

# MATEMÁTICA ANÁLISIS Y ENFOQUES NM



COLEGIO

**SAN  
AGUSTÍN**

EST. 1966



# Dispersión y Correlación



COLEGIO

**SAN  
AGUSTÍN**

EST. 1966



Competencia:

- Resuelve problemas de gestión de datos e incertidumbre.



COLEGIO

**SAN  
AGUSTÍN**

EST. 1966

Se ataron diferentes pesos a un resorte vertical y se midió la longitud del resorte. Los resultados se muestran en la tabla.

Carga (Kg)	1	2	3	4	5	7	9	11
Longitud (cm)	11	12.5	13.5	14.6	15.8	18.1	20.2	22.4

- Realice su gráfica de dispersión rotulando correctamente los ejes.
- Halle el punto  $\bar{M}(\bar{x}, \bar{y})$
- Utilice su Excel para hallar el coeficiente de correlación  $r$ , y comente sobre su resultado.
- Halle la ecuación de la recta de regresión.
- Utilice su ecuación para estimar la longitud del resorte cuando la carga tiene un peso de 6 kg.
- Estime cual sería la carga para una longitud de 21 cm.





# DISTRIBUCIONES BIDIMENSIONALES

- Al realizar experimentos de cualquier índole, siempre buscamos la relación entre dos o más variables.
- Al tener los valores relacionados entre las variables, debemos tratar de obtener una función que las relacione, para poder realizar algunas predicciones posteriores.



Mi amor no es amor de mercado, porque un amor sangrado no es amor de  
lucrar. **Silvio Rodríguez**



# DISTRIBUCIONES BIDIMENSIONALES

- F. Galton (1822-1917) fue el primero en asignar un número a un conjunto de variables y, de esta manera, obtener el grado de relación existente entre ellas. Así se propició el comienzo de la teoría de regresión.
- K. Pearson (1857-1936) hizo importantes aportes a la estadística como el coeficiente de correlación lineal o el test de Pearson para la bondad de ajuste en una distribución empírica.
- A Galton y Pearson se les considera hoy día los padres de la Estadística moderna.





# VARIABLES BIDIMENSIONALES

- Pulso y temperatura de un enfermo en un hospital.
- La fuerza y la aceleración de un cuerpo.
- Extensión y número de habitantes en las distintas ciudades de un país.
- Ingresos y gastos de una familia en un mes.
- Latitudes y temperaturas medias en las capitales de un conjunto de países.
- La cantidad de lana y carne en unas ovejas con cierta dieta suministrada.



Mi amor no es amor de mercado, porque un amor sangrado no es amor de lucrar. **Silvio Rodríguez**

# Correlación

---

Cuando dos conjuntos de datos parecen estar conectados, es decir, un conjunto de datos depende del otro, entonces hay varios métodos que se pueden usar para comprobar si hay o no alguna correlación.

Uno de estos métodos es el diagrama de dispersión.

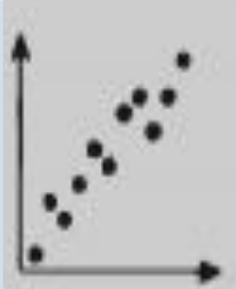
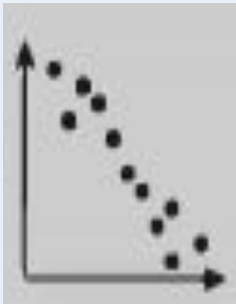

Los datos se pueden representar en un diagrama de dispersión con la variable independiente en el eje horizontal y la variable dependiente en el eje vertical. El patrón de los puntos dará una imagen visual de cuán estrechamente están relacionadas las variables, en caso de que sea así.





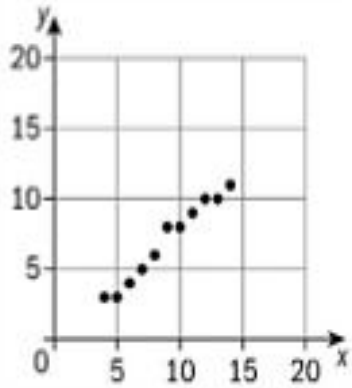
# Tipos de correlación.

