5. ESTADÍSTICA BIDIMENSIONAL

5.1. Describir y comparar conjuntos de datos de distribuciones bidimensionales, con variables discretas o continuas, procedentes de contextos relacionados con el mundo científico y obtener los parámetros estadísticos más usuales, mediante los medios más adecuados (lápiz y papel, calculadora, hoja de cálculo) y valorando, la dependencia entre las variables.

5.1.1. Elabora tablas bidimensionales de frecuencias a partir de los datos de un estudio estadístico, con variables discretas y continuas.

5.1.1.1.

Elabora una tabla de doble entrada a partir de la siguiente tabla bidimensional:

X	1	3	6	10	12
Y	5	10	18	22	25
f_i	1	2	4	3	2

a) *¿Qué porcentaje de los datos representa en valor (6,18) en la variable (X,Y)?*
b) Calcula la media y la desviación típica de la distribución marginal de Y.

5.1.1.2.

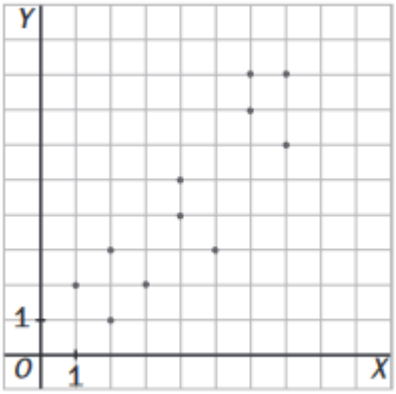
Las notas obtenidas por un grupo de alumnos en las materias de estadística (E) y macroeconomía (M), están reflejadas en la siguiente tabla:

E	3	4	6	7	5	8	7	3	5	4	8	5	5	8	8	8	5
M	5	5	8	7	7	9	10	4	7	4	10	5	7	9	10	5	7

a) Halla la tabla de frecuencias
b) Calcula las desviaciones típicas de las variables *E* y *M.*

5.1.1.3.

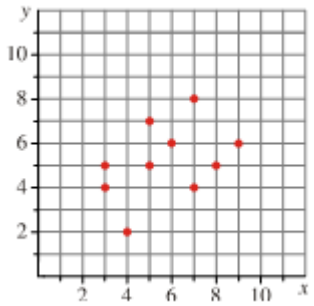
Dado el siguiente diagrama de dispersión:



Elabora una tabla de doble entrada que represente la situación.

5.1.1.4.

Dado el siguiente diagrama de dispersión:



Elabora una tabla de doble entrada que represente la situación.

5.1.2. Calcula e interpreta los parámetros estadísticos más usuales en variables bidimensionales.

5.1.2.1.

Una variable bidimensional viene dada por la siguiente tabla:

X	3	6	8	20	25
Y	0,2	0,4	0,5	1,2	1,7

a) Calcula las medias y las desviaciones típicas de las variables *X* e *Y*.
b) Calcula la covarianza de las variable bidimensional *(X, Y)*.

5.1.2.2.

Una variable bidimensional viene dada por la siguiente tabla:

X	1	2	3	4	5	6	7	8	9
Y	5	4	3	3	1	3	4	6	3

a) Calcula las medias y las desviaciones típicas de las variables *X* e *Y*.
b) Calcula la covarianza de las variable bidimensional *(X, Y)*.

5.1.2.3.

Calcula el coeficiente de correlación e interprétalo para la siguiente variable bidimensional:

X	6	8	11	14	15	18	20
Y	3	4	2	5	6	7	8

5.1.2.4.

Una variable bidimensional viene dada por la siguiente tabla:

X \ Y	1 - 16	16 - 31	31 - 46	46 - 60
10 - 19	1	1	1	
19 - 28		8	3	
28 - 37		3	7	1
37 - 45		2	3	

X representa el % de la población urbana en las provincias que gobierna el PSOE e Y representa la renta media por hogar.

a) Calcula las medias y las desviaciones típicas de las variables *X* e *Y*.
b) Calcula la covarianza de las variable bidimensional *(X, Y)*.

5.1.3. Calcula las distribuciones marginales y diferentes distribuciones condicionadas a partir de una tabla de contingencia, así como sus parámetros (media, varianza y desviación típica).

5.1.3.1.

Copia y completa la siguiente tabla y contesta a las preguntas:

Y \ X	A	B	C	Total
a	2		1	4
b		2	2	5
c			0	7
Total	7	6		

a) ¿Qué porcentaje de datos presenta la característica B en la variable unidimensional *X*?
b) ¿Qué porcentaje de datos presenta la característica c en la variable unidimensional *Y*?
c) ¿Qué porcentaje de datos presenta la característica (B,c) en la variable bidimensional *(X,Y)*?

