189 160 150




Pulsa para ir a la página siguiente.

2. Medidas de centralización

2.a. Media, mediana y moda

Un conjunto N de observaciones, N números, puede que por si solo no nos diga nada. En cambio, si además nos dicen que están situados alrededor de uno o varios valores centrales ya tenemos una referencia que sintetiza la información. Por eso se definen los siguientes **parámetros de centralización** (porque nos indican el **centro** de la distribución)

Media: _____



Media= _____

Moda: _____

En el caso de variable continua, consideraremos por moda _____

También puede ocurrir que haya dos modas o que no haya ninguna que destaque.

Mediana: _____

En la escena de la derecha vemos ejemplos de cómo calcular estos parámetros. Copia a continuación uno de los ejemplos:

<i>Datos</i>	<i>Media</i>	<i>Moda</i>	<i>Mediana</i>

En el caso de la mediana, para pocos datos lo mejor es proceder según el ejemplo de la escena, según sea una cantidad par o impar.

Para calcular la mediana si la cantidad de observaciones es grande, habrá que agrupar los datos primero en una tabla. Y determinar segmentos de longitud proporcional a su frecuencia, disponerlos de forma lineal y marcar el centro como muestra el siguiente ejemplo.



Pulsa en el botón



para ver un ejercicio resuelto.

Pulsa



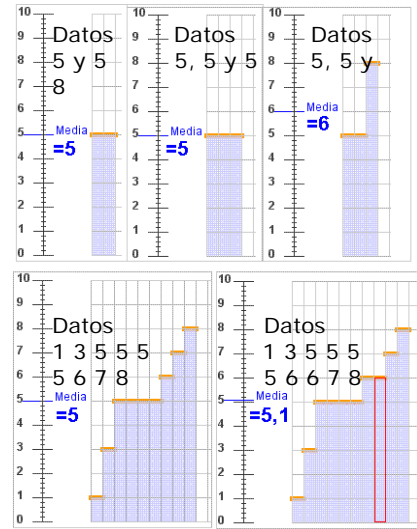
para ir a la página siguiente.

2.b Evolución de la media.

1 Para los datos 5 y 5 la media es ____.
 Si añadimos un 5 _____.
 Si añadimos un 8 _____.

2 Si tenemos 9 datos con media 5
 Necesitamos añadir un 6 para que la media pase a ser ____
 Si tenemos 19 datos con media 5
 Necesitamos un dato de valor 7 para que la media suba a ____

3 Para un conjunto de datos con media 5,
 si añadimos otro con media 5, por ejemplo 6 y 4,



En la escena de la derecha de la página puedes comprobar como se modifica la media en diversos ejemplos.

Elige el número del ejemplo:



A continuación puedes modificar el número de veces que aparece un dato pulsando las teclas y observa como varía en cada caso la media

Pulsa en el botón para hacer unos ejercicios.

En estos ejercicios tienes que calcular la media, puedes elegir si la variable es discreta o continua y ya te aparece hecho el recuento. Haz varios y a continuación copia un ejercicio de cada tipo en los recuadros siguientes:

Variable cuantitativa discreta			Variable cuantitativa continua		
Marca	Frecuencia		Marca	Frecuencia	
x_i	f_i	$x_i \cdot f_i$	x_i	f_i	$x_i \cdot f_i$
Total			Total		
Media \bar{x}			Media \bar{x}		

Pulsa para ir a la página siguiente.

2.c Evolución de la mediana

- 1 La mediana, para los datos 2, 3 y 4 es $Me = \underline{\hspace{2cm}}$.
Si cambiamos el 4 por 5 o por 6 o por cualquier otro valor mayor $\underline{\hspace{2cm}}$
- 2 Si añadimos otro dato y tenemos 2,3 4 y 4, por ejemplo, la $Me = \underline{\hspace{2cm}}$
Y si añadimos un quinto valor, un 4 o un 5 o un 6 o cualquier otro mayor que 4, la mediana en 2,3, 4, 4 y ?? pasa a ser $\underline{\hspace{2cm}}$
En cambio,.. Da igual que el valor ?? sea 5, 10 o 25.



En la escena de la derecha tienes ejemplos donde la mediana cambia y donde no. Además tu mismo puedes variar el valor o valores que quieras para observar como evoluciona. También tienes la posibilidad de realizar ejercicios de cálculo de esta, en la misma escena.

Pulsando en los botones **Número par de datos** y **Número impar de datos** obtienes ejemplos de datos y como calcular la mediana. Si pulsas **cambiar** puedes ver como calcular la mediana, elige el número del ejemplo: 1 2 3 4 5 6 7 8 9 10

A continuación puedes modificar el número de veces que aparece un dato pulsando las teclas - + y observar como varía en cada caso la mediana.

Pulsa en el botón para hacer unos ejercicios.

En estos ejercicios tienes que calcular la mediana. Puedes elegir si la variable es discreta o continua y ya te aparece hecho el recuento. Haz varios y a continuación copia un ejercicio de cada tipo. Puedes consultar la ayuda para resolverlos.

Variable cuantitativa discreta			Variable cuantitativa continua		
Marca	Frecuencia		Marca	Frecuencia	
x_i	f_i	F_i acumulada	x_i	f_i	F_i acumulada
<input style="width: 100%; height: 100%;" type="text"/>			<input style="width: 100%; height: 100%;" type="text"/>		

Pulsa para ir a la página siguiente.

2.d Media y mediana comparadas

Lee el texto y completa los valores de la media y la mediana en cada caso:

Datos	Media	Mediana
4, 6		
4, 6, 8		
4, 6, 11		
Los valores 8 y 11 se consideran observaciones _____.		
Si los datos estuvieran repartidos _____ respecto a un valor, ese valor sería _____.		
Si los valores a un lado de la mediana están más alejados de ella que los del otro lado, la media _____.		
_____ Hay una _____.		

