



Estadística: Estadística bidimensional

Matemáticas I

1.º Bachillerato

Contenidos

Estadística

Estadística bidimensional

1. Introducción

Por lo que has visto en los temas anteriores, sabes que es posible obtener información del análisis de un conjunto de datos, como pueden ser la cantidad de CO₂ en la atmósfera, o el ritmo de fusión de los glaciares. Situaciones de este tipo, en donde existe una cierta relación entre las variables, las hemos estudiado de forma aislada. En este tema realizaremos una observación conjunta, y para ello será necesario ampliar las medidas estadísticas conocidas con nuevos parámetros.



Imagen de slack12 en [Flickr](#). Licencia [CC 2.0 by-nc-nd](#)



Imagen de atzu en [Flickr](#). Licencia [CC 2.0 by-nc-sa](#)

2. Variables bidimensionales. Dependencia e independencia entre variables



Imagen de dcarrero en [Flickr](#). Licencia [CC 2.0 by-nc-sa](#)

Si vives en una región agrícola en donde se cultiva el olivo, seguro que tienes respuesta para la siguiente pregunta: ¿depende la producción de aceite de una zona olivarera del índice de precipitaciones anuales?

Otra pregunta: ¿crees que existe alguna relación entre la cualificación de un empleado en una empresa, y el salario que recibe?

Seguro que también podrías emitir un juicio sobre si crees que existe alguna relación entre las emisiones de CO_2 y el aumento global de las temperaturas en la Tierra.

No es cuestión de opiniones, sino que son preguntas a las que podemos contestar consultando datos reales y comparándolos entre sí. Al estudiar dos factores sobre una población, estamos determinando una **variable estadística bidimensional**.



Importante

Una **Variable Estadística Bidimensional (X,Y)** es el resultado del estudio de dos factores X e Y en los elementos de una población.

Para cada elemento de estudio obtenemos un par de valores que notaremos (x_i, y_i) , donde x_i es el valor para el factor X, e y_i para el factor Y.

Por ejemplo, de una zona olivarera podemos estudiar la producción de aceite (X), y el índice de precipitaciones de un mismo año agrícola (Y). Si en dicha zona la producción de aceite de oliva de 2001 fue de 80.000 toneladas, y ese mismo año el índice de precipitaciones fue de 450 mm, tendríamos el par (80.000, 450).



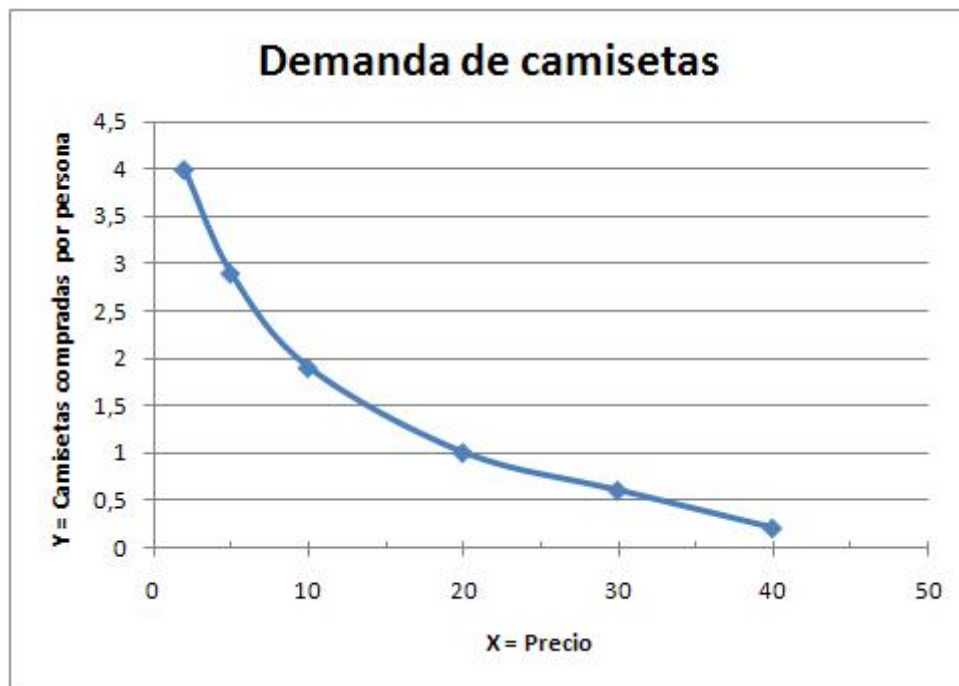
Comprueba lo aprendido

Al entrar en una tienda de ropa, ves en un cartel que venden una camiseta que te gusta por 20 €. A ese precio, puede que te la acabes comprando. Si costara sólo 10 euros,

seguramente te comprarías esa y puede que otro modelo del mismo precio. ¿Y si las vendieran por 3€? Es probable que compres más de dos... por si acaso.

Ésta es básicamente la Ley de la Demanda: la modificación del precio causa variaciones en la demanda de un artículo. Como tenemos dos factores X = precio del artículo, Y = número de artículos demandados, podemos establecer una variable estadística bidimensional.

