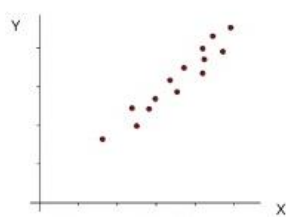


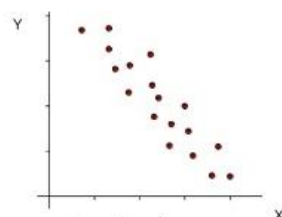
COVARIANZA

La **covarianza** indica el sentido de la correlación entre las variables

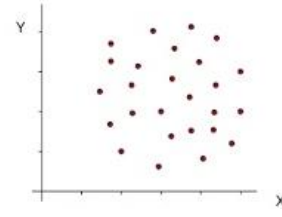
- Si $\sigma_{xy} > 0$ la correlación es positiva (directa)
- Si $\sigma_{xy} < 0$ la correlación es negativa (inversa)
- Si $\sigma_{xy} = 0$ los datos no están relacionados (correlación nula)



Correlación positiva
(o directa)



Correlación negativa
(o inversa)



Correlación nula

Para calcular la covarianza utilizamos la siguiente fórmula:

$$\sigma_{xy} = \frac{\sum_{i=1}^m \sum_{j=1}^n x_i \cdot y_j \cdot f_{ij}}{N} - \bar{X} \cdot \bar{Y}$$

La covarianza presenta como inconveniente, el hecho de que su valor depende de la escala elegida para los ejes (es decir, la covarianza variará si expresamos la altura en metros o en centímetros. También variará si el dinero lo expresamos en euros o en dólares), por eso, para estudiar la correlación de dos variables se suele utilizar también el **coeficiente de correlación**, que veremos en el siguiente apartado.

Ejemplo: (Pocos datos)

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

- Representa los datos en una nube de puntos.
- Indica el tipo de correlación que existe entre ambas variables.
- Calcula la covarianza y observa que la solución obtenida es coherente con el gráfico.

Solución:

a)

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
6	6	36	36	36
1	5	1	25	5
4	5	16	25	20
3	6	9	36	18
5	5	25	25	25
8	10	64	100	80
27	37	151	247	184

- Medias:

$$\bar{x} = \frac{27}{6} = 4,5$$

$$\bar{y} = \frac{37}{6} = 6,17$$

- Covarianza:

$$\sigma_{xy} = \frac{184}{6} = 4,5 \cdot 6,17 = 2,9$$

Ejemplo (muchos datos)

Las calificaciones de 40 alumnos en psicología evolutiva y estadística han sido las de la tabla adjunta.

Psicología x_i	3	4	5	6	6	7	7	8	10
Estadística y_i	2	5	5	6	7	6	7	9	10
Nº de alumnos f_i	4	6	12	4	5	4	2	1	2

- Representa los datos en una nube de puntos.
- Indica el tipo de correlación que existe entre ambas variables.
- Calcula la covarianza y observa que la solución obtenida es coherente con el gráfico.

Solución:

Construimos una tabla con las columnas necesarias

	xi	yi	fi	fi·xi	fi·yi	fi·xi ²	fi·yi ²	fi·xi·yi
	3	2	4	12	8	36	16	24
	4	5	6	24	30	96	150	120
	5	5	12	60	60	300	300	300
	6	6	4	24	24	144	144	144
	6	7	5	30	35	180	245	210
	7	6	4	28	24	196	144	168
	7	7	2	14	14	98	98	98
	8	9	1	8	9	64	81	72
	10	10	2	20	20	200	200	200
Σ			40	220	224	1314	1378	1336

Medias aritméticas

$$\bar{x} = \frac{\sum x_i \cdot f_i}{N} = \frac{220}{40} = 5,5 \quad \bar{y} = \frac{\sum y_i \cdot f_i}{N} = \frac{224}{40} = 5,6$$

Varianzas y desviaciones típicas

$$\sigma_x^2 = \frac{\sum f_i \cdot (x_i)^2}{N} - (\bar{x})^2 = \frac{1314}{40} - (5,5)^2 = 2,60$$

$$\sigma_x = \sqrt{2,6} = 1,61$$

$$\sigma_y^2 = \frac{\sum f_i \cdot (y_i)^2}{N} - (\bar{y})^2 = \frac{1378}{40} - (5,6)^2 = 3,09$$

$$\sigma_y = \sqrt{3,09} = 1,76$$

Covarianza σ_{xy}

$$\sigma_{xy} = \frac{\sum f_i \cdot x_i \cdot y_i}{N} - \bar{x} \cdot \bar{y} = \frac{1336}{40} - (5,5 \cdot 5,6) = 2,6$$

La covarianza es positiva, correlación positiva fuerte.

Ejercicios:

1. Las notas obtenidas por 10 alumnos en Matemáticas y en Música son:

Matemáticas x_i	6	4	8	5	3,5	7	5	10	5	4
Música y_i	6,5	4,5	7	5	4	8	7	10	6	5

- Representa los datos en una nube de puntos.
 - Indica el tipo de correlación que existe entre ambas variables.
 - Calcula la covarianza y observa que la solución obtenida es coherente con el gráfico.
2. La siguiente tabla recoge las medidas de los pesos en kg y las alturas en m de 20 alumnos:

Nº de alumnos	4	3	2	5	4	2
Peso (X) en kg	73	76	73	78	80	82
Altura (Y) en m.	1,65	1,68	1,70	1,72	1,76	1,80

- Representa los datos en una nube de puntos.
 - Indica el tipo de correlación que existe entre ambas variables.
 - Calcula la covarianza y observa que la solución obtenida es coherente con el gráfico.
3. En un curso de Bachillerato de 40 alumnos se ha querido estudiar la correlación de las notas de las asignaturas de Lengua (X) e Historia (Y) como materias clave en itinerarios de letras. Los resultados han sido los siguientes.

X \ Y		3	4	5	6	7	8	10
2		4						
5			7	11				
6					5	3		
7					5	2		
9							1	
10								2

- Representa los datos en una nube de puntos.
- Indica el tipo de correlación que existe entre ambas variables.
- Calcula la covarianza y observa que la solución obtenida es coherente con el gráfico.