Por ejemplo, si tenemos las observaciones

1. 20, 24 y 28.

$$Me = 24$$

2. Y para 20, 24, 28 y 30

$$Me = (24+28)/2 = 26$$

3. Para 20, 24, 28 y 100

$$Me = (24+28)/2 = 26$$

En cambio la media no se comporta de la misma forma para los mismos datos

1 $\bar{X} = 24$

2 $\bar{X} = 25,5$

3 $\bar{X} = 43$

Juega con la escena de la derecha.

Hay tres grupos de ejemplos, simétricos, asimétricos y atípicos.

Puedes observar la evolución de la mediana y la media

Elige el número del ejemplo:

Simétricos	Asimétricos	Atípicos
<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3	<input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> 6 <input type="radio"/> 7 <input type="radio"/> 8 <input type="radio"/> 9	<input type="radio"/> 10 <input type="radio"/> 11 <input type="radio"/> 12

Si quieres

puedes modificar el número de veces que aparece un dato pulsando las teclas



Pulsa para ir a la página siguiente.

2.e Medidas de posición: cuartiles y percentiles

Dado un conjunto de datos numéricos además de la mediana podemos considerar otras medidas de posición

- El primer valor que supera al 25% es el _____ Q_1
- El primer valor que supera al 75% es el _____ Q_3
- Para otros valores como el 10%, o el 80% hablamos de _____ P_{10} y P_{80} .

En la escena de la derecha tienes un ejemplo resuelto, si pulsas la flecha y pulsando en el botón **genera** puedes obtener muchos ejemplos resueltos, eligiendo si quieres que la variable sea discreta o continua.

Pulsa en el botón



para practicar el cálculo de las medidas de posición.

Pulsando en el botón **genera** obtienes nuevos datos, y en el botón **Discreta/Continua** intercambias el tipo de datos. Copia dos ejercicios en la tabla siguiente:

Variable cuantitativa discreta			Variable cuantitativa continua		
Marca	Frecuencia	Mediana	Marca	Frecuencia	Mediana
x_i	f_i		x_i	f_i	
		Cuartil Q_1			Cuartil Q_1
		Cuartil Q_3			Cuartil Q_3
		Percentil			Percentil
Total			Total		

EJERCICIOS

5. Calcula la media en cada caso:

- a) 4, 6, 8
- b) 4, 6, 8, 6
- c) 100, 120, 180, 200

6. Calcula la media en cada caso:

<p>a)</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th>Marca</th><th>Fr</th></tr> <tr><td>10</td><td>2</td></tr> <tr><td>20</td><td>4</td></tr> <tr><td>30</td><td>3</td></tr> <tr><td>40</td><td>2</td></tr> </table>	Marca	Fr	10	2	20	4	30	3	40	2	<p>b)</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th>Marca</th><th>Fr</th></tr> <tr><td>100</td><td>2</td></tr> <tr><td>200</td><td>4</td></tr> <tr><td>300</td><td>3</td></tr> <tr><td>400</td><td>2</td></tr> </table>	Marca	Fr	100	2	200	4	300	3	400	2
Marca	Fr																				
10	2																				
20	4																				
30	3																				
40	2																				
Marca	Fr																				
100	2																				
200	4																				
300	3																				
400	2																				

7. Determina la moda y la mediana

- a) 5,6,6
- b) 1,1,2,3
- c) 1,2,3,4,2
- d) 3,2,3,2,2,2

8. Calcula la moda y la mediana en cada caso:

<p>a)</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th>Marca</th><th>Fr</th></tr> <tr><td>10</td><td>2</td></tr> <tr><td>20</td><td>4</td></tr> <tr><td>30</td><td>3</td></tr> <tr><td>40</td><td>2</td></tr> </table>	Marca	Fr	10	2	20	4	30	3	40	2	<p>b)</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><th>Marca</th><th>Fr</th></tr> <tr><td>100</td><td>2</td></tr> <tr><td>200</td><td>3</td></tr> <tr><td>300</td><td>4</td></tr> <tr><td>400</td><td>1</td></tr> </table>	Marca	Fr	100	2	200	3	300	4	400	1
Marca	Fr																				
10	2																				
20	4																				
30	3																				
40	2																				
Marca	Fr																				
100	2																				
200	3																				
300	4																				
400	1																				

9. Calcula la mediana, cuartiles primero y 3º, y el percentil 30, 60 y 90 de los datos.

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

Pulsa para ir a la página siguiente.

3. Medidas de dispersión

3.a Varianza, Desviación típica y rango

“La estadística es una ciencia según la cual, si yo me como un pollo y tú no te comes ninguno, nos hemos comido como promedio medio pollo cada uno”.

La estadística indicará que todos comen lo mismo cuando las medidas de dispersión sean todas nulas.

Rango:

El intervalo definido por _____.

También se llama rango a _____.

Varianza:

La media aritmética de los _____

Si pulsas en el enlace [Fórmulas](#) se abre una ventana en la que puedes ver las dos fórmulas que nos permiten calcular la varianza y como son equivalentes entre sí.

Escribe en los cuadros esas dos fórmulas:

Desviación típica:

Cuanto mayores son la varianza o la desviación típica, los datos se separan más de la media, es decir, hay más dispersión.

Si pulsas en el enlace [Cálculo en distintos ejemplos](#) puedes generar ejemplos de variables discretas o continuas en los que verás dos métodos diferentes de cálculo de la varianza

¿Cuál es el método más manejable para el cálculo? _____

¿Por qué? _____

En la escena de la derecha tienes varios ejemplos de las medidas de dispersión y de su significado, léelos con atención.

Pulsa en el botón



para comparar distribuciones con iguales medidas de

centralización, en las que cambia la desviación típica. Copia a continuación dos de ellas

Pulsa



para ir a la página siguiente.

3.b Cálculo de las medidas de dispersión.