Cuando tenemos suficientes datos, podemos hacer una primera estimación de la dependencia entre ambas variables. En la siguiente gráfica hemos representado los datos obtenidos al estudiar esta variable bidimensional.



a) Cuando aumenta el precio de la camiseta ¿qué ocurre con la demanda?

- ☐ Disminuye
- ☐ Aumenta
- ☐ No varía

¡CORRECTO!

Recuerda que el precio está en el eje horizontal y la demanda en el vertical

Recuerda que el precio está en el eje horizontal y la demanda en el vertical

Solución

1. Opción correcta
2. Incorrecto
3. Incorrecto

b) Veamos si lo hace en la misma proporción. Por ejemplo, ¿si aumento el precio el doble la demanda cae a la mitad?

 [Sugerencia](#)

- ☐ Sí
- ☐ No

Si vendemos a 20€ nos comprarían 1 camiseta, pero por 40€ no se vendería ni media.

¡MUY BIEN! En este caso podríamos decir que hay una **dependencia negativa** (porque al aumentar una variable, la otra disminuye) y **aleatoria** (porque la dependencia no es exacta)

Solución

1. Incorrecto
2. Opción correcta

Como has podido comprobar, basta realizar una mirada rápida a algún tipo de datos para ya ser capaces de sacar algunas conclusiones. Como por ejemplo:

- **Dependencia positiva:** Al aumentar la variable X, también aumenta la Y.
- **Dependencia negativa:** Al aumentar la variable X, disminuye la Y.
- **Sin dependencia:** No se observa ninguna relación entre las dos variables.

- **Dependencia funcional:** Podemos encontrar una relación exacta entre ambas variables que siempre se cumple. Por ejemplo, si estudias la relación entre el número de cajas de leche y el número de litros que se compra de una marca, tenemos una dependencia funcional, porque cada caja tiene siempre el mismo número de litros.
- **Dependencia aleatoria:** No hay una regla exacta que determine la relación entre ambas variables, como en el ejemplo anterior.

En la autoevaluación anterior tendríamos una dependencia negativa aleatoria.



Comprueba lo aprendido

Determina el tipo de dependencia que existe en las siguientes variables estadísticas bidimensionales, con tan sólo observar los datos que aparecen en las tablas.

a) Al alumnado de una clase se le pregunta sobre las horas que dedican diariamente a estudiar y las calificaciones de matemáticas obtenidas. Los resultados medios son los siguientes:

X = Horas de estudio 0 0,5 1 1,5 2 2,5

Y = Calificación obtenida 4,5 5,1 5,9 6 6,7 8,3

- ☐ Dependencia positiva funcional
- ☐ Dependencia positiva aleatoria
- ☐ Sin dependencia

Si fuera funcional, podríamos determinar exactamente la calificación que obtendría un alumno que estudiara 3 horas ¿crees que sería posible con estos datos?

¡CORRECTO! Al aumentar el número de horas, aumenta la calificación (positiva), pero no podemos apreciar ninguna relación exacta (aleatoria).

Comprueba si al aumentar el número de horas, aumenta o disminuye la calificación.

Solución

1. Incorrecto
2. Opción correcta
3. Incorrecto

b) En una empresa pagan las horas extra según la siguiente tabla, donde X = número de horas extra e Y = sueldo recibido.

X = N° de horas 1 2 3 4 5

Y = Sueldo 15 30 45 60 75

- ☐ Dependencia positiva funcional
- ☐ Dependencia positiva aleatoria
- ☐ Sin dependencia

¡CORRECTO! Podemos decir que cada hora siempre se cobra a 15€, por lo tanto la relación se puede determinar exactamente.

¿No crees que podrías determinar exactamente la relación que hay entre las horas de trabajo y el sueldo que percibe?

Comprueba si al aumentar el número de horas extra, también aumenta o disminuye el sueldo.

Solución

1. Opción correcta
2. Incorrecto
3. Incorrecto

c) Queremos saber si existe alguna relación entre el peso de un hombre y el número de hijos que tiene. Para ello, después de preguntar a una población, hemos obtenido los siguientes datos:

X = Peso 70 75 80 85 90 95

Y = N° de hijos 3 1 0 2 2 1

- ☐ Dependencia positiva funcional
- ☐ Dependencia positiva aleatoria
- ☐ Sin dependencia

¿Al aumentar el peso siempre aumenta el número de hijos?

¿Al aumentar el peso siempre aumenta el número de hijos?

¡CORRECTO! Como ya podías imaginar, no existe ninguna relación entre el peso de un hombre y el número de hijos que tiene

Solución

1. Incorrecto
 2. Incorrecto
 3. Opción correcta
-

2. Tablas de doble entrada

Como has visto en el apartado anterior, tan sólo con unos cuantos datos ya se pueden establecer relaciones entre dos variables. Pero lo normal para obtener resultados fiables es contar con una gran cantidad de datos estadísticos. En estos casos no es cómodo hacer una **tabla simple** como las que hemos utilizado en los últimos ejemplos, en los que sólo había seis o siete datos. Vamos a ver cómo organizar la información en una **tabla de doble entrada** cuando tenemos muchos pares de datos.

