

Conceptos básicos de estadística

Teoría y Estructuración de Portafolios

Docente: Natalia María Acevedo Prins.

Medidas de dispersión y posición de los datos

Estadística descriptiva:

Es la rama de la estadística que recolecta, analiza y caracteriza un conjunto de datos. Trata de extraer conclusiones muy generales del conjunto.

Medidas de tendencia o posición:

Medidas de tendencia central y no central.

- Media
- Mediana
- Moda

Medidas de dispersión:

Cuan dispersos se encuentran los datos, si están centrados o no. Pueden determinar acercamientos al riesgo.

- Alcance interfractil (percentiles)
- Desviación estándar
- Coeficiente de asimetría
- Curtosis.

Medidas de posición o tendencia central

Media aritmética:

La mas intuitiva y conocida de las medidas descriptivas. También llamado Promedio aritmético.

Es útil dentro de los histogramas cuando los datos ya vienen agrupados ya que Excel no tiene una función para estos

$$\text{Promedio} = \frac{\sum_{i=1}^n X_i}{n}$$

Media ponderada:

Promedio que permite asignar pesos a los valores de manera tal que el promedio es calculado considerando la importancia de cada valor sobre el total.

Participación (P_i)	Valor (X_i)
20%	\$15,000
30%	\$8,000
25%	\$25,000
25%	\$3,500

Total: \$12,525

$$\text{media ponderada} = \sum_{i=1}^n X_i P_i$$

Medidas de posición o tendencia central de los datos

Mediana

Valor de una serie de datos que divide a dicho conjunto en dos partes igual. La cantidad de observaciones es la misma hacia arriba y hacia debajo de la mediana.

$$\text{Mediana} = \bar{\bar{x}} = \frac{n + 1}{2}$$

- Tiene la ventaja de no estar influenciada por valores extremos.

Moda

Valor de las observaciones que mas se repite en un conjunto de datos.

Si hay dos valores que sean los que mas se repiten con la misma cantidad se dice que la distribución es “Bimodal”.

No hay moda en rendimientos ya que comúnmente no hay valores que se repitan.