Recorrido

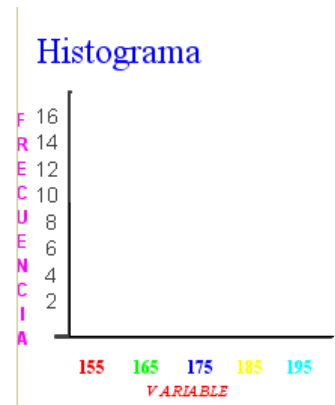
Pulsando en el botón **genera** obtienes nuevos datos, y en el botón **Discreta/Continua** intercambias el tipo de datos. Copia aquí dos ejercicios de cada tipo

Variable estadística discreta				Variable estadística continua			
Máximo		Máximo		Máximo		Máximo	
Mínimo		Mínimo		Mínimo		Mínimo	
Recorrido		Recorrido		Recorrido		Recorrido	

Desviación típica

En la escena de la derecha puedes generar unos datos, calcular la desviación típica y ver el diagrama de columnas. Copia a continuación dos ejercicios

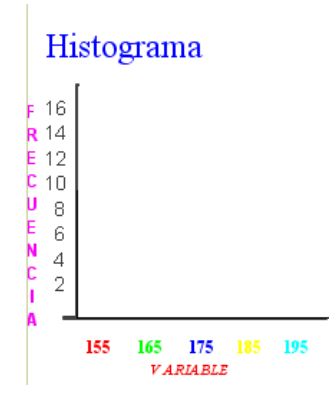
Intervalo	Marca	Frecuencia		
	x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$
Total				



Media Desviación típica


Mínimo Máximo Recorrido

Intervalo	Marca	Frecuencia		
	x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$
Total				



Media Desviación típica

Mínimo Máximo Recorrido

Pulsa en el botón  para hacer unos ejercicios.


Pulsando en el botón **Genera** obtienes nuevos datos, y en el botón **Discreta/Continua** intercambias el tipo de datos. Copia aquí dos ejercicios de cada tipo

Variable discreta					Variable discreta				
Marca	Frecuencia				Marca	Frecuencia			
x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	$f_i \cdot x_i^2$	x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	$f_i \cdot x_i^2$
Total					Total				

Media Desviación típica Media Desviación típica

Variable discreta					Variable discreta				
Marca	Frecuencia				Marca	Frecuencia			
x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	$f_i \cdot x_i^2$	x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	$f_i \cdot x_i^2$
Total					Total				

Media Desviación típica Media Desviación típica

Pulsa  para ir a la página siguiente.

3.c Media y desviación típica.

Para muestras unimodales (una sola moda) y casi simétricas, alrededor de la media podemos considerar un intervalo que contenga la mayoría de los datos.

Por ejemplo, para una muestra con media 100 y desviación típica 10, la mayor parte de los datos estarán entre 90 y 110, aproximadamente el 68%; entre 80 y 120 estará el 95% aproximadamente. Y casi todos entre 70 y 130.

Hay una forma de distribución de datos llamada normal que cumple con lo anterior, y de una manera u otra, de todas las poblaciones grandes se pueden extraer datos que se ajustan a ella. En cursos superiores verás la importancia de estas distribuciones.

En la escena de la derecha tienes unos ejemplos en donde aparece la media y unas franjas de color a su alrededor. Elige el número del ejemplo:

A continuación puedes modificar el número de veces que aparece un dato pulsando las teclas y observa como varía en cada caso la media y las franjas de su alrededor

Pulsa en el botón para hacer unos ejercicios .

Pulsando en el botón **genera** obtienes nuevos datos. A continuación haz en las tablas dos de ellos, y después comprueba el resultado en la escena

Marca	Frecuencia				
x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	$f_i \cdot x_i^2$	Media
					Desviación típica
					$[\bar{x} + \sigma, \bar{x} - \sigma] = [\quad , \quad]$
					Nº de datos
					$[\bar{x} + 2\sigma, \bar{x} - 2\sigma] = [\quad , \quad]$
					Nº de datos
Total					

Marca	Frecuencia				
x_i	f_i	$x_i \cdot f_i$	$f_i \cdot (\bar{x} - x_i)^2$	$f_i \cdot x_i^2$	Media
					Desviación típica
					$[\bar{x} + \sigma, \bar{x} - \sigma] = [\quad , \quad]$
					Nº de datos
					$[\bar{x} + 2\sigma, \bar{x} - 2\sigma] = [\quad , \quad]$
					Nº de datos
Total					

Pulsa para ir a la página siguiente.

4. Representatividad

4.a Representatividad. Muestreo estratificado

Una muestra es **representativa** de la población cuando _____

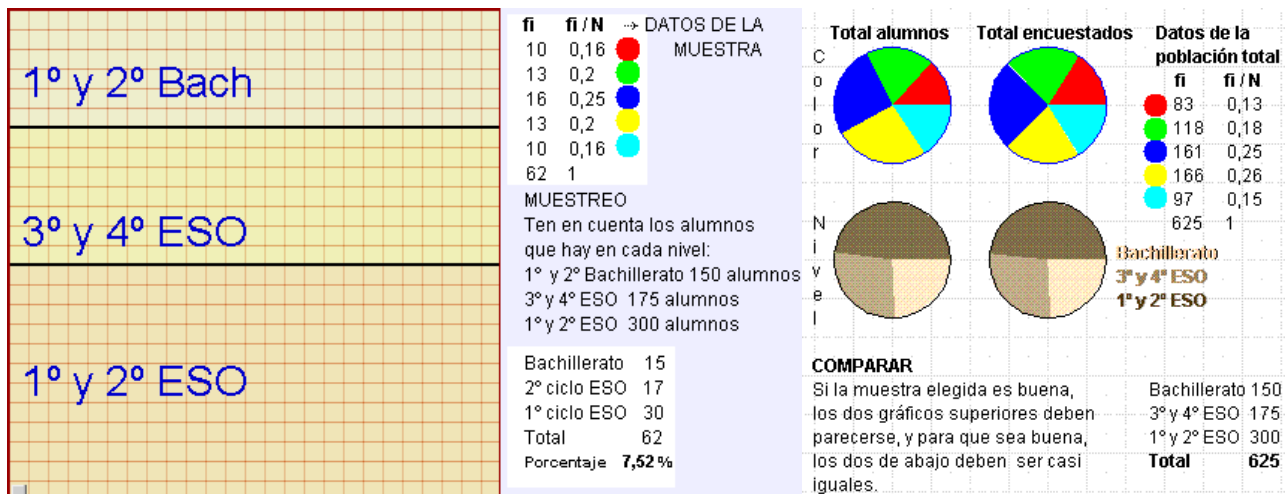
¿De qué depende que el estudio de una población sea o no representativo? _____