5.1.3.2.

La siguiente distribución bidimensional se expresa en la siguiente tabla de correlaciones:

X/ Y	< 60	60 - 80	80 - 100	100 - 150	> 150
50 - 100	20	18	2	1	0
100 - 200	25	40	30	2	1
200 - 350	5	10	15	25	3
350 - 500	0	5	15	20	8
> 500	0	1	2	7	10

X representa los ingresos familiares mensuales en unidades de 10€. La variable Y representa los m^2 de las casas de la vivienda familiar.
a) Calcula las distribuciones marginales de X e Y.
b) Obtén la distribución de la superficie de la familia condicionada al intervalo modal de los ingresos.
c) Calcula la distribución de los ingresos condicionada al intervalo mediano de la superficie de las viviendas.

5.1.3.3.

En una academia para aprender a conducir se han estudiado las semanas de asistencia a clase de sus alumnos y las semanas que tardan en aprobar el examen teórico (desde que se apuntaron a la autoescuela). Los datos correspondientes a seis alumnos son:

***X:* Asistencia**	**6**	**1**	**4**	**3**	**5**	**8**
***Y:* Aprobado**	**6**	**5**	**5**	**6**	**5**	**10**

a)Determina las distribuciones marginales y calcula la desviación típica para cada una

5.1.4. Decide si dos variables estadísticas son o no dependientes a partir de sus distribuciones condicionadas y marginales

5.1.4.1.

La siguiente distribución bidimensional se expresa en la siguiente tabla de correlaciones:

X/ Y	< 60	60 - 80	80 - 100	100 - 150	> 150
50 - 100	20	18	2	1	0
100 - 200	25	40	30	2	1
200 - 350	5	10	15	25	3
350 - 500	0	5	15	20	8
> 500	0	1	2	7	10

X representa los ingresos familiares mensuales en unidades de 10€. La variable Y representa los m^2 de las casas de la vivienda familiar.
a) Calcula las distribuciones marginales de X e Y. (Hecho en 5.1.3.2)
b) ¿Son independientes los ingresos familiares y el tamaño de la vivienda donde habitan?

5.1.4.2.

A partir de la siguiente distribución de correlaciones:

X \ Y	1	2	3
-1	0	1	0
0	1	0	1
1	0	1	0

¿Son independientes la variables X e Y?

5.1.5. Usa adecuadamente medios tecnológicos para organizar y analizar datos desde el punto de vista estadístico, calcular parámetros y generar gráficos estadísticos.

5.1.5.1.

Realiza el siguiente ejercicio ayudándote de la aplicación geogebra.

Una empresa realiza un estudio de los efectos de la publicidad sobre sus ventas, los resultados son los siguientes:

Gasto	1	2	3	4	5	6
N.º productos	9	18	32	27	40	46

Los datos están dados en miles de euros.

a) Calcula el coeficiente de correlación.
b) Determina la recta de regresión del Nº de productos sobre el gasto en publicidad.
c) Si se invierte 15000€ en publicidad. ¿Qué ventas se estima que se produciría? ¿Es fiable la predicción?

5.2. **Interpretar la posible relación entre dos variables y cuantificar la relación lineal entre ellas mediante el coeficiente de correlación, valorando la pertinencia de ajustar una recta de regresión y, en su caso, la conveniencia de realizar predicciones, evaluando la fiabilidad de las mismas en un contexto de resolución de problemas relacionados con fenómenos científicos.**

5.2.1. Distingue la dependencia funcional de la dependencia estadística y estima si dos variables son o no estadísticamente dependientes mediante la representación dela nube de puntos.

5.2.1.1.

Observa la siguiente variable bidimensional:

N.º de cigarrillos consumidos al día	3	6	8	20	25
Índice de mortalidad	0,2	0,4	0,5	1,2	1,7

a) Representa la nube de puntos.
b) Indica el tipo de correlación.

5.2.1.2.

Una variable bidimensional viene dada por la siguiente tabla:

N.º de horas de estudio	1	2	3	4	5
N.º de horas de televisión	5	4	3	3	1

c) Representa la nube de puntos.
d) Indica el tipo de correlación.

5.2.1.3.