---

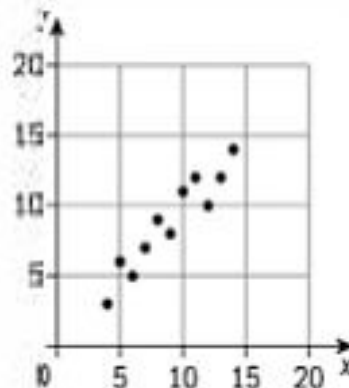
Correlación	Definición	Gráfico
Positiva	Cuando la variable dependiente crece a medida que crece la variable independiente.	
Negativa	Cuando la variable dependiente decrece a medida que crece la variable independiente.	
No existe	Cuando los puntos están dispersos en forma aleatoria en el diagrama.	

---

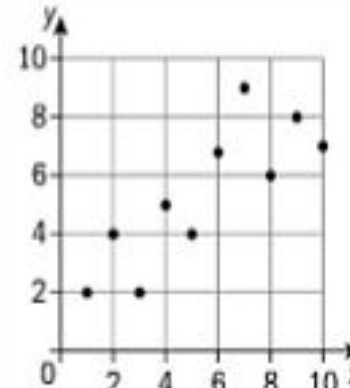
La correlación también se puede describir como fuerte, moderada o débil.



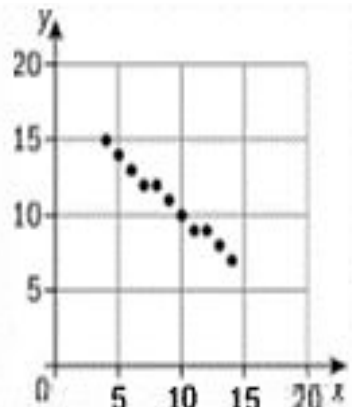
**Positiva fuerte**



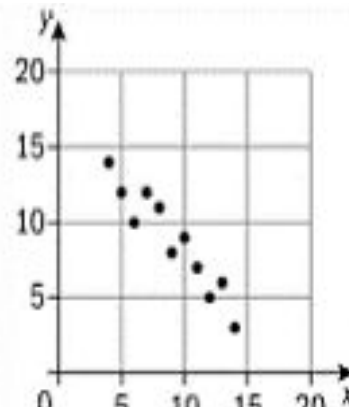
**Positiva moderada**



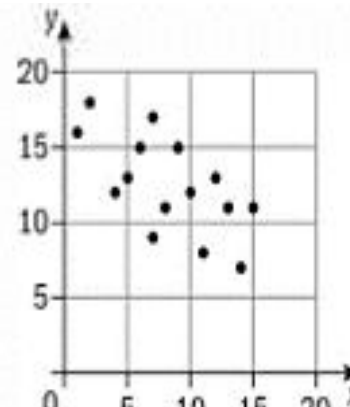
**Positiva débil**



**Negativa fuerte**



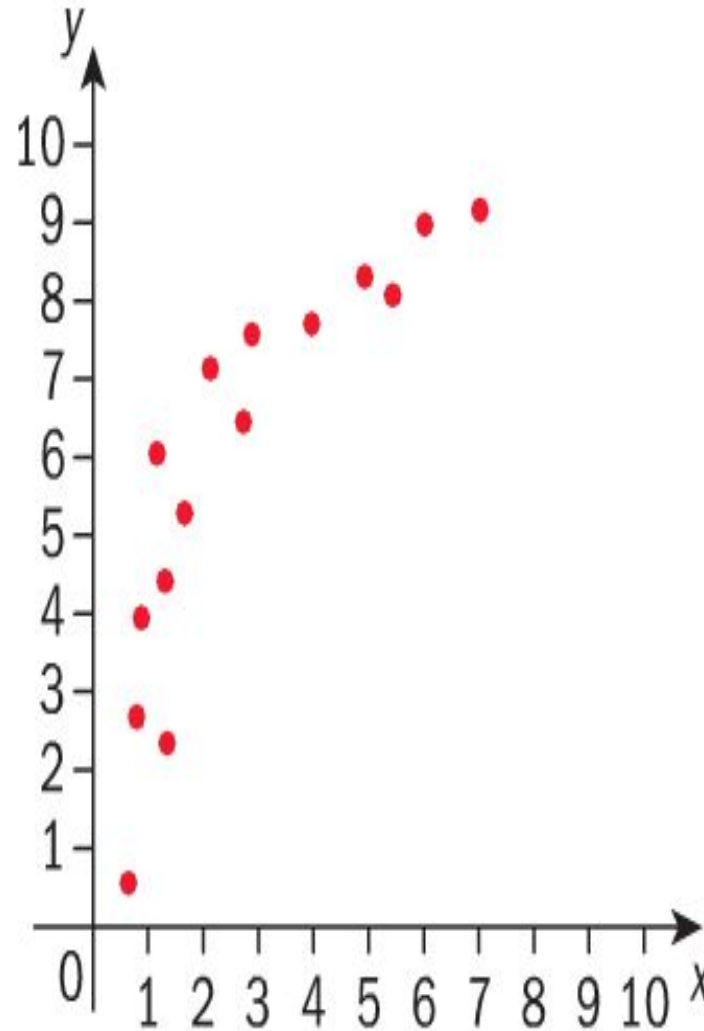
**Negativa moderada**



**Negativa débil**



The points on this graph would be represented by a curve.  
There is a **non-linear** relationship between the variables.