Por ejemplo, si queremos estudiar el poder adquisitivo de una población, y solo elegimos a individuos de una determinada zona, o principalmente de una determinada zona, ¿cómo será la muestra? _____

Si hay tres zonas con 12.000, 18.000 y 20.000 habitantes, escribe en que porcentaje debemos elegir a los individuos de cada zona para elaborar una muestra representativa


Un **muestreo estratificado** es _____

En la escena tienes 625 cuadros que representan a los alumnos de un instituto ficticio, siguiendo las instrucciones puedes observar la diferencia entre un muestreo representativo y otro que no lo es.



Si comparamos los gráficos en ambos ejemplos de muestra, ¿en que tipo de muestra se parecen más a los de la población total? _____

¿Por qué? _____

Pulsa  para ir a la página siguiente.

4.b Muestreo aleatorio. Sesgo

¿Cuándo se dice que la muestra está sesgada? _____

Explica en que consiste un muestreo aleatorio total: _____

En la escena puedes animar una elección totalmente aleatoria o realizar muestreos simulando encuestas al hacer clic.

xi fi fi / N → DATOS DE LA MUESTRA

0	13	0,2
1	20	0,32
2	9	0,14
3	10	0,16
4	10	0,16
62	1	

EN ESTE MUESTREO
No tienes que tener en cuenta los niveles, solo que cada alumno sea elegido de entre todos aleatoriamente. Aún así habrá correlación en los niveles entre muestra y población.

Bachillerato	13
2º ciclo ESO	12
1º ciclo ESO	37
Total	62
Porcentaje	7,84%

Total alumnos **Total encuestados** **Datos de la población total**

xi	fi	fi / N
0	130	0,2
1	135	0,21
2	113	0,18
3	112	0,17
4	135	0,21
625	1	

COMPARAR
Aunque no hayas tenido en cuenta los niveles para tu elección, además de comparar por nº hermanos puedes hacerlo por niveles.

Bachillerato	150
3º y 4º ESO	175
1º y 2º ESO	300
Total	625

Pulsa en el botón  para hacer un ejercicio sobre representatividad .

Copia en este cuaderno un ejercicio y compruébalo después en la escena

De una población queremos extraer una muestra de tamaño _____.

Si proceden de 5 áreas distintas, A, B, C, D y E con porcentajes del total de la población de _____%, _____%, _____%, _____% y _____%


¿A cuantos de cada zona hay que entrevistar?

A= _____, B= _____, C= _____, D= _____ y E= _____

EJERCICIO

13. Una gran empresa tiene trabajadores en cuatro áreas. Operarios, Representantes, administración y dirección. Las condiciones de trabajo son bastantes diferentes en cada área, por lo que el grado de satisfacción no es igual en cada una de ellas. Para averiguarlo, si hay 1000, 500, 300 y 200 trabajadores en las áreas de operarios, representantes, administrativos y directivos, ¿cuántos hay que seleccionar de cada área para una muestra de tamaño?

- a) 200
- b) 100
- c) 300

Pulsa  para ir a la página siguiente.

5. Estadística bidimensional

5.a Distribuciones bidimensionales

Lee en pantalla la explicación sobre este tipo de distribuciones y contesta:

¿Qué es una **distribución bidimensional**?

Las variables que intervienen, ¿tienen que estar relacionadas? _____

¿Qué es un diagrama de dispersión? _____

¿Con qué otro nombre se conoce al diagrama de dispersión? _____


¿Cuándo hay muchos datos y los pares de valores se repiten, acudimos a _____

_____.

¿Qué tipos de gráficos se utilizan en el caso de muchos datos o de pares repetidos? _____

_____.

En la escena de la derecha puedes ver los distintos tipos de gráficos. Pulsa en "Ver gráfico siguiente" para ir de un tipo a otro.

Pulsa  para ir a la página siguiente.

5.b Correlación lineal

Lee en pantalla la explicación sobre este concepto. En la escena de la derecha puedes mover los puntos para variar así la correlación y observar cómo varía su valor y como va cambiando la nube de puntos. Contesta las siguientes preguntas después de leer y comprender las explicaciones:

¿Cuál es el objetivo de un **estudio bidimensional**?

¿Qué es la correlación? _____

¿Cómo se aprecia la correlación? _____

¿En qué caso se dice que la correlación es lineal? _____

¿Cómo se sabe si la correlación lineal es fuerte o no? _____

Si la recta es creciente, se dice que la correlación lineal es _____

Si la recta es decreciente, se dice que la correlación lineal es _____

¿Cómo se llama el parámetro que se utiliza para cuantificar la correlación?

¿Qué valores puede tomar este coeficiente? _____

¿En qué casos la correlación es más fuerte? _____

Completa la siguiente tabla con los tipos de correlación en función de los valores de "r":