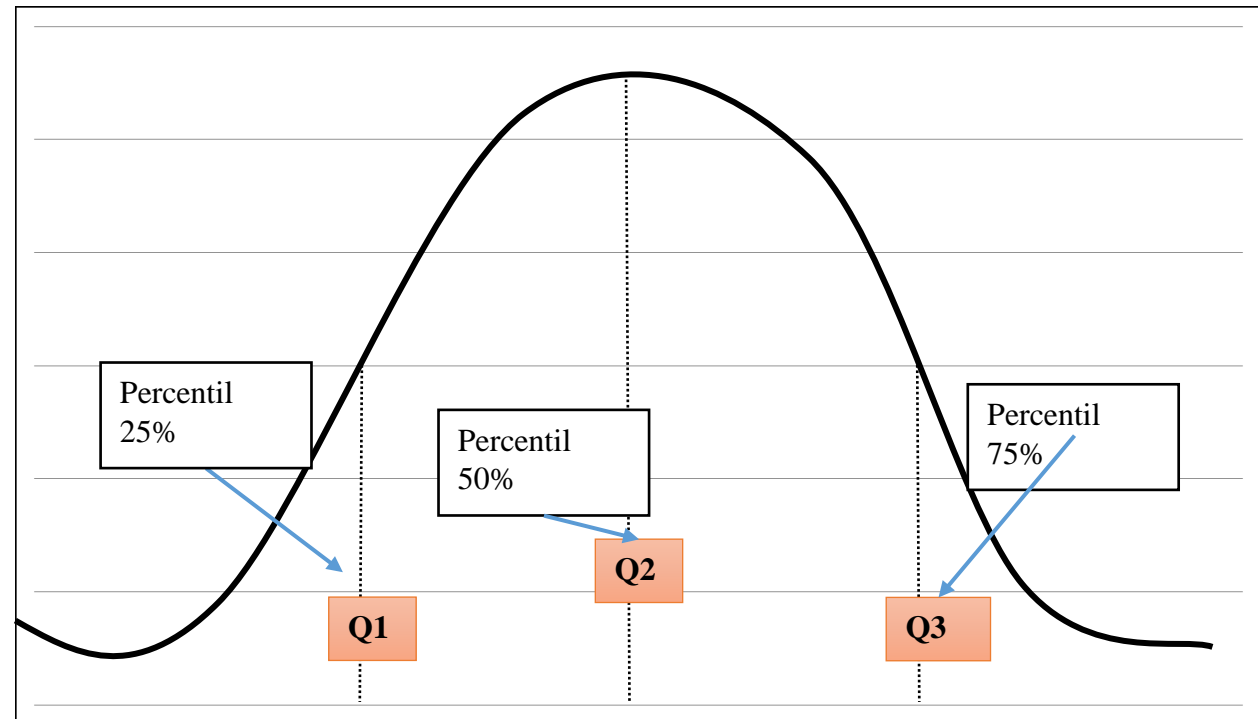
Medidas de dispersión

Alcance Percentil o Cuartiles:

Se construyen con las frecuencias relativas acumuladas de los datos ordenados de forma ascendente y que tomarán valores entre 0% y 100%. 1% corresponderá al percentil 1.

Si los datos se dividieran en 4 partes iguales a cada grupo de probabilidades se le conoce como cuartil

La mediana divide el grupo de datos en 2 partes iguales; el 50% está por encima de la mediana y el otro 50% por debajo. A esta se le conoce como percentil 50%.



Medidas de dispersión

Varianza

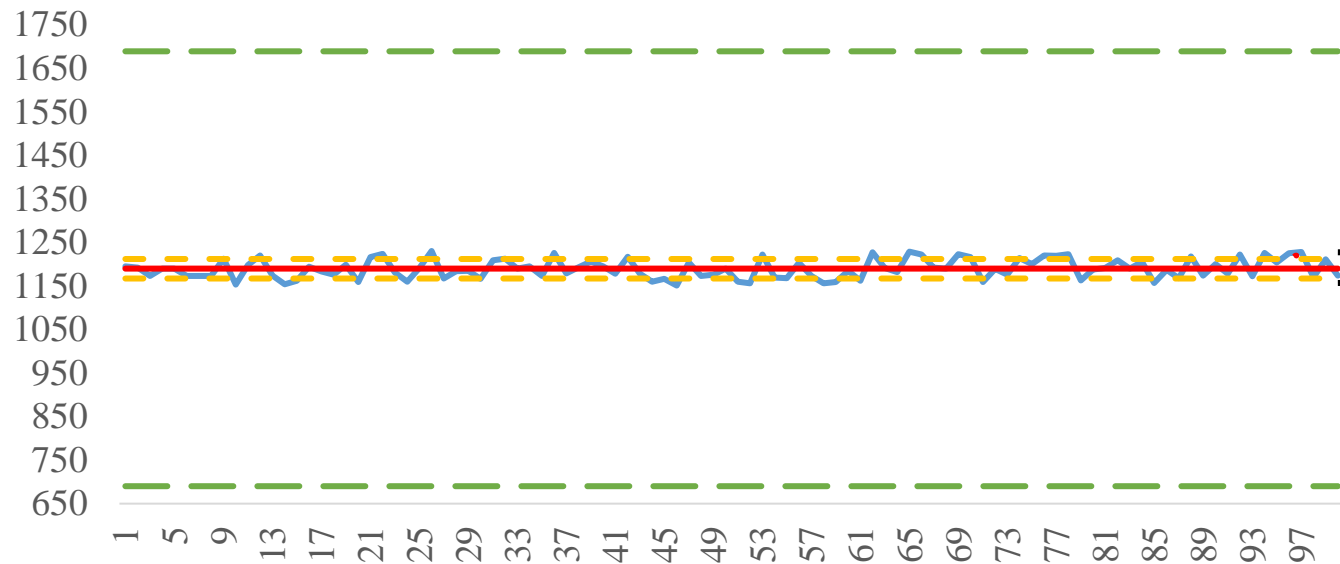
Mide la distancia que existe entre cada uno de los datos y la media aritmética.