# Coeficiente de correlación momento – producto de Pearson

Es útil conocer la fuerza de correlación entre dos conjuntos de datos que se cree que están relacionados.

El **Coeficiente de correlación momento – producto de Pearson,  $r$** , es una forma de hallar un valor numérico que se puede usar para determinar la fuerza de una correlación lineal entre dos conjuntos de datos.

El **Coeficiente de correlación momento – producto de Pearson,  $r$**  puede tomar valores entre -1 y +1 inclusive.

Cuando  $r = -1$ , hay una **correlación negativa perfecta** entre los dos conjuntos.

Cuando  $r = 0$  **no hay correlación**.

Cuando  $r = +1$ , hay una **correlación positiva perfecta** entre los dos conjuntos.

Cuando  $r$  está entre 0 y 0.25, la **correlación es muy débil**.

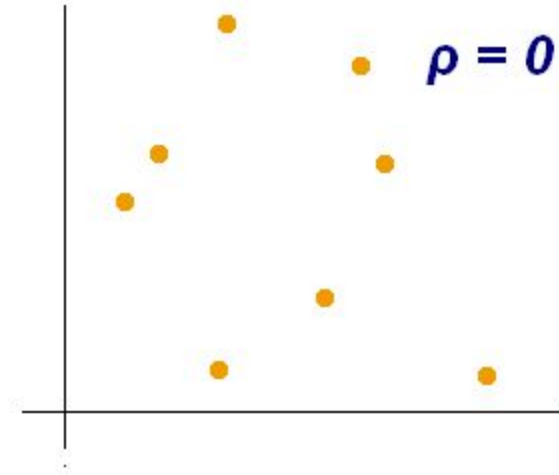
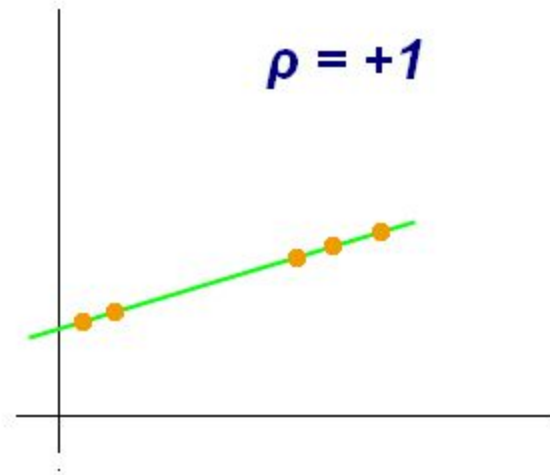
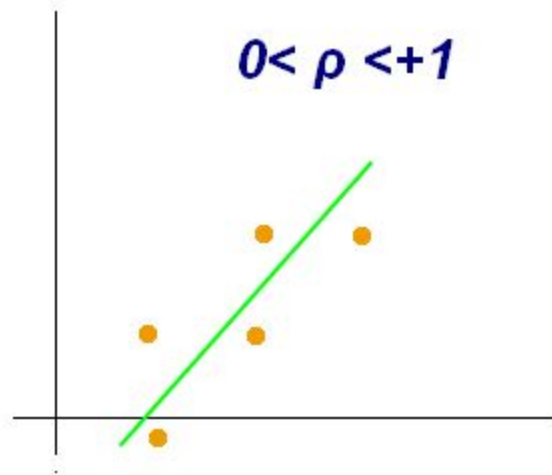
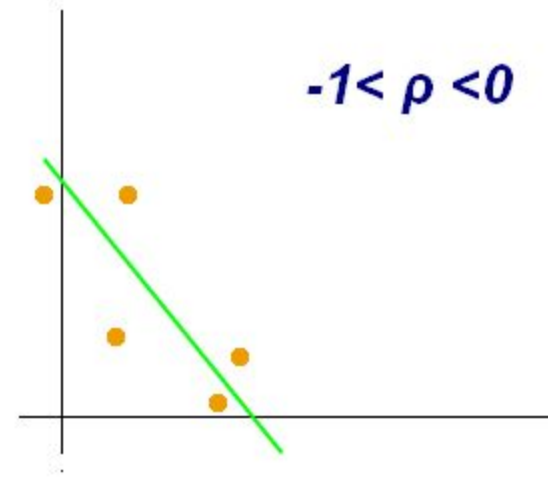
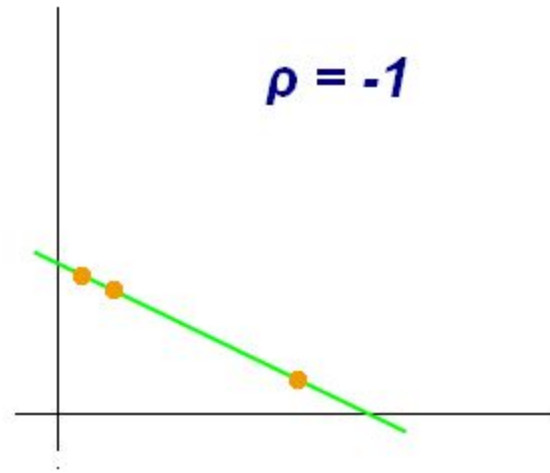
Cuando  $r$  está entre 0.25 y 0.5, la **correlación es débil**.