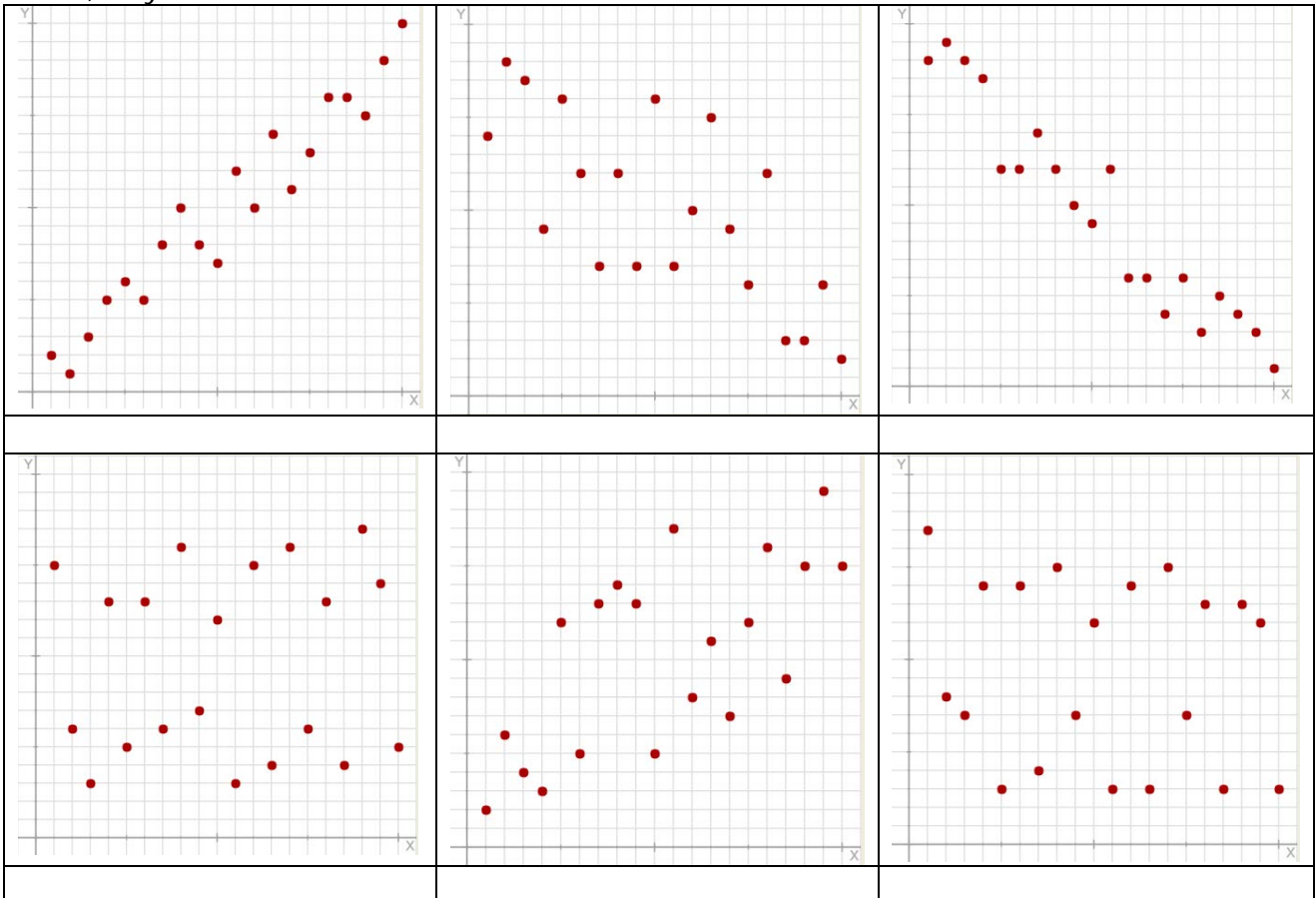
Si $r = 1$ ó $r = -1$	
Si $0,5 < r < 1$	
Si $-1 < r < 0,5$	
Si $r = 0$ ó próximo	


Pulsa en el botón  para hacer ejercicios.

Aparece una nube de puntos de una distribución bidimensional. Tienes que elegir la solución correcta entre las propuestas indicando cómo crees que es la correlación.

EJERCICIOS:


Indica de qué tipo de correlación es cada uno de las siguientes nubes de puntos, entre las opciones: Muy Fuerte y Directa, Muy Fuerte e Inversa, Fuerte y Directa, Fuerte e Inversa, Débil, Muy Débil Casi Nula.



Pulsa  para ir a la página siguiente.

5.c Rectas de regresión

Lee en pantalla la explicación sobre este concepto. En la escena de la derecha puedes ver una nube de puntos que podrás modificar si quieres arrastrando los puntos.

Una vez situados en dónde quieras, modifica el valor de la pendiente en el control: 

Cuando hayas encontrado su valor aparecerá dibujada una recta en color azul, habrás encontrado la **recta de regresión** que se ajusta a esa nube de puntos.

Si pulsas en **Ver cálculos** podrás observar cómo se ha calculado la ecuación de esta recta. Y en esa nueva escena, en la parte inferior puedes pulsar en **Recta de regresión de X sobre Y** para ver esta otra, que como observarás puede ser diferente de la anterior.

Contesta las siguientes preguntas:

¿Por qué punto pasa la Recta de regresión de Y sobre X? _____

¿Cuál es su ecuación? _____

¿Cuál es el valor de la pendiente de esa recta? _____

¿Qué nombre recibe ese la pendiente de esa recta? _____

¿Para qué sirve la recta de regresión? _____

¿En qué casos es más fiable el valor de la estimación?

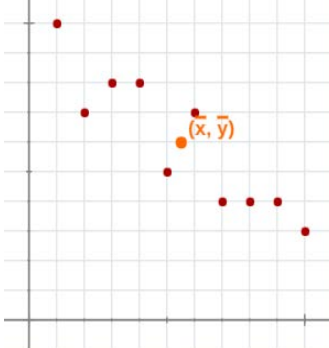
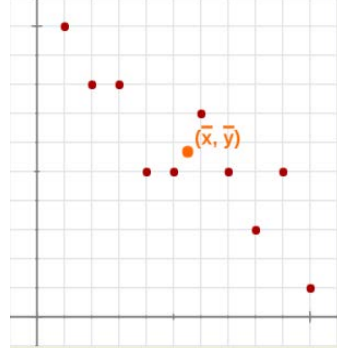
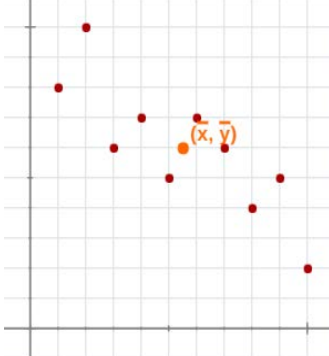
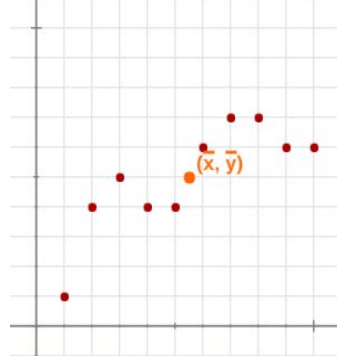
- _____
- _____


Pulsa en el botón  para hacer ejercicios.