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Desviación estándar

Raíz cuadrada de la varianza. Esta es una forma sencilla de explicar la dispersión de los datos.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$



— datos — media • lim inf varianza
• lim sup varianza — lim inf desvest — lim sup desvest

Desviación
estándar

Está dada en términos
absolutos

Varianza

Es difícil de concluir con esta
medida porque altera las
unidades de medida

Medidas de dispersión

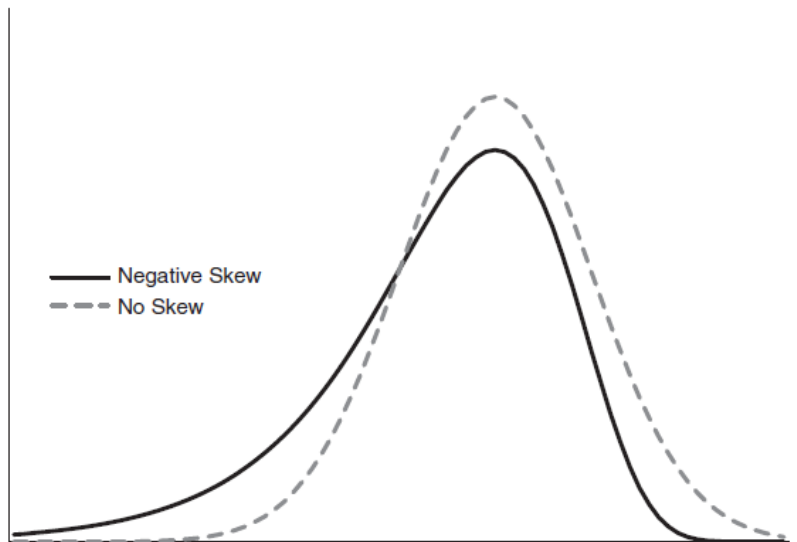
Coeficiente de asimetría o sesgo:

Mide el grado de asimetría de una distribución respecto a su media y si los datos descienden mas rápido por la derecha o por la izquierda.

$$\text{Asimetría} = \frac{n}{(n-1) \times (n-2)} \times \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^3$$