Cuando  $r$  está entre 0.5 y 0.75 hay una **correlación moderada**.

Cuando  $r$  está entre 0.75 y 1, la **correlación es fuerte**.







---

# RECTA DE REGRESIÓN

- **CÁLCULO:**
- Se tienen muchas maneras de calcular los coeficientes de la recta de regresión.
- Uno de los métodos más usados es el llamado de mínimos cuadrados.
- Usaremos la calculadora para obtener los coeficientes.



# La recta de regresión

---

La recta de regresión de  $y$  sobre  $x$  es una versión más precisa de la recta de ajuste óptimo, comparada con la hallada por aproximación.

Se representa:  $y = ax + b$



# Caso 1

---

Diez alumnos entrenan para una caminata de beneficencia. La tabla muestra la cantidad promedio de horas por semana que cada integrante entrena y el tiempo que tardan en completar la caminata.

Tiempo de entrenamiento (horas)	9	8	12	3	25	6	10	5	6	21
Tiempo en completar la caminata (minutos)	15,9	14,8	15,3	18,4	13,8	16,2	14,1	16,1	16	14,2





# Caso 1

a) Sitúe estos puntos en un diagrama de dispersión y comente acerca del tipo de correlación.

Usando la calculadora sigue los siguientes pasos:

Presiona el menú STAT, luego EXE.

Anota en List 1 todos los datos de Tiempo de entrenamiento.

Anota en List 2 todos los datos de Tiempo en completar la caminata.

- Presiona F1 (GRPH)
- Presiona F1 (GRPH1)
- Observando el gráfico podemos decir que hay una correlación negativa moderada.

	List 1	List 2	List 3	List 4
SUB				
1	9	15.9		
2	8	14.8		
3	12	15.3		
4	3	18.4		

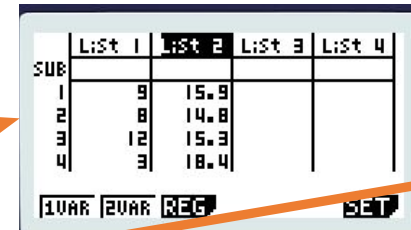
GRPH CALC TEST INTR DIST ▸



# Caso 1

b) Halle la media de x (Tiempo de entrenamiento) y la media de y (Tiempo de completar la caminata)

- Presionamos EXIT-EXIT-F2(CALC)
- Presiona SET: Escribimos el orden
  - F1 (1 VAR)



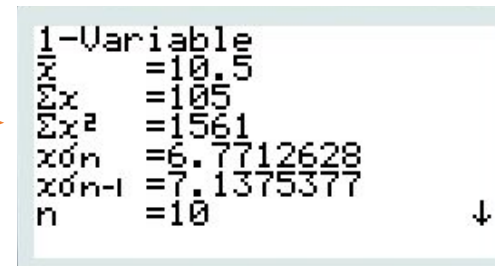
	List 1	List 2	List 3	List 4
SUB				
1	9	15.9		
2	8	14.8		
3	12	15.3		
4	3	18.4		



```
1Var XList : List1
1Var Freq : 1
2Var XList : List2
2Var YList : List2
2Var Freq : 1
LIST
```