Aparece una nube de puntos de una distribución bidimensional. Tienes que elegir la solución correcta entre las propuestas indicando cuál crees que es la ecuación de la recta de regresión de Y sobre X.

EJERCICIOS:

En cada una de las nubes de puntos siguientes indica cuál es la ecuación que corresponde a la recta de regresión de Y sobre X:

	<p>Centro de gravedad: (5,5 , 6)</p> <p><input type="checkbox"/> $y = -0,7x + 2,1$</p> <p><input type="checkbox"/> $y = 0,7x + 2,1$</p> <p><input type="checkbox"/> $y = -0,7x + 9,9$</p> <p><input type="checkbox"/> $y = 0,7x + 9,9$</p>		<p>Centro de gravedad: (5,5 , 5,7)</p> <p><input type="checkbox"/> $y = -0,8x + 1,3$</p> <p><input type="checkbox"/> $y = 0,8x + 1,3$</p> <p><input type="checkbox"/> $y = -0,8x + 10,1$</p> <p><input type="checkbox"/> $y = 0,8x + 10,1$</p>
	<p>Centro de gravedad: (5,5 , 6)</p> <p><input type="checkbox"/> $y = 0,6x + 2,7$</p> <p><input type="checkbox"/> $y = -0,6x + 9,3$</p> <p><input type="checkbox"/> $y = 0,6x + 9,3$</p> <p><input type="checkbox"/> $y = -0,6x + 2,7$</p>		<p>Centro de gravedad: (5,5 , 5)</p> <p><input type="checkbox"/> $y = 0,5x + 2,3$</p> <p><input type="checkbox"/> $y = -0,5x + 2,3$</p> <p><input type="checkbox"/> $y = 0,5x + 7,8$</p> <p><input type="checkbox"/> $y = -0,5x + 7,8$</p>

Pulsa  para ir a la página siguiente.



Recuerda lo más importante – RESUMEN

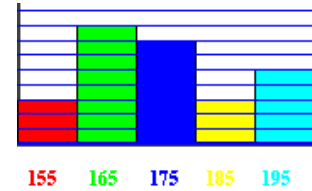
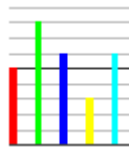
Población.

Muestra

Variables estadísticas

Tipos

Tipos de gráficos



Media, moda y desviación típica

Media

Moda

Desviación típica

$$\bar{X} =$$

$$M_o =$$

$$\sigma = \sqrt{\quad}$$

Cuartil, mediana, percentil

Cuartiles

Mediana

Percentiles

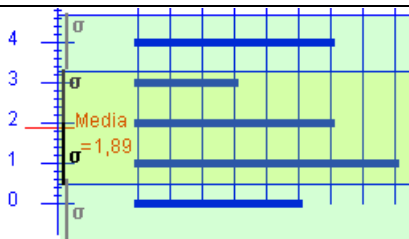
$$Q_1 =$$

$$M_e =$$

$$P_i =$$

$$Q_3 =$$

Media y desviación típica: Observa el ejemplo



$$[\bar{x} + \sigma, \bar{x} - \sigma] = [\quad , \quad]$$


% de datos

$$[\bar{x} + 2\sigma, \bar{x} - 2\sigma] = [\quad , \quad]$$

% de datos

Representatividad

Una muestra es representativa de la población cuando _____

Pulsa  para ir a la página siguiente.



Para practicar

Ahora vas a practicar resolviendo distintos EJERCICIOS. En las siguientes páginas encontrarás EJERCICIOS de:

Medidas de centralización y dispersión. Representatividad Interpretación de gráficos del INE

Completa el enunciado con los datos con los que te aparece cada EJERCICIO en la pantalla y después resuélvelo. Es importante que primero lo resuelvas tú y después compruebes en el ordenador si lo has hecho bien.



Medidas de centralización y dispersión. Representatividad.

1. Tipo de variable (haz dos ejercicios)

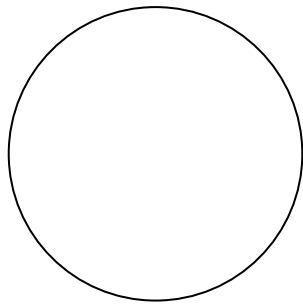
Clasifica las siguientes variables: <table border="1" style="width: 100%;"> <tr> <td> nº de hijos → flor preferida → peso → temperatura media → sabor → altura → </td> <td> Velocidad → Aceleración → nº de valvulas → nº de plazas → tipo de vehículo → nº de ruedas → carga neta → tipo de tapicería → </td> </tr> </table>	nº de hijos → flor preferida → peso → temperatura media → sabor → altura →	Velocidad → Aceleración → nº de valvulas → nº de plazas → tipo de vehículo → nº de ruedas → carga neta → tipo de tapicería →
nº de hijos → flor preferida → peso → temperatura media → sabor → altura →	Velocidad → Aceleración → nº de valvulas → nº de plazas → tipo de vehículo → nº de ruedas → carga neta → tipo de tapicería →	

Clasifica las siguientes variables estadísticas de un partido de fútbol: <table border="1" style="width: 100%;"> <tr> <td> nº de espectadores en el campo → </td> <td> jugador preferido → nº de goles → tiempo transcurrido → </td> </tr> </table>	nº de espectadores en el campo →	jugador preferido → nº de goles → tiempo transcurrido →
nº de espectadores en el campo →	jugador preferido → nº de goles → tiempo transcurrido →	