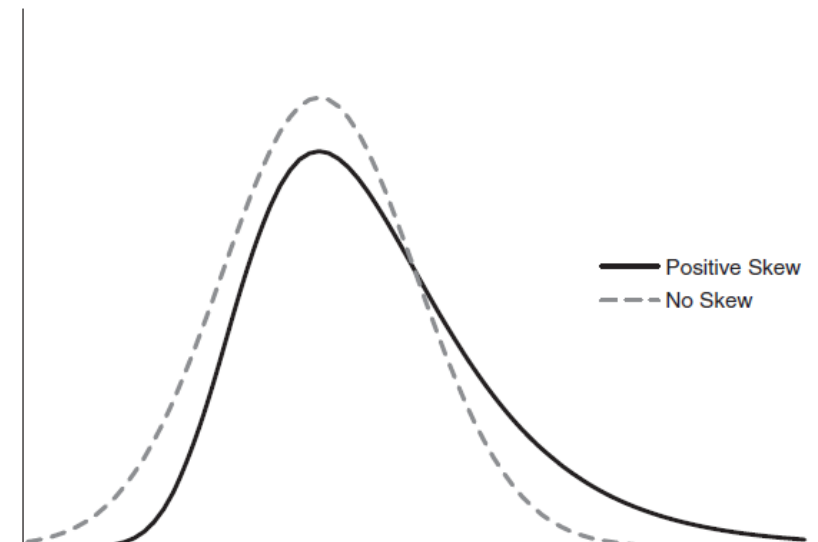
Coeficiente de Asimetría < 0

Negative Skewness



Coeficiente de Asimetría > 0

Positive Skewness

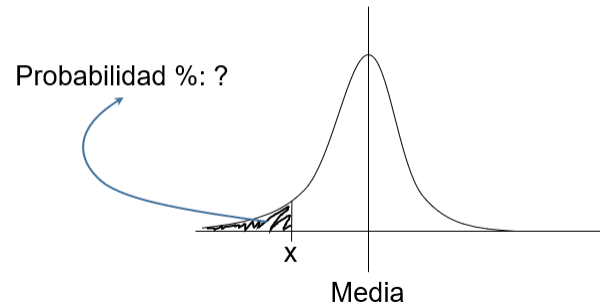
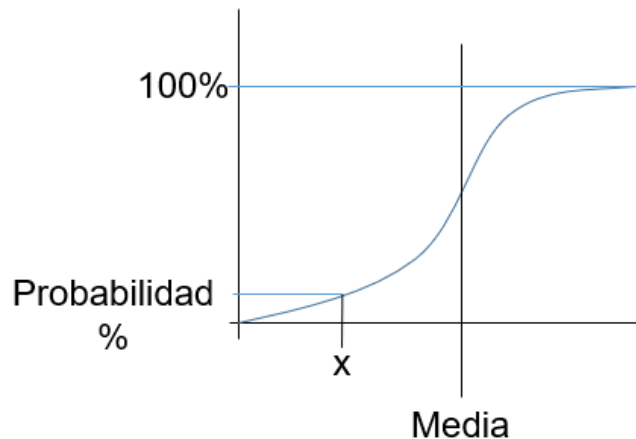


Distribuciones de probabilidad continua

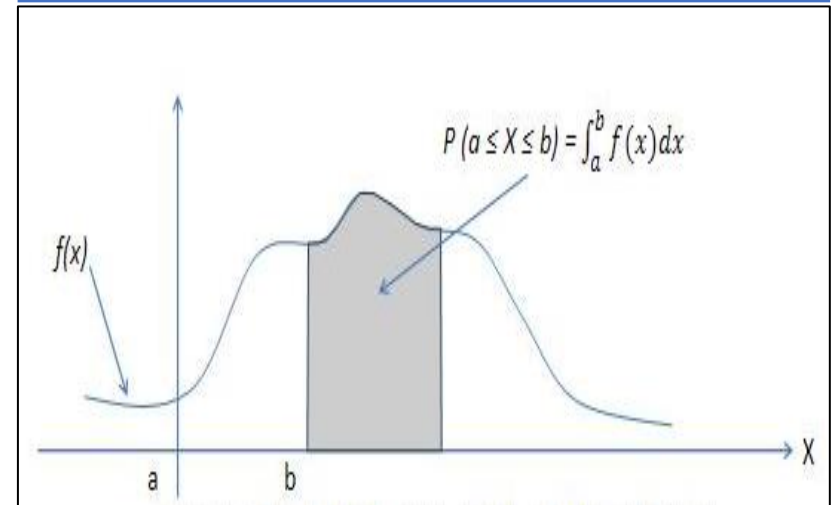
Una función de probabilidad continua no puede ser expresada en forma tabular.

para describir una distribución de probabilidad continua, se usa una ecuación o fórmula llamada:

Función de Densidad de Probabilidad (FDP) o Función de Densidad: $P(x)$



El área bajo la curva de la FDP es igual a 1.



Función de Probabilidad Acumulativa (FPA): es una regla o ecuación que describe la suma de todas las probabilidades.