En la siguiente presentación verás con un ejemplo cómo se crean estas tablas.



[Tablas de doble entrada](#) from [saulvalper](#)



Caso práctico

Vamos a trabajar con la tabla del ejemplo anterior para sacar algunas conclusiones. Recuerda que X="número de días por mes en los que se supera el límite permitido de concentración de NO₂", e Y="número de días por mes en los que se supera el límite permitido de concentración de ozono".

		Y					
		$y_1=0$	$y_2=1$	$y_3=2$	$y_4=3$	$y_5=4$	f_i
X	$x_1=0$	7	1	2	2	0	12
	$x_2=1$	4	3	1	1	5	14
	$x_3=2$	3	0	0	2	0	5
	$x_4=3$	0	2	2	1	0	5
	f_j	14	6	5	6	5	36

a) ¿Cuántos meses tuvieron 2 días con niveles excesivos de NO_2 pero ninguno con nivel excesivo de ozono?

Respuesta: 3 meses. Tenemos que fijarnos en la fila de $x_3=2$ y en la columna de $y_1=0$.

b) ¿Cuántos meses tuvieron sólo un día de exceso de concentración de NO_2 en aire?

Respuesta: 14 meses. Como sólo nos piden información de X, tendremos que sumar todas las casillas que corresponden a $x_2=1$, que coincide con la suma parcial que tenemos en la última columna.



Comprueba lo aprendido

En una de las estaciones meteorológicas del Alto Guadalquivir se han recogido medidas de temperatura media ($^{\circ}\text{C}$) y precipitaciones medias (l/m^2) cada mes. Los datos de los años 2007 y 2008 son los siguientes:

(7,5 ; 7,7)	(10,2 ; 56,5)	(11 ; 28,6)	(12,8 ; 93,5)	(17,6 ; 76,5)	(22,5 ; 5)
(27,5 ; 0,2)	(26,3 ; 1,9)	(22,2 ; 44,4)	(16,7 ; 29,9)	(10,2 ; 61,3)	(7,5 ; 6,1)
(8,48 ; 62,51)	(10,89 ; 38,34)	(11,72 ; 16,46)	(14,71 ; 122,7)	(16,55 ; 64,66)	(23,87 ; 5,38)
(26,86 ; 10,54)	(27,16 ; 0,14)	(20,63 ; 48,58)	(16,18 ; 58,55)	(8,57 ; 67,71)	(6,46 ; 48,56)

El primer par significa que en Enero de 2007 la media de temperatura fue de 7,5°C y la media de precipitaciones fue de 7,7 l/m².

Con estos datos, completa la tabla de doble entrada en la que las variables son X = "Temperatura media mensual" e Y = "Precipitaciones medias mensuales". Fíjate que en este caso las variables se han agrupado por intervalos. En la primera casilla tendrás que contar el número de meses en los que la temperatura media está entre 0 y 10 grados, y las precipitaciones entre 0 y 30 l/m², que son los pares (7,5 ; 7,7) y (7,5 ; 6,1), por lo que en esa casilla pondremos un 2.

		Y					
		[0-30)	[30-60)	[60-90)	[90-120)	[120-150]	f_i
X	[0-10)	2	1	2	0	0	5
	[10-20)	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
	[20-30]	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
	f_j	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	24



Importante

De la tabla de doble entrada del ejercicio resuelto de arriba podemos crear las siguientes tablas de valores de las variables x e y.

x	f _i
0	12
1	14
2	5
3	5

y	f _j
0	14
1	6
2	5
3	6
4	5

A la tabla correspondiente a la variable x se le llama **distribución marginal de frecuencias absolutas de la variable x**, y se ha creado tomando los valores de la

variable x y los de la columna f_i que figura a la derecha de la tabla de doble entrada.

A la tabla correspondiente a la variable y se le llama **distribución marginal de frecuencias absolutas de la variable y** , y se ha creado tomando los valores de la variable y y los de la fila f_j que figura en la parte inferior de la tabla de doble entrada.

Análogamente se definen las **distribuciones marginales de frecuencias relativas de la variable x** y las **distribuciones marginales de frecuencias relativas de la variable y** .

De las tablas de arriba podemos hallar las medias y desviaciones típicas de cada una de las variables x e y , en este caso a estos parámetros se les llama **media marginal** y **desviación típica marginal**.



Importante

De la tabla de doble entrada de arriba podemos crear las siguientes tablas:

x	$y=1$
0	1
1	3
2	0
3	2

y	$x=1$
0	4
1	3
2	1
3	1
4	5

A la tabla de la derecha se le llama **distribución de frecuencias absolutas de la variable x condicionada a que $y=1$** , y se ha creado tomando los valores de la variable x y los de la columna $y_2=1$ que figura en la tabla de doble entrada.

A la tabla correspondiente a la variable y se le llama **distribución de frecuencias absolutas de la variable y condicionada a que $x=1$** , y se ha creado tomando los valores de la variable y y los de la fila $x_2=1$ que figura en la la tabla de doble entrada.

Análogamente se definen las **distribuciones de frecuencias relativas de la variable x condicionada a que $y=1$** y las **distribuciones de frecuencias relativas de la variable y condicionada a que $x=1$** .