- Presionamos EXIT

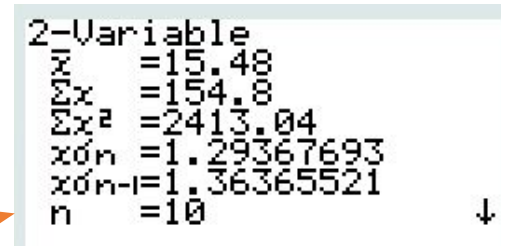
Observamos que la media de la variable x: Tiempo de entrenamiento es 10,5



```
1-Variable
x̄ = 10.5
Σx = 105
Σx² = 1561
x̄n = 6.7712628
x̄n-1 = 7.1375377
n = 10
↓
```

- Presionamos EXE – F2 (2 VAR)

Observamos que la media de la variable Y: Tiempo en completar la caminata es 15,48



```
2-Variable
x̄ = 15.48
Σx = 154.8
Σx² = 2413.04
x̄n = 1.29367693
x̄n-1 = 1.36365521
n = 10
↓
```

Hallar las medias de las variables nos sirve para colocarlas en el gráfico: (10,5; 15,48)



# Caso 1

c) Halle el coeficiente de correlación  $r$  y comente acerca de este valor

- Presiona exit-SET: ORDENAR
- Presionamos EXE
- Presiona F3 (REG)
- Presiona F1 (X)
- Presiona F1 ( $ax+b$ )



```
1Var XList :List1
1Var Freq :List2
2Var XList :List1
2Var YList :List2
2Var Freq :1

LIST

LinearReg
a =-0.1465648
b =17.0189312
r =-0.7671384
r²=0.58850145
MSe=0.86085496
y=ax+b
```

$r = -0,767$ , significa que hay una correlación negativa fuerte-perfecta.



# Caso 1

d) Halle la ecuación de la recta de regresión.

Según lo que hallamos en la calculadora:

$$y = ax + b \quad y = -0,15x + 17$$

En este menú no podemos ver la gráfica.

Para ver la gráfica:

- Presionamos EXIT, hasta llegar al menú GRPH

```
LinearReg
a = -0.1465648
b = 17.0189312
r = -0.7671384
r² = 0.58850145
MSe = 0.86085496
y = ax + b
```

COPY

	List 1	List 2	List 3	List 4
SUB				
1	9	15.9		
2	8	14.8		
3	12	15.3		
4	3	18.4		

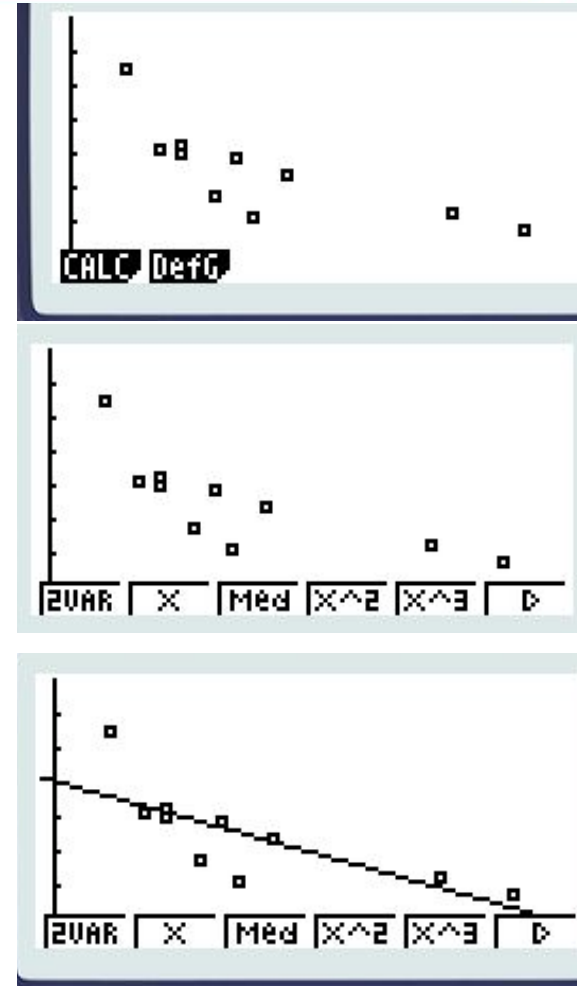
GRPH CALC TEST INTR DIST ▶





# Caso 1

- Presiona F1 (GRPH)
- Presiona F1 (GRPH1)
- Presiona F1 (CALC)
- Presiona F2 (X)
- Presiona F1 ( $ax+b$ )
- Presiona F6 (DRAW)



Se observa la gráfica con la recta de regresión.