Distribución normal (Gaussiana)

Variable distribuida normalmente:

Función de Densidad de Probabilidad

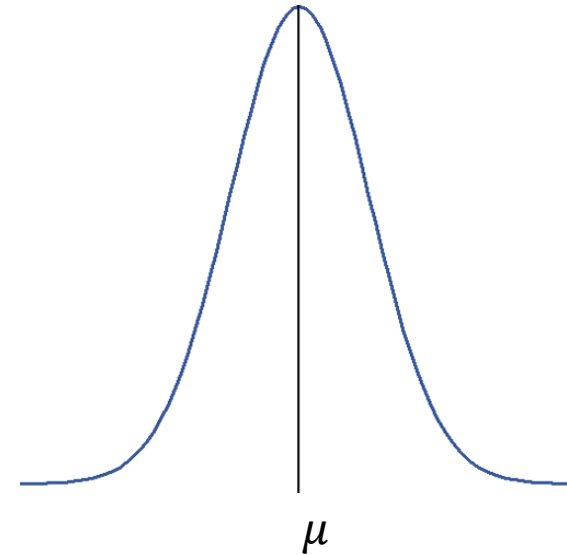
$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]}$$

Probabilidad entre los valores a y b:

$$P(a \leq X \leq b) = \int_a^b P(x) dx$$

Valor esperado: $E[X] = \mu$

Varianza: $VAR[X] = \sigma^2$



Sesgo = 0

Curtosis = 3

Media = Moda = Mediana

$$z \sim N(0, \sigma^2) \quad X_1 \sim N(\mu_1, \mu_1) \text{ y } X_2 \sim N(\mu_2, \mu_2)$$

$$z \sim N(0,1) \quad X = aX_1 + bX_2 \sim \text{Normal}$$

Distribución normal (Gaussiana)

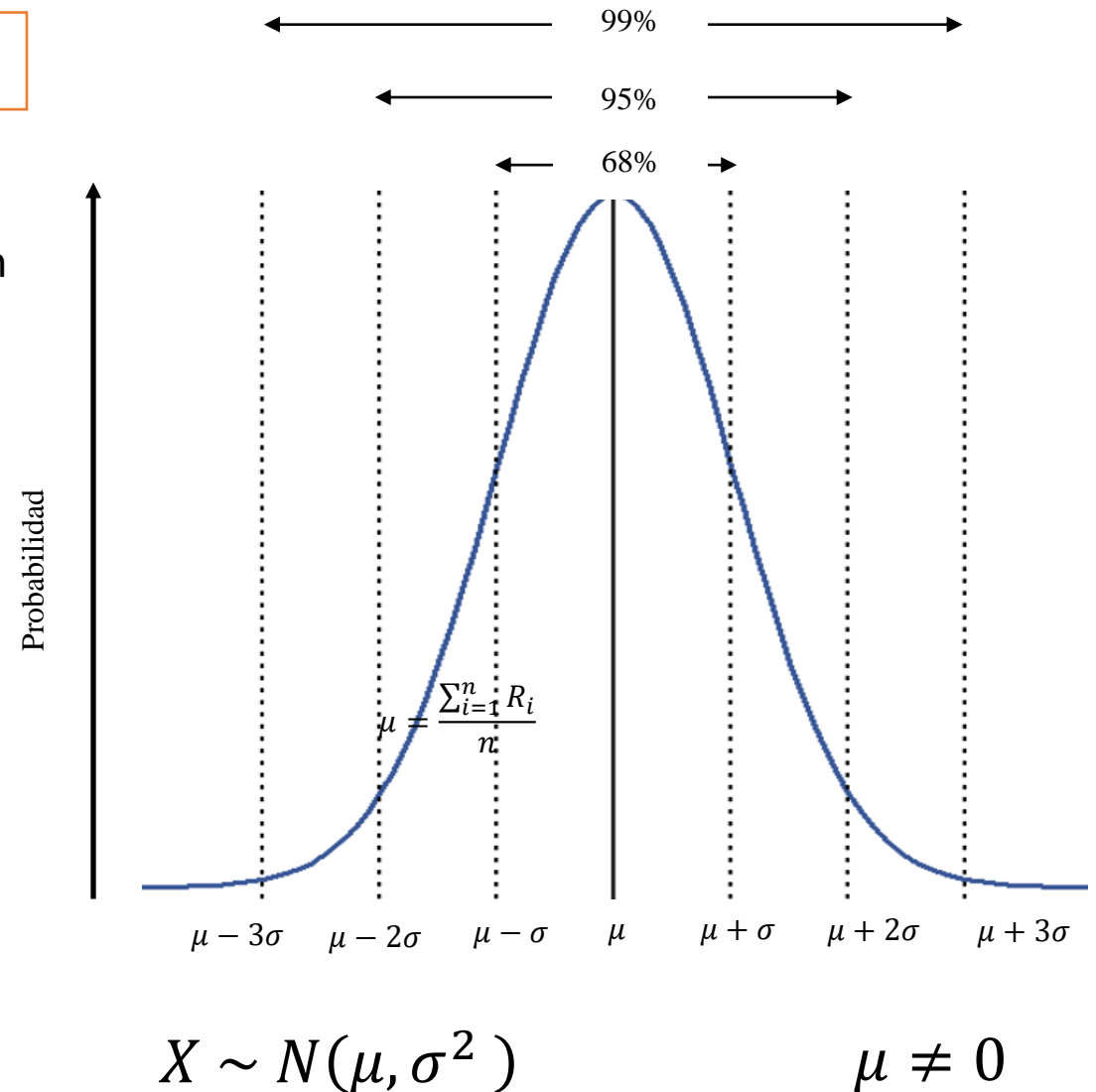
La distribución normal está centrada alrededor de la media: μ