En general a la variable x condicionada a que y tome un valor cualquiera se la denota como X_i/Y_j y a la variable y condicionada a que x tome un valor cualquiera se la denota como Y_j/X_i .



Caso práctico

Las calificaciones obtenidas por un grupo de alumnos en Biología y Física son:

Biología 3 4 6 7 5 8 7 3 5 4 8 5 5 8 8 8 5

Física 5 5 8 7 7 9 10 4 7 4 10 5 7 9 10 5 7

Se pide:

- Escribir la tabla de doble entrada de frecuencias absolutas.
- Hallar las distribuciones marginales así como la media y la desviación típica de las mismas.
- Hallar la distribución de la variable "calificaciones de Biología" condicionada a que la calificación de Física haya sido un 5. Y la distribución de la variable "calificaciones de Física" condicionada a que la calificación de Biología haya sido un 5.

a. La tabla de doble entrada es la siguiente (las notas de Biología y Física figuran en azul):

		Física							
		4	5	6	7	8	9	10	
Biología	3	1	1	0	0	0	0	0	2
	4	1	1	0	0	0	0	0	2
	5	0	1	0	4	0	0	0	5
	6	0	0	0	0	1	0	0	1
	7	0	0	0	1	0	0	1	2
	8	0	1	0	0	0	2	2	5
		2	4	0	5	1	2	3	17

b. Si llamamos X_i = calificaciones de Biología y Y_j = calificaciones de Física, las tablas de las distribuciones marginales pedidas son las siguientes:

Distribución marginal de frecuencias absolutas de X_i .

X_i	f_i
3	2
4	2
5	5
6	1
7	2
8	5

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{n} = \frac{99}{17} = 5,82$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n x_i^2 f_i}{n} - \bar{x}^2} = \sqrt{\frac{629}{17} - 5,82^2} = 1,045$$

Distribución marginal de frecuencias absolutas de Y_i .

Y_j	f_j
4	2
5	4
6	0
7	5
8	1
9	2
10	3

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{n} = \frac{119}{17} = 7$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n x_i^2 f_i}{n} - \bar{x}^2} = \sqrt{\frac{903}{17} - 7^2} = 4,118$$

c. Las distribuciones condicionadas pedidas son las siguientes:

$X_i/Y_j=5$	f_i
3	1
4	1
5	1
6	0
7	0
8	1

$Y_j/X_i=5$	f_j
4	0
5	1
6	0
7	4
8	0
9	0
10	0



Curiosidad

Si quieres ver los datos reales que toman estas estaciones de medición u otras estadísticas sobre Medio Ambiente, Población, Turismo, Comercio y muchas más en Andalucía, accede en el siguiente enlace al [Instituto de Estadística y Cartografía de Andalucía](#).

3. Gráficas y nubes de puntos

Si te fijas en los pasos que hemos dado hasta este momento, verás que es un proceso muy lógico:

1. Nos planteamos una pregunta sobre la relación entre dos parámetros.
2. Tomamos suficientes datos de ambos parámetros sobre la población que nos interesa.
3. Organizamos estos datos en una tabla simple o de doble entrada.

El siguiente paso será visualizar estos datos en una **gráfica**, de modo que nos resulte más fácil dar respuesta a nuestra pregunta inicial.

Volvamos a la tabla de doble entrada que vimos en el ejercicio resuelto del apartado anterior, donde comparábamos el número de días mensuales en los que se superaba la concentración máxima de NO_2 y de Ozono en el aire:

		Y					
		$y_1=0$	$y_2=1$	$y_3=2$	$y_4=3$	$y_5=4$	f_i
X	$x_1=0$	7	1	2	2	0	12
	$x_2=1$	4	3	1	1	5	14
	$x_3=2$	3	0	0	2	0	5
	$x_4=3$	0	2	2	1	0	5
	f_j	14	6	5	6	5	36

El par $(0,0)$ se podría representar como un punto en una gráfica habitual de ejes cartesianos, pero en este caso tenemos que hacer ver de algún modo que la frecuencia de ese par es 7. A continuación verás algunos ejemplos:

a) Histograma tridimensional:

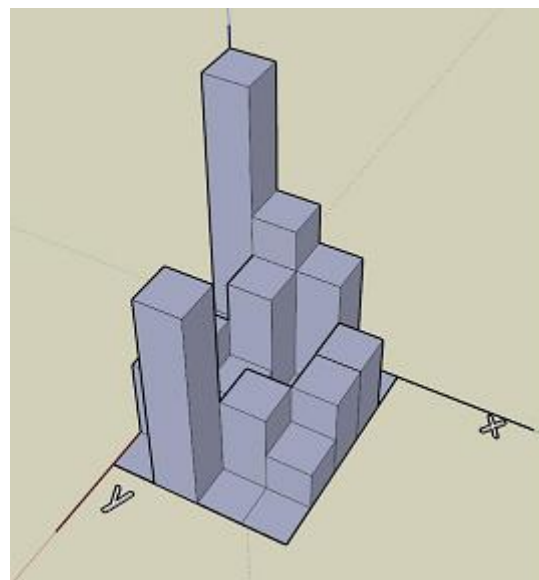
Para representar la información partimos de tres ejes cartesianos.