Un grupo de 10 amigos se ha presentado a una prueba de oposición. Anotaron el número de horas que dedicaron a estudiar la semana antes del examen y la nota obtenida en la prueba. La información se recoge en la siguiente tabla:

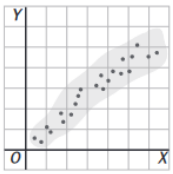
Horas de estudio	21	15	10	15	20	30	18	20	25	16
Nota	9	7	5	2	7	8	8	6	5	4

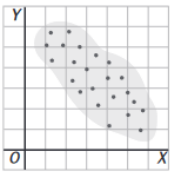
Representa los datos mediante una nube de puntos e indica cuál de estos valores te parece más apropiado para el coeficiente de correlación: 0,92; -0,44; -0,92; 0,44.

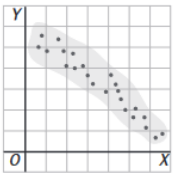
5.2.1.4.

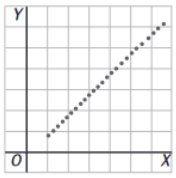
Asocia a cada nube de puntos el coeficiente de correlación más adecuado:

$r = 1$ $r = 0{,}92$ $r = -0{,}25$ $r = -0{,}78$

a)


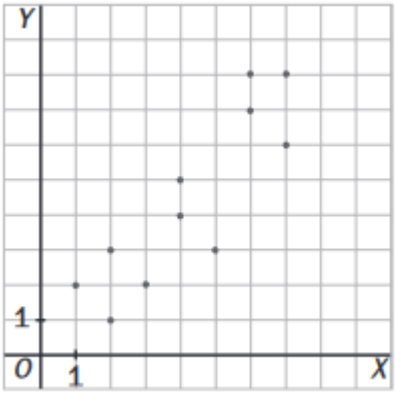
c)


b)


d)


5.2.1.5.

Dado el siguiente diagrama de dispersión:



	a) ¿Qué tipo de correlación tienen las dos variables? ¿Fuerte o débil? ¿Positiva o negativa?

5.2.2. Cuantifica el grado y sentido de la dependencia lineal entre dos variables mediante el cálculo e interpretación del coeficiente de correlación lineal.

5.2.2.1.

Calcula el coeficiente de correlación e interprétalo para la siguiente variable bidimensional:

X	6	8	11	14	15	18	20
Y	3	4	2	5	6	7	8

5.2.2.2.

Una variable bidimensional viene dada por la siguiente tabla:

X	13	14	11	13	14	14	15	22
Y	54	52	54	53	53	50	49	42

a) Calcula el coeficiente de correlación.
b) Halla la recta de regresión de *Y* sobre *X*.
c) Si x=12 ¿Cuánto valdrá y?
d) ¿Es fiable la predicción del apartado c? ¿Por qué?

5.2.2.3.

Se han realizado unas pruebas de habilidad (puntúan de 0 a 5) en un grupo de alumnos. Las siguientes puntuaciones corresponden a las obtenidas por seis alumnos en dos de ellas:

1ª Prueba	5	5	4	3	2	4
2ª Prueba	4	3	4	4	3	2

Calcula el coeficiente de correlación. ¿Cómo es la relación entre las variables?

5.2.2.4.

Se ha realizado una encuesta preguntando por el número de personas que habitan el hogar familiar y el número de habitaciones que tiene la casa. La tabla siguiente recoge la información obtenida:

N.º de personas	3	5	4	6	5	4
N.º de habitaciones	2	3	4	4	3	3

Halla la covarianza y el coeficiente de correlación. ¿Cómo es la relación entre las dos variables?

5.2.3. Calcula las rectas de regresión de dos variables y obtiene predicciones a partir de ellas.

5.2.3.1.

Una variable bidimensional viene dada por la siguiente tabla:

X	13	14	11	13	14	14	15	22
Y	54	52	54	53	53	50	49	42

a) Calcula el coeficiente de correlación.

b) Halla la recta de regresión de *Y* sobre *X*.
c) Si x=12 ¿Cuánto valdrá y?
d) ¿Es fiable la predicción del apartado c? ¿Por qué?

5.2.3.2.

Se ha analizado en distintos modelos de impresoras cuál es el coste por página (en céntimos de euro) en blanco y negro y cuál es el coste por página si esta es en color. La siguiente tabla nos da los seis primeros pares de datos obtenidos:

***X:* B Y N**	**8**	**11**	**17**	**21**	**14**	**10**
***Y:* Color**	**33**	**49**	**95**	**106**	**58**	**53**

a) Halla la recta de regresión de *Y* sobre *X*.

5.2.3.3.

Se ha medido el peso, en kilogramos, y el volumen, en litros, de distintos tipos de maletas, obteniendo los resultados que se recogen en esta tabla:

***X:* Volumen**	**97**	**102**	**94**	**107**	**92**	**98**
***Y:* Peso**	**6,9**	**7,1**	**6,7**	**7,4**	**5,8**	**6,1**

a) Halla la recta de regresión de *Y* sobre *X*.
b) Calcula el peso para un volumen de 120. (sabemos que r=0,79)

5.2.4. Evalúa la fiabilidad de las predicciones obtenidas a partir de la recta de regresión mediante el coeficiente de determinación lineal.

