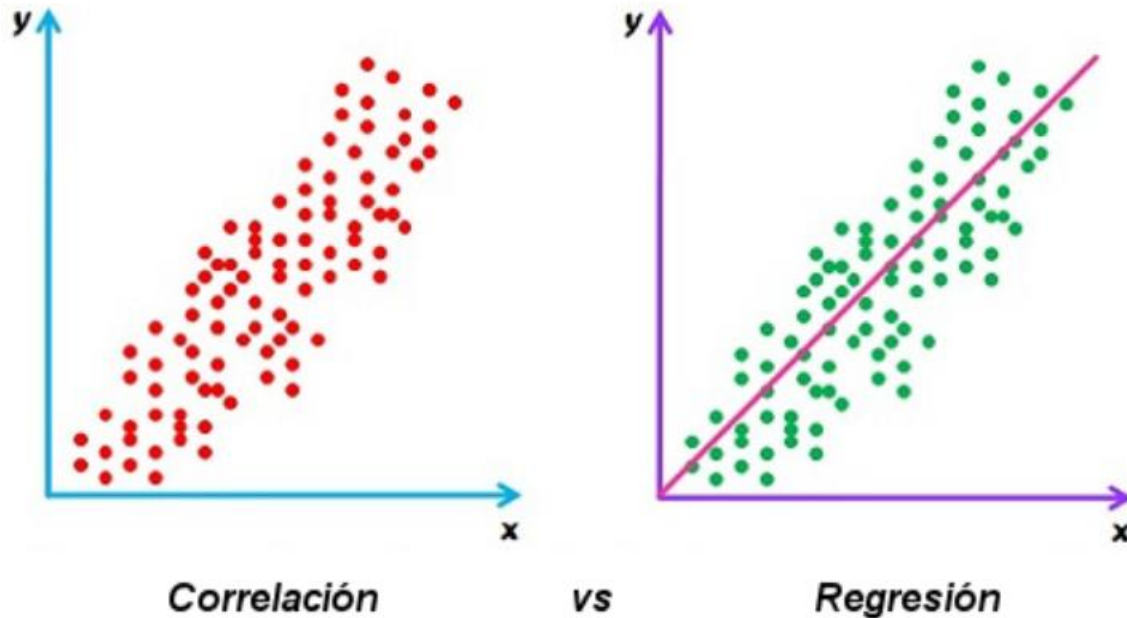


CORRELACIÓN Y REGRESIÓN LINEAL



CORRELACIÓN ENTRE VARIABLES CUANTITATIVAS

Se considera que dos variables cuantitativas están relacionadas entre sí cuando los valores de una de ellas varían de forma sistemática con respecto a los valores homónimos de la otra. Dicho de otro modo, si tenemos dos variables, A y B , existe relación entre ellas si al aumentar los valores de A también lo hacen los de B , o por el contrario si al aumentar los valores de A disminuyen los de B .

- Para variables métricas, el gráfico de dispersión es la manera más sencilla de comprobar la relación entre las dos variables, pudiendo esta adoptar diferentes formas.
- El método más usual para medir la intensidad de la relación lineal entre dos variables métricas es la correlación momento- producto o correlación de Pearson.

DIAGRAMAS DE DISPERSIÓN ESTADÍSTICA

Gráfico de puntos para variables cuantitativas

Disposición:

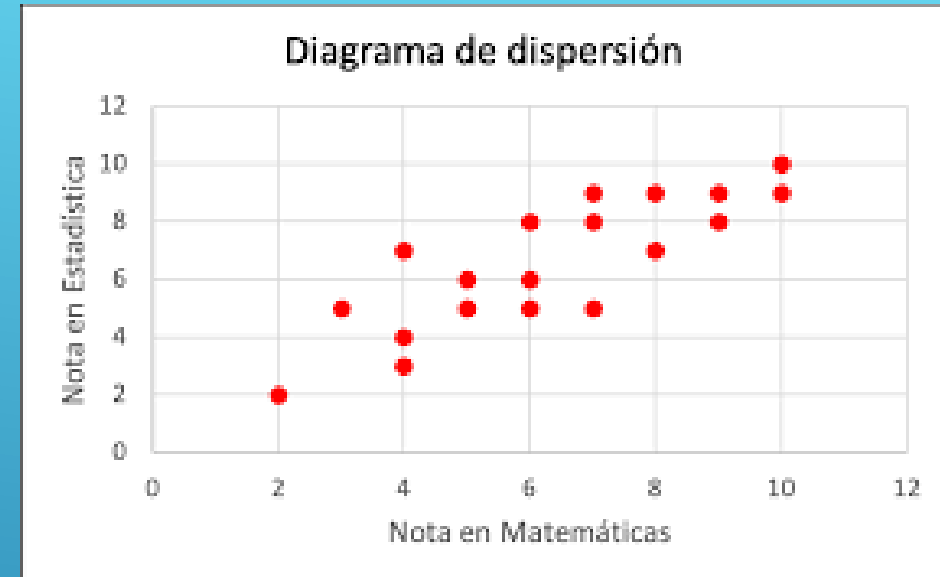
Eje de abscisas: variable independiente (X)