En los ejes X e Y marcamos los posibles valores de cada variable (en nuestro caso 0, 1, 2 y 3 para X, y 0, 1, 2, 3 y 4 para Y). Cada cuadrado representa un par de valores.

La altura de cada cuadrado será la correspondiente frecuencia de ese par de valores.

Fíjate cómo en nuestro caso el par con mayor frecuencia es el $(0,0)$, que se repite 7 veces, y por tanto es el prisma de mayor altura.

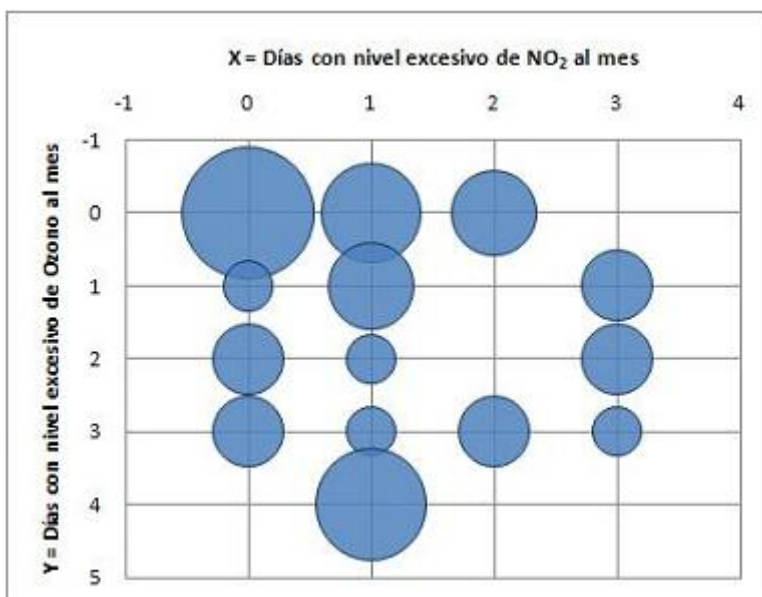
El siguiente sería el $(1,4)$ que tiene frecuencia 5.



b) Diagrama de dispersión o de Burbujas:

En este caso partimos de un par de ejes cartesianos X e Y en los que representamos los valores de ambos parámetros.

En lugar de puntos, representamos circunferencias en las que su **superficie es proporcional a la frecuencia**. Ojo, no son proporcionales los radios sino las superficies.



Los pares de datos que tienen frecuencia 0 no se representan.

En el caso en el que tengamos variables continuas, el diagrama de dispersión será diferente, pues en lugar de representar los datos que hemos acumulado en la tabla, haremos una gráfica con todos los pares de datos que tomamos en la fase previa.

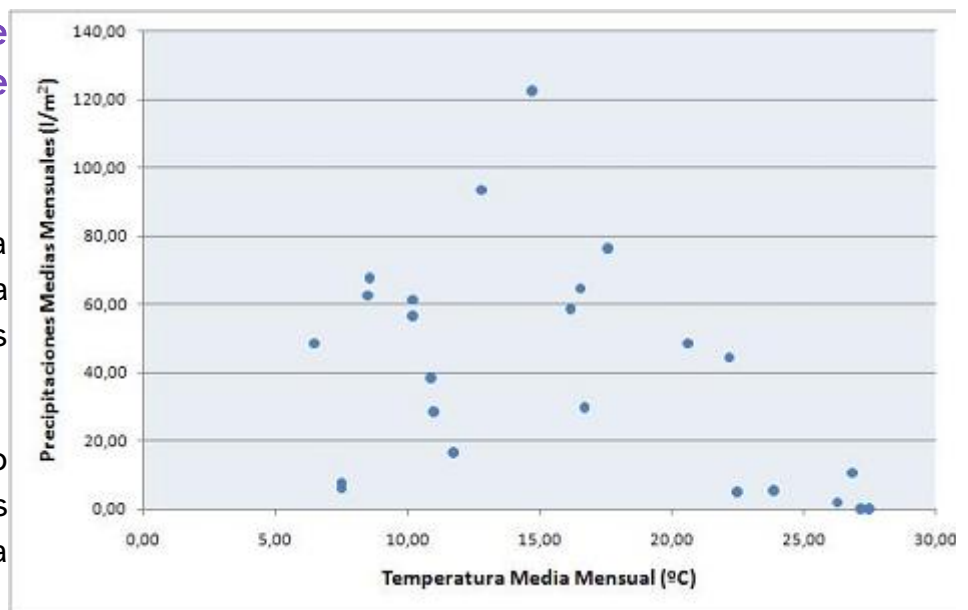
Por ejemplo: vamos a representar los valores de Temperatura y Precipitaciones medias mensuales en una determinada estación climatológica que tenías en la autoevaluación del apartado anterior.