# Caso 1

---

e) Utilizando su ecuación, estime cuántos minutos tardará en completar la caminata un alumno que entrena 18 horas por semana.

$$x = 18$$

$$y = -0,15x + 17$$

$$y = -0,1465648(18) + 17$$

$$y = 14,4$$

También se puede trasladar la función al menú Graph y predecir



## Caso 2

---

La tabla muestra la cantidad de ratones que están a la venta en una veterinaria después de una cantidad determinada de semanas.

Tiempo (x semanas)	3	5	6	9	11	13
Cantidad de ratones (y)	41	57	61	73	80	91



## Caso 2

---

a) Sitúe estos puntos en un diagrama de dispersión y comente acerca del tipo de correlación.

Usamos los mismos procedimientos de la calculadora.

Observando el gráfico podemos decir que hay una correlación positiva moderada.





## Caso 2

---

b) Halle la media de x y la media de y

Observamos que la media de la variable x: Tiempo es 7,83

Observamos que la media de la variable Y: Cantidad de ratones es 67,2

Hallar las medias de las variables nos sirve para colocarlas en el gráfico: (7,83; 67,2)

c) Halle el coeficiente de correlación r y comente acerca de este valor

$r = 0,999$ , significa que hay una correlación fuerte.



## Caso 2

---

d) Halle la ecuación de la recta de regresión de  $y$  sobre  $x$ .

$$y = 4,63x + 30,9$$

e) Utilice su recta de regresión para predecir la cantidad de ratones a la venta después de 10 semanas

$$x = 10$$

$$y = 4,63x + 30,9$$

$$y = 4,62929061(10) + 30,9$$

$$y = 77,2 \approx 77$$



## Caso 3

---

La tabla muestra la distancia que recorre un tren entre varios sitios en India y el costo del viaje.

Distancia (Km)	204	1407	1461	793	1542	343	663	780
Costo (rupias)	390	2200	2270	1390	2280	490	1200	1272

- a) Halle el coeficiente de correlación  $r$ , y comente sobre su resultado
- b) Halle la ecuación de la recta de regresión.
- c) Utilice su ecuación para estimar el costo de un viaje de 1 000km.



## Caso 4

---

Se ataron diferentes pesos a un resorte vertical y se midió la longitud del resorte. Los resultados se muestran en la tabla.

Carga (Kg)	0	2	3	5	6	7	9	11
Longitud (cm)	15	16.5	17.5	18.5	18.8	19.2	20	20.4

- a) Halle el coeficiente de correlación  $r$ , y comente sobre su resultado
- b) Halle la ecuación de la recta de regresión.
- c) Utilice su ecuación para estimar la longitud del resorte cuando la carga tiene un peso de 8 kg.



## Caso 5

Luis es un nadador entusiasta. Para su proyecto de estudios matemáticos, quiere investigar si hay una correlación entre la longitud del brazo de un nadador y el tiempo que este tarda en nadar 200 m. Elige 15 socios del club de natación para que naden 200 m. sus tiempos (y segundos) y las longitudes de sus brazos (x cm) se muestran en la siguiente tabla.

