

# Tema 6. RELACIONES ENTRE VARIABLES: CORRELACIÓN



## CONTENIDO:

1. DIAGRAMA DE DISPERSIÓN
2. COVARIANZA
3. COEFICIENTE DE CORRELACIÓN LINEAL

### Lecturas recomendadas:

- Capítulo 8 de *Introducción a la Estadística para las Ciencias Sociales*, de D. Peña y J. Romo.

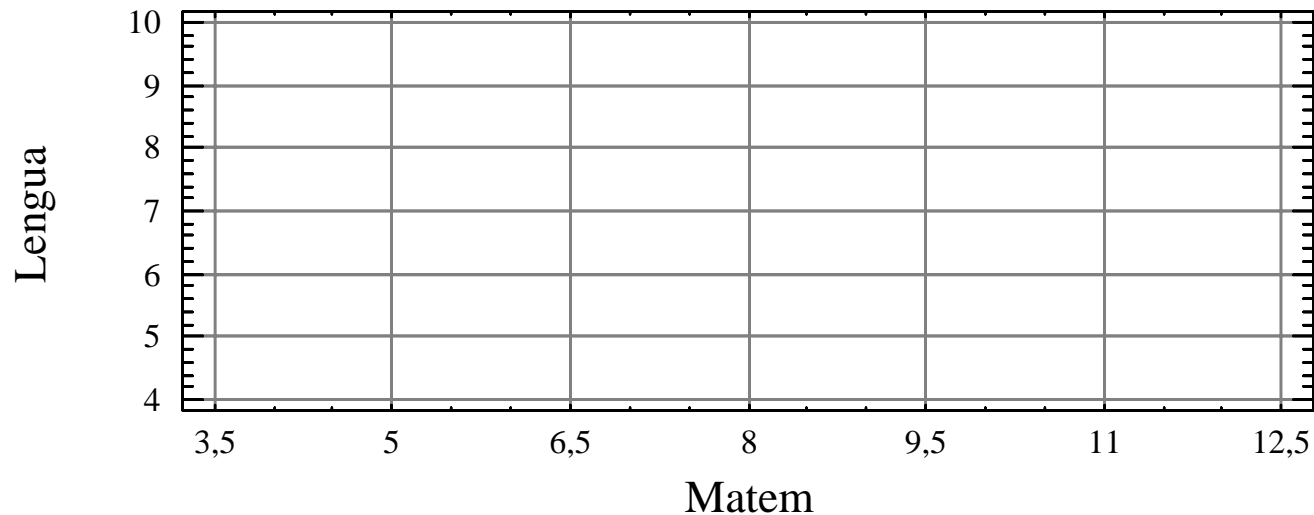
# 1. Relación entre variables continuas

## EJEMPLO 1: DIAGRAMA DE DISPERSIÓN

Notas de 10 alumnos/as en matemáticas y lenguaje. Representa las notas de cada alumno en el gráfico:

Alumno	1	2	3	4	5	6	7	8	9	10
X=Nota Matem.	6	4	8	5	3,5	7	5	10	5	4
Y=Nota Lengua	6,5	4,5	7	5	4	8	7	10	6	5

Plot of Lengua vs Matem



# 1. Relación entre variables continuas

## EJERCICIO 1: COVARIANZA

Calcula las notas medias y represéntalas en el gráfico.

¿Qué alumnos tienen notas por encima de la media tanto en Matemáticas como en Lengua? Para esos alumnos, ¿Qué signo toma el producto de las diferencias a la media?

$$(X_i - \bar{X})(Y_i - \bar{Y})$$

¿Qué alumnos tienen notas por debajo de la media tanto en Matemáticas como en Lengua? ¿Qué signo toma el producto anterior para estos alumnos?

¿Qué alumnos tienen la nota de Matemáticas por debajo de la media pero la de Lengua por encima de la media? ¿Qué signo toma el producto anterior para estos alumnos?

Ahora calcula el promedio de los productos anteriores para todos los alumnos. ¿Qué mide esta cantidad?

# 1. Relación entre variables continuas

## EJERCICIO 1: COVARIANZA

$X_i$	6	4	8	5	3,5	7	5	10	5	4	
$Y_i$	6,5	4,5	7	5	4	8	7	10	6	5	
$X_i Y_i$											

Ahora calcula el promedio de los productos anteriores para todos los alumnos.

$$s_{x,y} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

¿Qué signo tomaría esta cantidad si todos los individuos estuviesen por encima o por debajo de la media en las dos asignaturas? ¿Qué mide esta cantidad?

# 1. Relación entre variables continuas

## COVARIANZA

$S_{x,y} > 0$  → X e Y tienden a moverse en la misma dirección

$S_{x,y} < 0$  → X e Y tienden a moverse en direcciones opuestas.