(7,5 ; 7,7)	(10,2 ; 56,5)	(11 ; 28,6)	(12,8 ; 93,5)	(17,6 ; 76,5)	(22,5 ; 5)
(27,5 ; 0,2)	(26,3 ; 1,9)	(22,2 ; 44,4)	(16,7 ; 29,9)	(10,2 ; 61,3)	(7,5 ; 6,1)
(8,48 ; 62,51)	(10,89 ; 38,34)	(11,72 ; 16,46)	(14,71 ; 122,7)	(16,55 ; 64,66)	(23,87 ; 5,38)
(26,86 ; 10,54)	(27,16 ; 0,14)	(20,63 ; 48,58)	(16,18 ; 58,55)	(8,57 ; 67,71)	(6,46 ; 48,56)

c) Diagrama de dispersión o Nube de puntos:

Al igual que el Diagrama de Burbujas, se representa sobre un par de ejes cartesianos.

En este caso, cada punto representa un par de datos de la Variable Estadística Bidimensional.



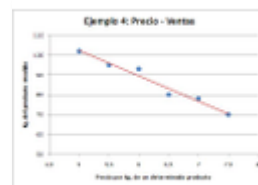
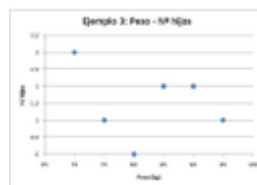
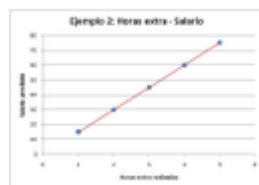
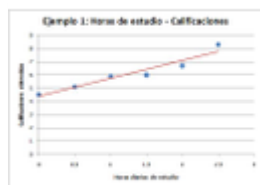
Las Nubes de Puntos también nos ayudan a ver la dependencia entre las variables. Si recuerdas, en el primer apartado vimos que la dependencia podía ser:

- **Dependencia positiva:** Al aumentar la variable X, también aumenta la Y.
- **Dependencia negativa:** Al aumentar la variable X, disminuye la Y.
- **Sin dependencia:** No se observa ninguna relación entre las dos variables.
- **Dependencia funcional:** Podemos encontrar una relación exacta entre ambas variables que siempre se cumple. Por ejemplo, si estudias la relación entre el número de cajas de leche y el número de litros que se compra de una marca, tenemos una dependencia funcional, porque cada caja tiene siempre el mismo número de litros.
- **Dependencia aleatoria:** No hay una regla exacta que determine la relación entre ambas variables, como en el ejemplo anterior.

Mira las siguientes gráficas, que representan los ejemplos del primer apartado. Verás que es mucho más fácil ver así la dependencia:



Galería de imágenes



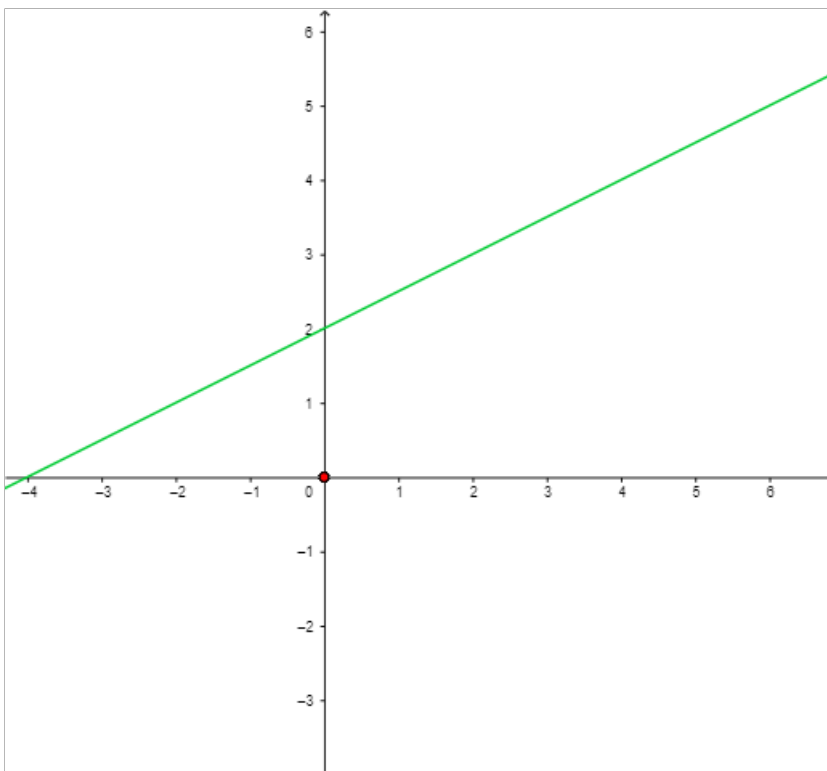
Comprueba lo aprendido

Una compañía química está estudiando el uso de un fertilizante líquido que han inventado en una determinada planta. Para ello miden dos variables: X = "cantidad diaria de fertilizante que se aporta a la planta (en ml)" e Y = "crecimiento de la planta al cabo de 10 días (en cm)".