2. Recuento de datos (haz dos ejercicios)

Haz un recuento de los datos en una tabla.  <p>Nº de hermanos</p>	
Haz un recuento de los datos en una tabla.  <p>Largo de la tela</p>	

3. Diagrama de sectores

Haz un diagrama de sectores para los datos del color preferido de la tabla <table border="1" style="width: 100%;"> <thead> <tr> <th>Marca</th> <th>Frecuencia</th> </tr> </thead> <tbody> <tr> <td>x_i</td> <td>f_i</td> </tr> <tr> <td> </td> <td> </td> </tr> <tr> <td> </td> <td> </td> </tr> <tr> <td> </td> <td> </td> </tr> <tr> <td> </td> <td> </td> </tr> <tr> <td> </td> <td> </td> </tr> <tr> <td>Total</td> <td> </td> </tr> </tbody> </table>	Marca	Frecuencia	x_i	f_i											Total		
Marca	Frecuencia																
x_i	f_i																
Total																	

4. Diagrama de barras

Haz un diagrama de barras para los datos de la tabla.

Marca	Frecuencia
x_i	f_i
Total	

5. Histograma

Con los datos de la tabla haz un histograma

Intervalo	Marca	Frecuencia
	x_i	f_i
	Total	

6. Moda

¿Cuál es la moda en cada grupo?

A= {rojo, azul, verde, azul}

B= {blanco, negro, azul}

C= {rojo, verde, amarillo, rojo, azul, rojo, azul, azul}

A →

B →

C →

¿Cuál es la moda en cada grupo?

A= {1, 2, 3, 4, 1, 1, 2, 3}

B= {1, 1, 2, 2, 3, 4, 4}

C= {1, 2, 3, 3, 3, 7, 7, 7, 4}

A →

B →

C →

7. Mediana

¿Cuál es la mediana en cada caso?

A= {1, 2, 3, 4, 5}

B= {1, 2, 3, 4}

C= {1, 2, 3, 4, 5, 6}

D= {1, 2, 3, 3}

E= {1, 2, 3, 3, 3}

A →

B →

C →

D →

E →

¿Cuál es la mediana en cada caso?

A= {1, 2, 7, 10}

B= {3, 6, 7}

C= {1, 2, 3, 4, 5, 6, 1, 2, 1}

A →

B →

C →

8. Igual media

¿Cuál es la mediana en cada caso?

A= { , } ; B= { , } ; C= { , }

A →

B →

C →

9. Concepto de media

Calcula la media para los datos:

$x_1 =$ $f_1 =$

$x_2 =$ $f_2 =$

$x_3 =$ $f_3 =$

10. Cálculo de la media

Calcula la media: Distribución discreta			
Marca	Frecuencia		
x_i	f_i		
Total			
Calcula la media: Distribución continua			
Intervalo	Marca		Frecuencia
	x_i		f_i
	Total		

11. Caso simple de desviación típica

¿Cuál es la desviación típica en cada caso? $A = \{ \quad , \quad \} ; B = \{ \quad , \quad \} ; C = \{ \quad , \quad \}$	$A \rightarrow$	$B \rightarrow$	$C \rightarrow$
--	-----------------	-----------------	-----------------

12. Concepto de desviación típica

Calcula la desviación típica para los datos: $x_1 = \quad \quad f_1 = \quad$ $x_2 = \quad \quad f_2 = \quad$ $x_3 = \quad \quad f_3 = \quad$	
---	--

13. Cálculo de desviación típica

Calcula la desviación típica: Distribución discreta			
Marca	Frecuencia		
x_i	f_i		
Total			
Calcula la desviación típica: Distribución continua			
Intervalo	Marca		Frecuencia
	x_i		f_i
	Total		

14. Representatividad

Tomamos una muestra de tamaño 2000 de una población donde la edad influye en la característica del estudio. El ___ % de la población es mayor, el ___ % joven y el ___ % media. ¿A cuántos entrevistaré de cada grupo de edad?	$Jóvenes \rightarrow$ $Medios \rightarrow$ $Mayores \rightarrow$
--	--

Pulsa para ir a la página siguiente.

Interpretación de gráficos del INE.

(En cada apartado aparece una imagen y en el texto preguntas sobre ella. Pulsando en OTRO EJERCICIO aparecen más preguntas sobre la misma imagen)

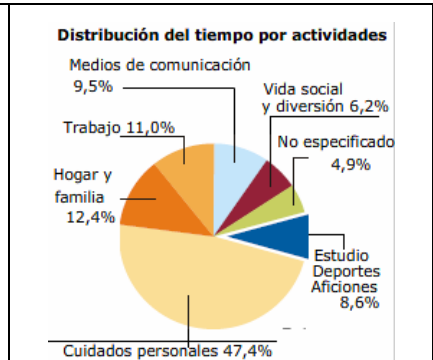
1. ¿Qué hacemos? Observa el gráfico de sectores del INE y responde a las preguntas:

¿Cuál es la variable estudiada?
¿y la frecuencia?

¿A qué grupo de actividades dedicamos más tiempo los españoles?
¿Cuál es la moda?

Calcula cuánto tiempo dedicamos al hogar y la familia:

¿Cuántos grados ocupa este sector en el diagrama?

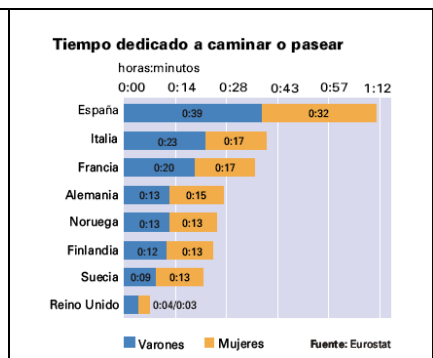


2. ¿Cuánto paseamos? En el gráfico es fácil ver que somos lo europeos que más paseamos.

¿En qué países pasean más las mujeres que los hombres?

Calcula el tiempo medio que se dedica en cada país a pasear.

¿Qué país está en el percentil 50?