5.2.4.1.

En seis institutos de la misma zona se ha estudiado la nota media de los estudiantes de 1º de bachillerato en Matemáticas y en Inglés, obteniéndose la información que se recoge en la siguiente tabla:

***X:* Matemáticas**	**6,5**	**5,2**	**6**	**6,5**	**7**	**6**
***Y:* Inglés**	**7**	**5**	**5**	**6**	**7,5**	**5**

a) Halla la recta de regresión de *Y* sobre *X*.
b) Calcula la nota de matemáticas para una nota de inglés de 5,5. ¿Es fiable la predicción? (sabemos que r=0,79)

5.2.4.2.

Se ha realizado una encuesta preguntando por el número de personas que habitan el hogar familiar y el número de habitaciones que tiene la casa. La tabla siguiente recoge la información obtenida:

N.º de personas	**3**	**5**	**4**	**6**	**5**	**4**
N.º de habitaciones	**2**	**3**	**4**	**4**	**3**	**3**

Determina (de forma lineal) cuantas personas vivirían en una casa de 6 habitaciones. ¿Es fiable la predicción?

5.2.4.3.

Se ha analizado en distintos modelos de impresoras cuál es el coste por página (en céntimos de euro) en blanco y negro y cuál es el coste por página si esta es en color. La siguiente tabla nos da los seis

primeros pares de datos obtenidos:

X: B Y N	8	11	17	21	14	10
Y: Color	33	49	95	106	58	53

a) Halla la recta de regresión de *Y* sobre *X*. (Hecho en el 5.2.3.2.)
b) ¿Cuánto nos costaría imprimir una página en color en una impresora en la que el coste por página en blanco y negro fuera de 12 céntimos de euro? ¿Es fiable la estimación? (Sabemos que $r = 0{,}97$).

5.2.4.4.

Se ha medido el peso, en kilogramos, y el volumen, en litros, de distintos tipos de maletas, obteniendo los resultados que se recogen en esta tabla:

X: Volumen	97	102	94	107	92	98
Y: Peso	6,9	7,1	6,7	7,4	5,8	6,1

a) Halla la recta de regresión de *Y* sobre *X*. (Hecho en el 5.2.3.3.)
b) Calcula el peso para un volumen de 120. (sabemos que r=0,79)

5.3 Utilizar el vocabulario adecuado para la descripción de situaciones relacionadas con la estadística, analizando un conjunto de datos o interpretando de forma crítica informaciones estadísticas presentes en los medios de comunicación, la publicidad y otros ámbitos, detectando posibles errores y manipulaciones tanto en la presentación de los datos como de las conclusiones.

5.3.1. Describe situaciones relacionadas con la estadística utilizando un vocabulario adecuado.

5.3.1.1.

Escoge una variable de las dos primeras tablas y, a partir de la tercera indica que tipo de relación hay:

1. Número de cigarros fumados al día.
2. Número de horas sentado al cabo del día.
3. Velocidad a la que voy en el coche.
4. Nota en una asignatura.

A. Peso del individuo.
B. Número de horas diarias de móvil.
C. Capacidad pulmonar.
D. Espacio que recorro.

W. Dependencia funcional.
X. Correlación negativa fuerte.
Y. Correlación positiva media.
Z. Correlación positiva medio-fuerte.

5.3.1.2.

Relaciona los elementos de las siguientes tablas:

	a) $r = -1$ b) $r = 0{,}95$ c) $r = 0{,}26$ d) $r = -0{,}89$ e) $r = -0{,}62$ I. Correlación positiva fuerte II. Independientes III. Dependencia funcional IV. Correlación negativa débil V. Correlación negativa fuerte
5.3.1.3.	Un profesor de 1º de bachillerato ha realizado un estudio sobre la altura y el peso de sus alumnos obteniendo las siguientes conclusiones: Variable: Altura (m) — Media: 1,72 — Desviación típica: 0,29 Variable: Peso (kg) — Media: 65,4 — Desviación típica: 4,68 ¿Qué varía más el peso o la altura? ¿Por qué?
5.3.1.4.	En una variable bidimensional (X,Y), su coeficiente de correlación lineal es 0,48 y la pendiente de su recta de regresión es 1,34. Sabiendo que la suma de las desviaciones típicas de X e Y es 7,33. Calcula cada una de ellas y la covarianza.