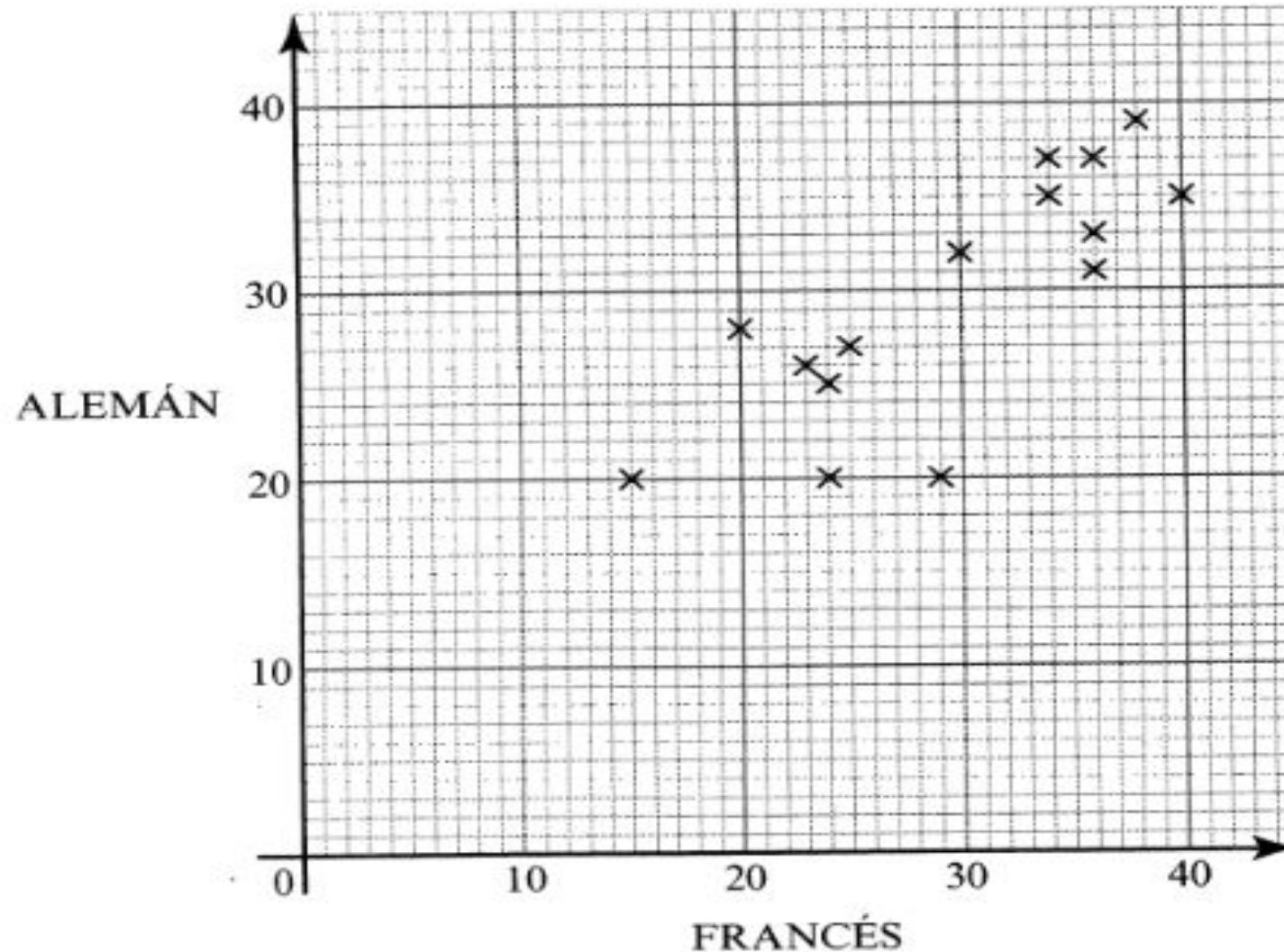
Longitud del brazo (cm)	78	72	74	67	79	58	62	67	71	69	75	65	73	59	60
Tiempo (Segundos)	130	135	132	143	133	148	140	139	135	145	129	140	130	145	142

- Halle la media de  $x$  e  $y$
- Halle el coeficiente de correlación  $r$ , y comente sobre su resultado
- Halle la ecuación de la recta de regresión de  $y$  sobre  $x$ .
- Utilizando su ecuación, estime cuantos segundos tardará en nadar 200 m un nadador que tiene una longitud de brazo igual a 70 cm.



## Caso 6

La siguiente figura es una gráfica de las notas conseguidas por un grupo de alumnos en un examen de francés y en un examen de alemán. La nota media en el examen de francés es de 29 y la nota media en el examen de alemán es de 31.