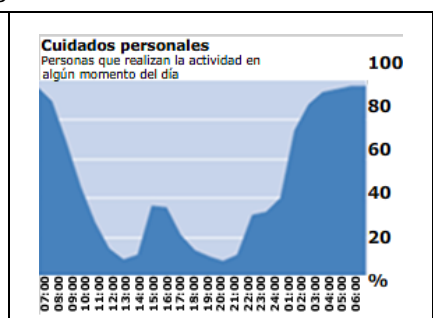


3. Cuidado personal. Observa el gráfico y responde a las preguntas:

¿Crees que el dormir se ha contado como actividad de cuidado personal?

A las 15:00 hay un máximo local en la gráfica ¿a qué se debe?

A la hora de la comida el 38% de las personas se dedica al cuidado personal. ¿Significa esto que un 62% de las personas no come?



4. Vida social. Observa el gráfico y responde a las preguntas:

¿Cuáles son las comunidades en las que se dedica menos tiempo a la vida social y a la diversión?

¿Cuánto tiempo dedican a la diversión o a la vida social la mayor parte de las comunidades?

¿Cuál es el tiempo medio que se dedica en España a esta actividad?



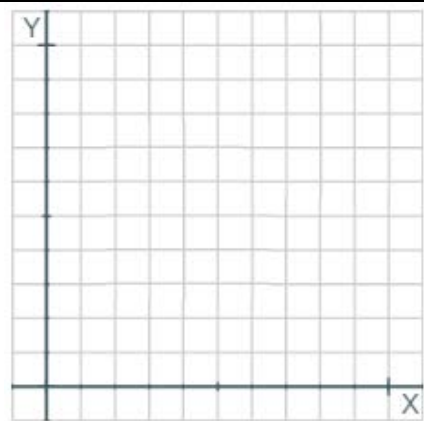
DISTRIBUCIONES BIDIMENSIONALES.

1. Notas

Las notas de 8 alumnos en lengua e Inglés fueron:

LENGUA (X)									
INGLÉS (Y)									

Dibuja la nube de puntos y calcula el coeficiente de correlación lineal:

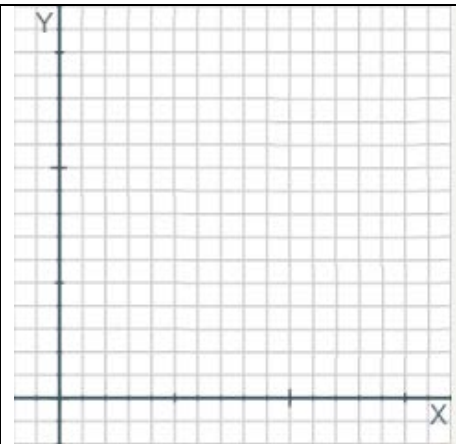


2. ¿Deporte o TV?

Las horas semanales que dedican 10 personas a hacer deporte y a ver la TV son:

DEPORTE (X)										
TV (Y)										

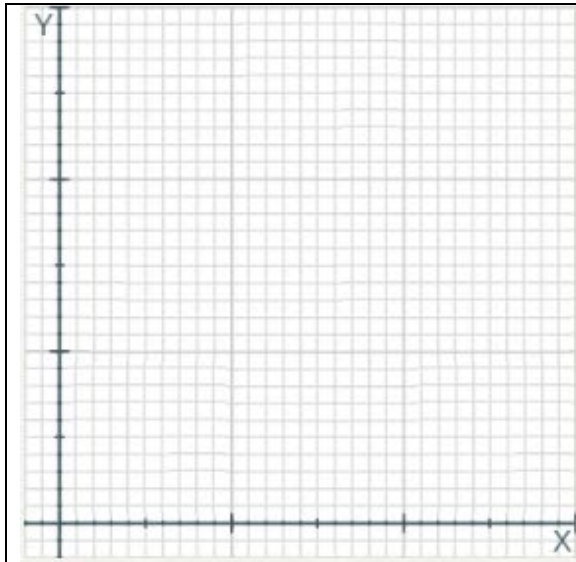
Dibuja la nube de puntos y calcula el coeficiente de correlación lineal:



3. Climatología

En un lugar se ha medido las temperaturas medias y las precipitaciones cada mes durante un año. Dibuja una la nube de puntos y calcula el coeficiente de correlación lineal:

°C (X)												
Litros/m ² (Y)												



4. Rectas de regresión

De una distribución bidimensional (X,Y) conocemos:

$X =$ $Y =$ $\sigma_x =$ $\sigma_y =$ $\sigma_{xy} =$


- a) Calcula el coeficiente de correlación lineal
- b) Calcula la recta de regresión de Y sobre X y el valor estimado de y para $x =$ ____

5. Con las cuentas hechas

De una distribución bidimensional (X,Y) conocemos:

$X =$ $Y =$ $\sigma_x =$ $\sigma_y =$ $\sigma_{xy} =$

- a) Calcula la recta de regresión de Y sobre X
- b) Calcula el valor estimado de y para $x =$ ____
- c) ¿Es fiable esta predicción?

Pulsa  para ir a la página siguiente.

Autoevaluación



Completa aquí cada uno de los enunciados que van apareciendo en el ordenador y resuélvelo, después introduce el resultado para comprobar si la solución es correcta.

1 ¿Cuántos grados corresponden en el diagrama de sectores al valor de frecuencia _____?

2 La mediana es _____

3 ¿Cuál es la moda?

4 ¿Cuál es el porcentaje de la muestra que corresponde a las dos primeras marcas?

5 ¿Cuál es el percentil más pequeño que deja por debajo los valores menores a 3?

6 ¿Cuál es la media?

7 Calcula la desviación típica

8 Asocia cada nube de puntos con su coeficiente de correlación:

1) _____ 2) _____ 3) _____

9 Calcula la covarianza

10 El centro de gravedad de una distribución bidimensional es (____,____) y la pendiente de la recta de regresión de Y sobre X es _____. Estima el valor de y para $x = \underline{\hspace{1cm}}$