Eje de ordenadas: variable dependiente (Y)

Frecuentemente X es una variable controlada (no aleatoria)

Un punto por cada observación (par de valores X-Y)

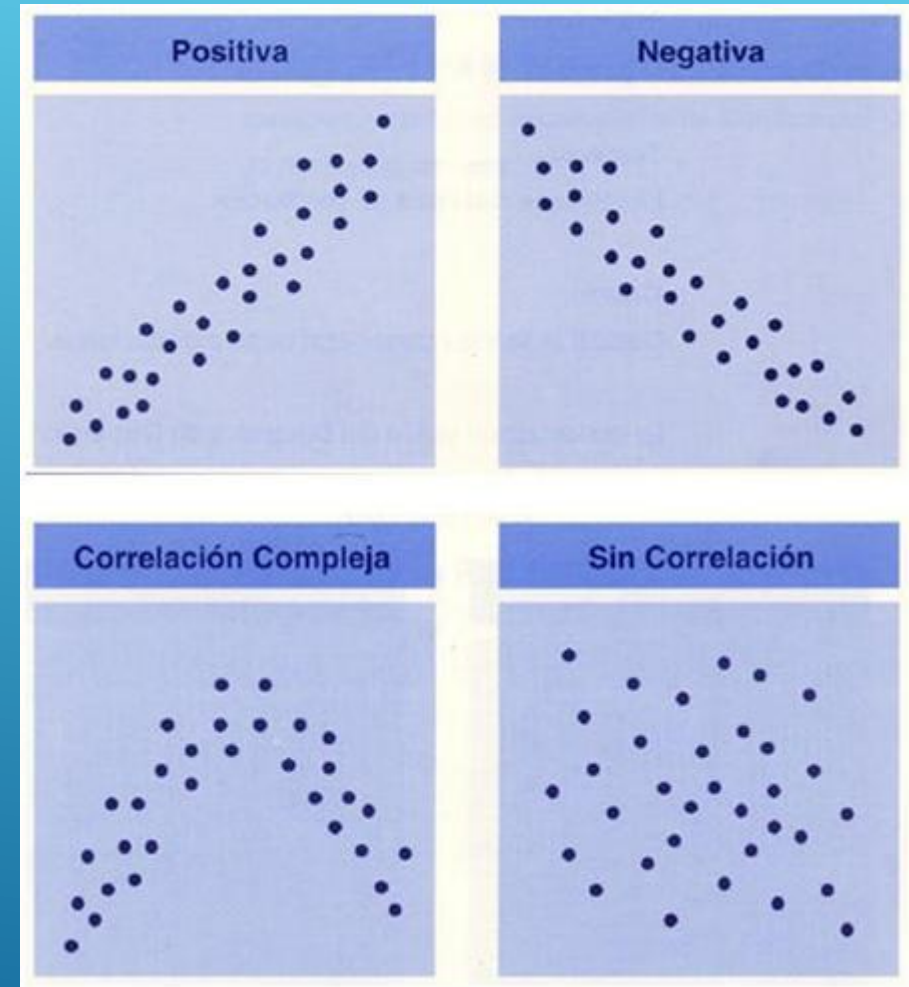
Aproximación al tipo de relación existente entre las variables



CORRELACIÓN ENTRE VARIABLES CUANTITATIVAS

Los componentes fundamentales de una relación entre dos variables cuantitativas son:

La Fuerza
El Sentido
La Forma



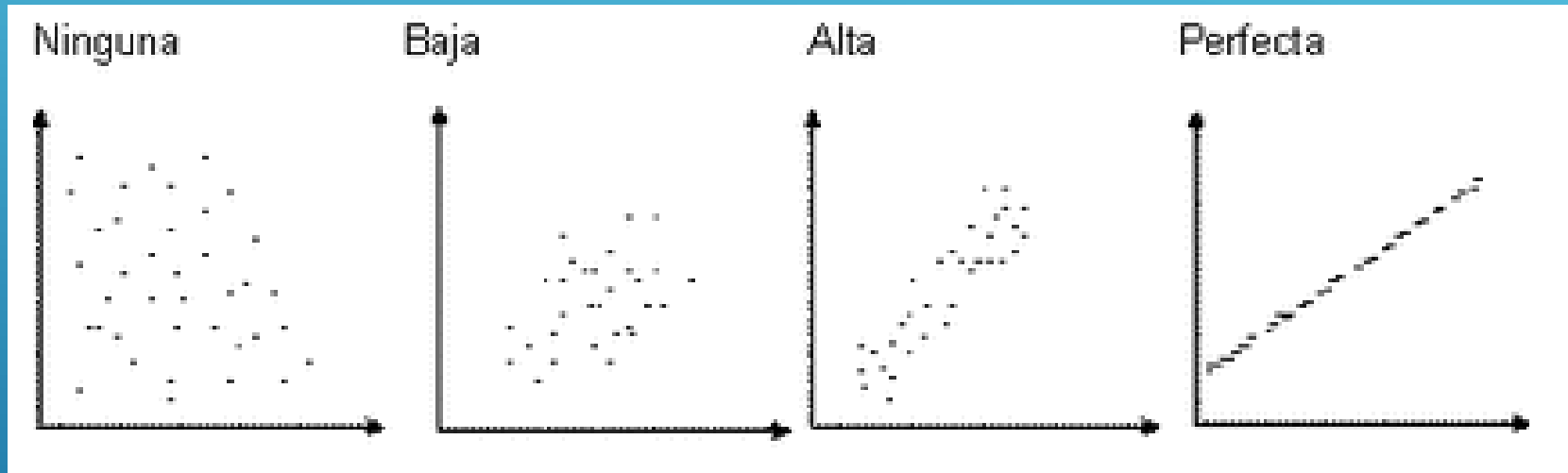
CORRELACIÓN ENTRE VARIABLES CUANTITATIVAS

La fuerza mide el grado en que los pares de observaciones quedan representados en una línea. Si la nube de observaciones es estrecha y alargada, una línea recta representará adecuadamente a la nube de puntos y a la relación y por tanto ésta será fuerte.

El sentido de la relación se refiere a cómo varían los valores de B con respecto a A . Si al crecer los valores de la variable A lo hacen los de B , será una relación positiva o directa. Si al aumentar A , disminuye B , será una relación negativa o inversa.

La forma establece el tipo de línea a emplear para definir el mejor ajuste. Se pueden emplear tres tipos de líneas: una línea recta, una curva monotónica o una curva no monotónica

FORMAS TÍPICAS DE LOS DIAGRAMAS DE DISPERSIÓN ESTADÍSTICA



EL COEFICIENTE DE CORRELACIÓN LINEAL DE PEARSON

El Coeficiente de Correlación Lineal de Pearson es un índice estadístico que permite medir la fuerza de la relación lineal entre dos variables. Su resultado es un valor que fluctúa entre -1 (correlación perfecta de sentido negativo) y $+1$ (correlación perfecta de sentido positivo). Cuanto más cercanos al 0 sean los valores, indican una mayor debilidad de la relación o incluso ausencia de correlación entre las dos variables.

Su cálculo se basa en la expresión:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum (x^2) - (\sum x)^2][n\sum (y^2) - (\sum y)^2]}}$$

Interpretación:

1. El valor absoluto de r indica la fuerza de la relación entre Y y X .
2. El signo la dirección de la relación (directa o inversamente proporcional)

$r = 1$ correlación positiva perfecta.
 $r = -1$ correlación negativa perfecta.
 $r = 0$ no hay relación lineal entre Y y X .