## Caso 6

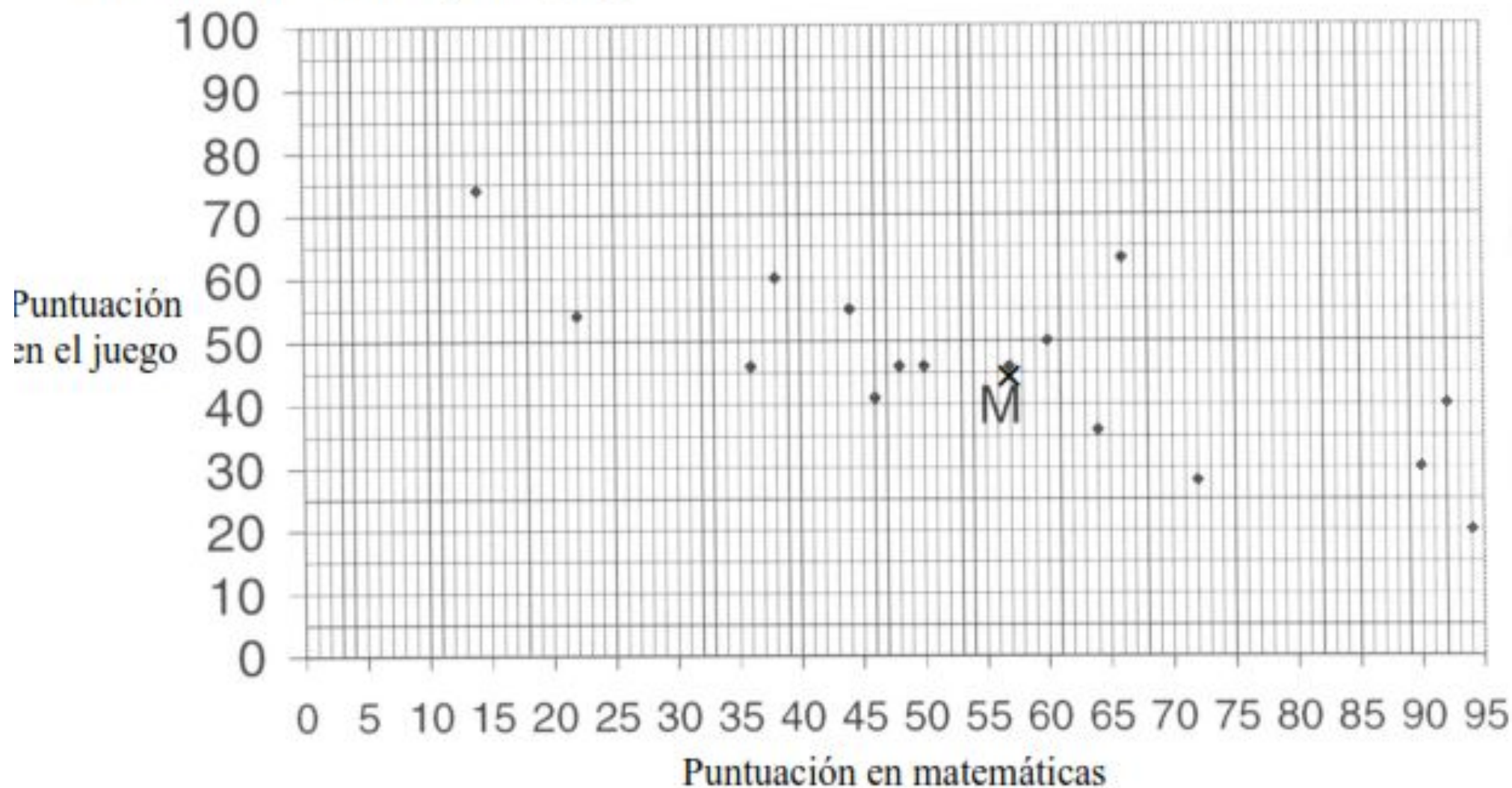
---

- (a) Describa la relación entre las notas conseguidas en los dos exámenes.
- (b) En la gráfica señale el punto  $M$  que representa la media de la distribución.
- (c) Dibuje una recta de máximo ajuste conveniente.
- (d) Idris consiguió una nota de 32 en el examen de francés. Use su gráfica para dar una estimación de la nota de Idris conseguida en el examen de alemán.



# Caso 7

Un grupo de 15 estudiantes realizó un examen de matemáticas. Luego, los estudiantes jugaron a un juego de computador. En el diagrama que sigue se muestran las puntuaciones obtenidas en el examen y en el juego.



# Caso 7

---

La puntuación media obtenida en el examen de matemáticas fue de 56,9 y la puntuación media obtenida en el juego de computador fue de 45,9. El punto M tiene coordenadas (56,9 , 45,9).

(a) Describa la relación entre ambos conjuntos de puntuaciones.

Una recta de ajuste óptimo pasa por el punto (0, 69).

(b) Trace esta recta de ajuste óptimo en el diagrama.

Jane llegó tarde al examen y obtuvo 45 en matemáticas.

(c) Por medio de la gráfica o de otra manera, estime la puntuación que espera obtener Jane en el juego de computador; exprese su respuesta redondeada al entero más próximo.