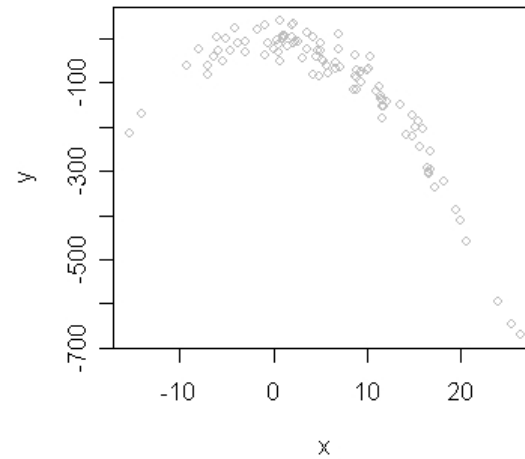
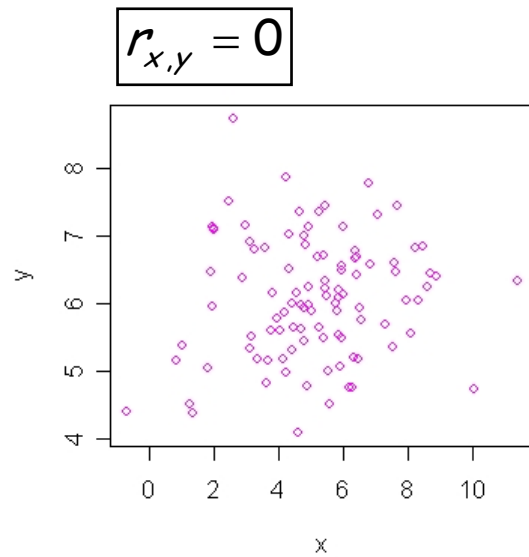
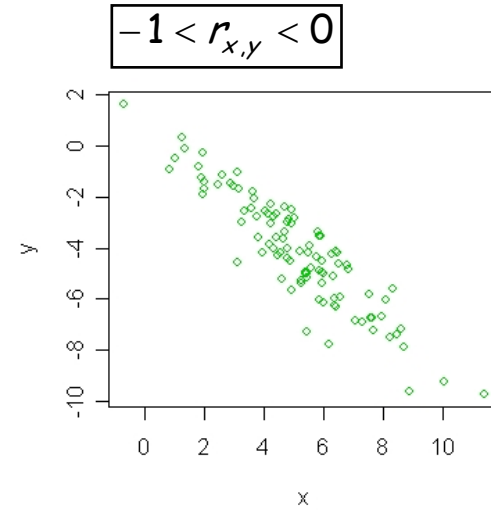
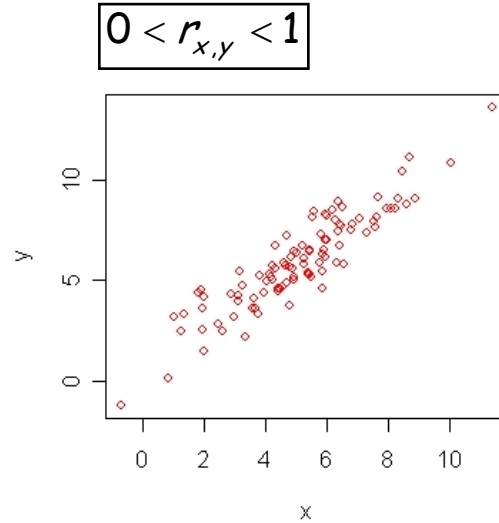
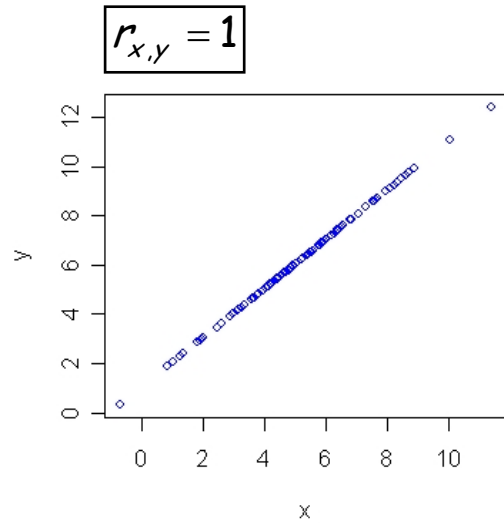
$S_{x,y} = 0$  → X e Y no están relacionadas linealmente.

## COEFICIENTE DE CORRELACIÓN LINEAL

$$r_{x,y} = \frac{s_{x,y}}{s_x s_y}$$

$$-1 < r_{x,y} < 1$$

# 1. Relación entre variables continuas



# ACTIVIDADES

## ACTIVIDAD 1:

Calcula la covarianza y el coeficiente de correlación lineal entre los ingresos (X) y los gastos (Y) anuales de 5 individuos medidos en euros mostrados en la tabla:

Individuo	X=Ingresos	Y=Gastos
1	10000	90000
2	7000	8000
3	15000	13000
4	21000	20000
5	14000	13500

Ahora pasa los datos de la tabla a miles de euros y vuelve a calcular la covarianza y el coeficiente de correlación lineal ¿Son iguales que antes?

# ACTIVIDADES

## ACTIVIDAD 2:

Inventa o busca datos de dos variables medidas a un conjunto de unidades, que estén correlados positivamente. Haz el gráfico de dispersión, y calcula la covarianza y el coeficiente de correlación lineal entre las dos variables.