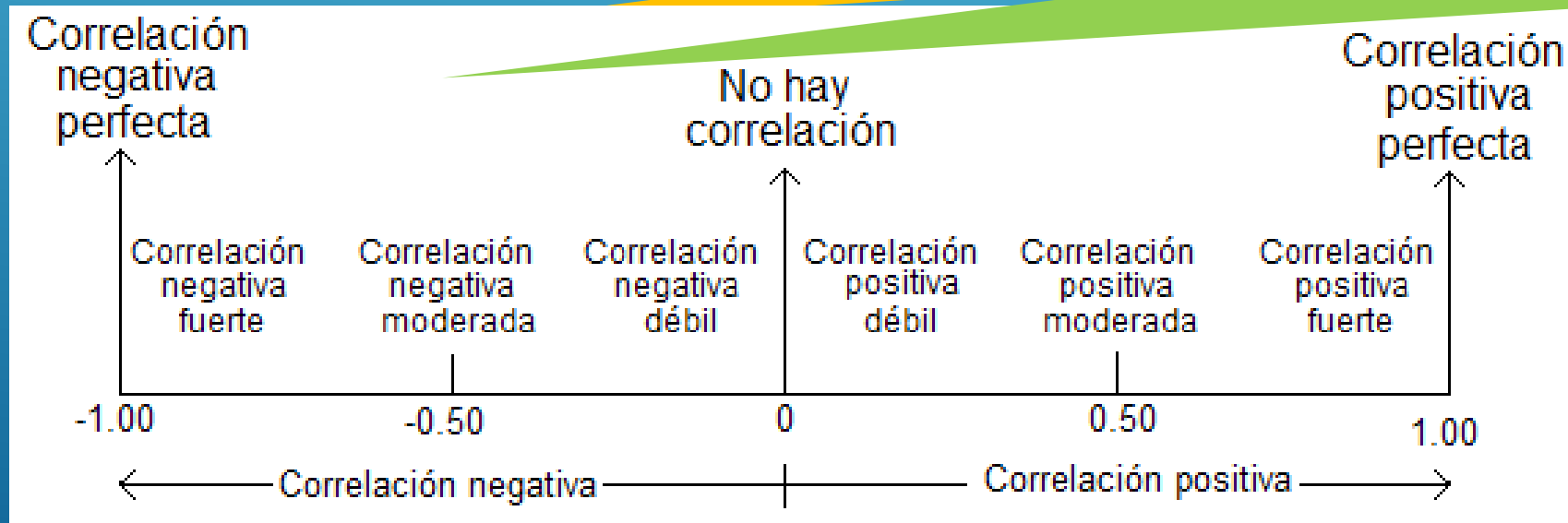
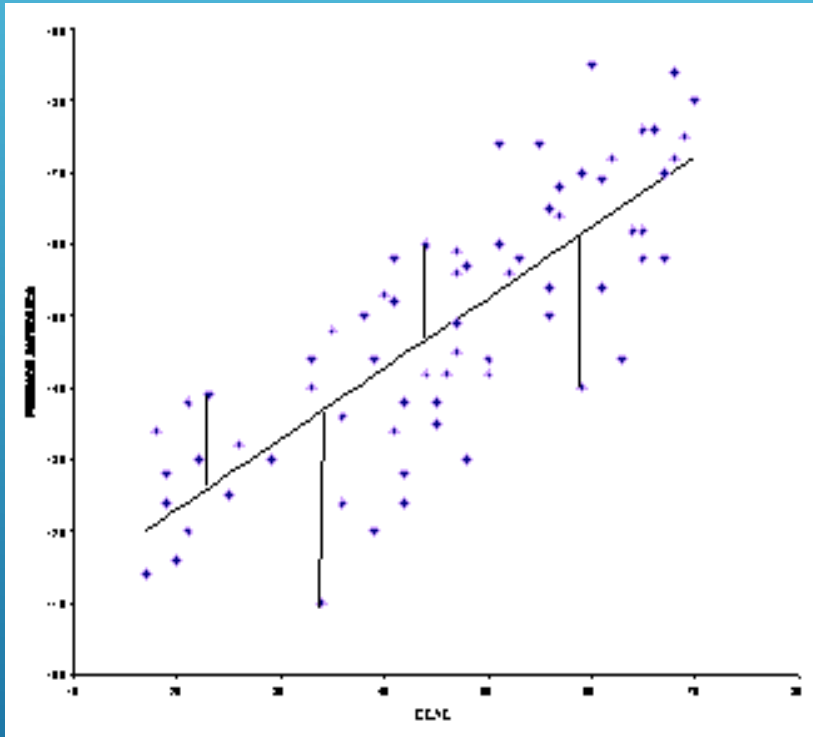


Gráfico de Dispersión y Recta de Regresión Lineal



Para el cálculo de la recta de regresión se aplica el método de mínimos cuadrados entre dos variables. Esta línea es la que hace mínima la suma de los cuadrados de los residuos, es decir, es aquella recta en la que las diferencias elevadas al cuadrado entre los valores calculados por la ecuación de la recta y los valores reales de la serie, son las menores posibles.

$$y = a + bx$$

ECUACIÓN DE REGRESIÓN

La ecuación de regresión o recta de mejor ajuste: $y = a + bx$ predice el comportamiento de los valores y de acuerdo al valor de x ; de acuerdo a los coeficientes a y b .

$$a = \frac{\sum y - b \sum x}{n}$$

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

Coeficiente de Determinación

$R^2 =$ es el cuadrado del coeficiente de correlación de Pearson

Es el porcentaje en que la variabilidad de la variable respuesta está explicada por la variable regresora.