Los resultados son los siguientes pares de datos:

X	1	2	3	4	5	6	7	8	9	10
Y	2	2.9	3	4	4	4.5	5	4.8	4.9	5.2

Utiliza la siguiente gráfica para representar estos puntos en los ejes de coordenadas. Para ello, escribe cada par en las columnas X e Y. Ojo, para escribir los números decimales debes utilizar el punto, no la coma.



a) A la vista de la gráfica, ¿crees que existe dependencia funcional?

- ☐ Sí
- ☐ No

¿Crees que hay una relación exacta entre las dos variables?

Como puedes observar, los puntos no se adaptan a una recta o una parábola exactamente. Por tanto no hay dependencia funcional.

Solución

1. Incorrecto
2. Opción correcta

b) ¿Cómo clasificarías la dependencia según la nube de puntos?

- ☐ Positiva
- ☐ Negativa
- ☐ Sin dependencia

Muy bien. Al aumentarla cantidad de fertilizante, también aumenta el tamaño de la planta.

Observa si la nube de puntos crece o decrece

Observa si la nube de puntos crece o decrece

Solución

1. Opción correcta
2. Incorrecto
3. Incorrecto

c) La línea verde de la gráfica representa los resultados que se obtienen con el fertilizante de la competencia, ¿con cuál crees que se obtienen mejores resultados?

- ☐ Con el de la compañía
- ☐ Con el de la competencia

Observa que los resultados de la compañía quedan por debajo de los de la competencia.

Muy bien. Nuestra nube de puntos está por debajo de los resultados de la competencia.

Solución

1. Incorrecto
2. Opción correcta

Resumen



Importante

Una **Variable Estadística Bidimensional (X,Y)** es el resultado del estudio de dos factores X e Y en los elementos de una población.

Para cada elemento de estudio obtenemos un par de valores que notaremos (x_i, y_i) , donde x_i es el valor para el factor X, e y_i para el factor Y.

Por ejemplo, de una zona olivarera podemos estudiar la producción de aceite (X), y el índice de precipitaciones de un mismo año agrícola (Y). Si en dicha zona la producción de aceite de oliva de 2001 fue de 80.000 toneladas, y ese mismo año el índice de precipitaciones fue de 450 mm, tendríamos el par (80.000, 450).



Importante

A partir de la tabla de doble entrada

		Y					f_i
		$y_1=0$	$y_2=1$	$y_3=2$	$y_4=3$	$y_5=4$	
X	$x_1=0$	7	1	2	2	0	12
	$x_2=1$	4	3	1	1	5	14
	$x_3=2$	3	0	0	2	0	5
	$x_4=3$	0	2	2	1	0	5
	f_j	14	6	5	6	5	36

podemos crear las siguientes tablas de valores de las variables x e y.

x	f _i
0	12
1	14
2	5
3	5

y	f _j
0	14
1	6
2	5
3	6
4	5

A la tabla correspondiente a la variable x se le llama **distribución marginal de frecuencias absolutas de la variable x**, y se ha creado tomando los valores de la variable x y los de la columna f_i que figura a la derecha de la tabla de doble entrada.

A la tabla correspondiente a la variable y se le llama **distribución marginal de frecuencias absolutas de la variable y**, y se ha creado tomando los valores de la variable y y los de la fila f_j que figura en la parte inferior de la tabla de doble entrada.

Análogamente se definen las **distribuciones marginales de frecuencias relativas de la variable x** y las **distribuciones marginales de frecuencias relativas de la variable y**.

De las tablas de arriba podemos hallar las medias y desviaciones típicas de cada una de las variables x e y, en este caso a estos parámetros se les llama **media marginal** y **desviación típica marginal**.



Importante

De la tabla de doble entrada de arriba podemos crear las siguientes tablas:

x	y=1
0	1
1	3
2	0
3	2

y	x=1
0	4
1	3
2	1
3	1
4	5

A la tabla de la derecha se le llama **distribución de frecuencias absolutas de la variable x condicionada a que y=1**, y se ha creado tomando los valores de la variable x y los de la columna $y_2=1$ que figura en la tabla de doble entrada.

A la tabla correspondiente a la variable y se le llama **distribución de frecuencias absolutas de la variable y condicionada a que x=1**, y se ha creado tomando los valores de la variable y y los de la fila $x_2=1$ que figura en la la tabla de doble entrada.

Análogamente se definen las **distribuciones de frecuencias relativas de la variable x condicionada a que y=1** y las **distribuciones de frecuencias relativas de la variable y**

condicionada a que $x=1$.

En general a la variable x condicionada a que y tome un valor cualquiera se la denota como X_i/Y_j y a la variable y condicionada a que x tome un valor cualquiera se la denota como Y_j/X_i .



Importante

Para representar una distribución bidimensional podemos utilizar los siguientes gráficos:

Histograma tridimensional

Para representar la información partimos de tres ejes cartesianos. En los ejes X e Y marcamos los posibles valores de cada variable.

La altura de cada cuadrado será la correspondiente frecuencia de ese par de valores.

Diagrama de dispersión o de Burbujas

En este caso partimos de un par de ejes cartesianos X e Y en los que representamos los valores de ambos parámetros. En lugar de puntos, representamos circunferencias en las que su superficie es proporcional a la frecuencia. Ojo, no son proporcionales los radios sino las superficies. Los pares de datos que tienen frecuencia 0 no se representan.

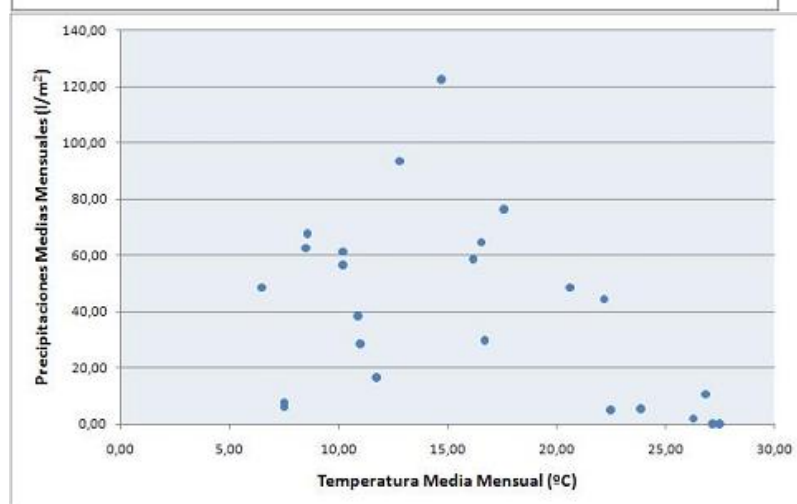
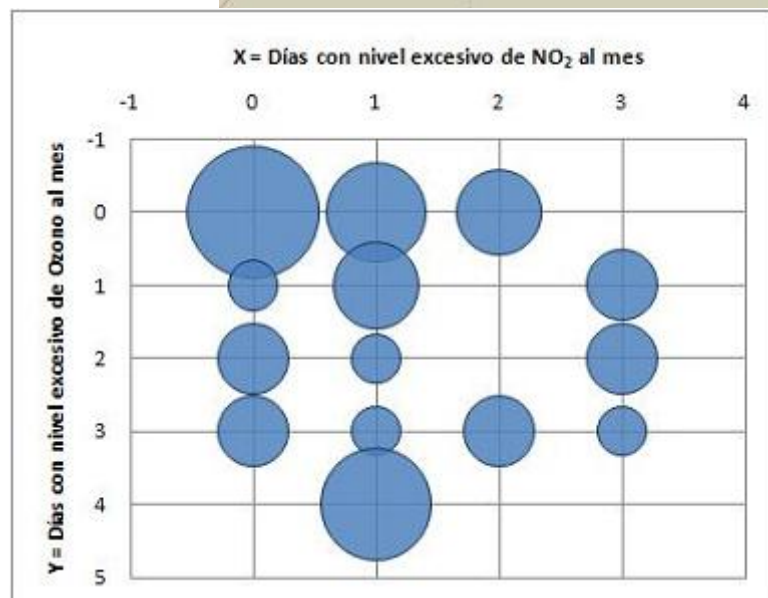
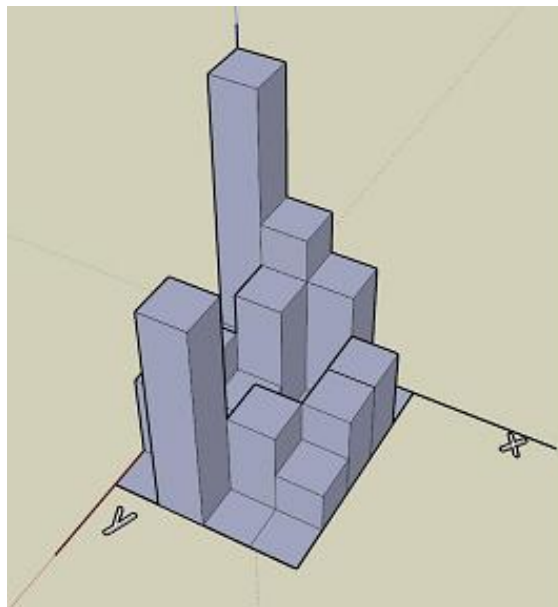
Diagrama de dispersión o nube de puntos

Al igual que el Diagrama de Burbujas, se representa sobre un par de ejes cartesianos. En este caso, cada punto representa un par de datos de la Variable Estadística Bidimensional.

Histograma
tridimensional

Diagrama de
burbujas

Diagrama de
dispersión



Aviso legal

Las páginas externas no se muestran en la versión imprimible

Aviso Legal

El presente texto (en adelante, el "**Aviso Legal**") regula el acceso y el uso de los contenidos desde los que se enlaza. La utilización de estos contenidos atribuye la condición de usuario del mismo (en adelante, el "**Usuario**") e implica la aceptación plena y sin reservas de todas y cada una de las disposiciones incluidas en este Aviso Legal publicado en el momento de acceso al sitio web. Tal y como se explica más adelante, la autoría de estos materiales corresponde a un trabajo de la **Comunidad Autónoma Andaluza, Consejería de Educación y Deporte (en adelante Consejería de Educación y Deporte)**.

Con el fin de mejorar las prestaciones de los contenidos ofrecidos, la