La variación o dispersión alrededor de la media se expresa en unidades de la desviación estándar: σ

En portafolios de inversión, la media es el rendimiento promedio y la desviación estándar se define como volatilidad.

Volatilidad histórica:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (R_i - \mu)^2}{n - 1}}$$



Medidas de dispersión

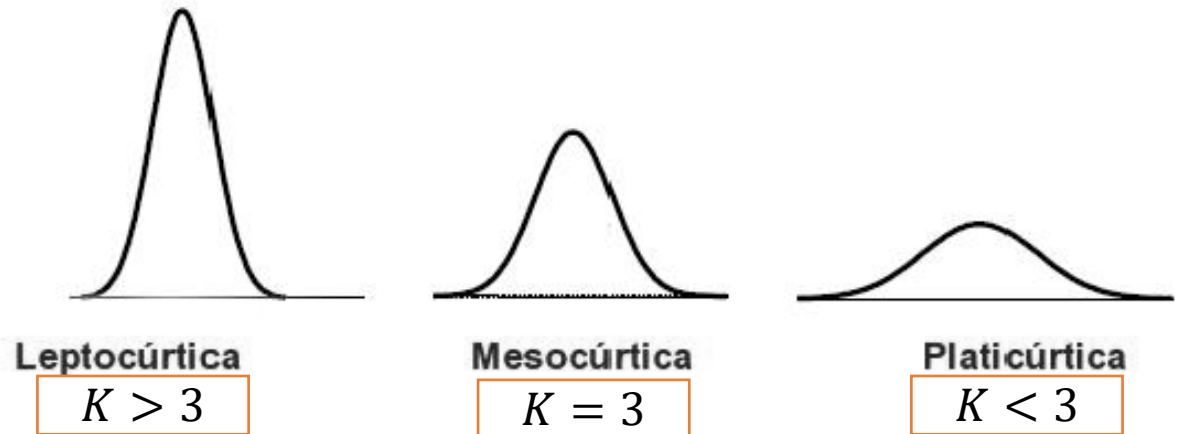
Curtosis

Otra característica adicional de la forma de la distribución es el grado de altura es decir “elevamiento” o “aplanamiento” de los datos.

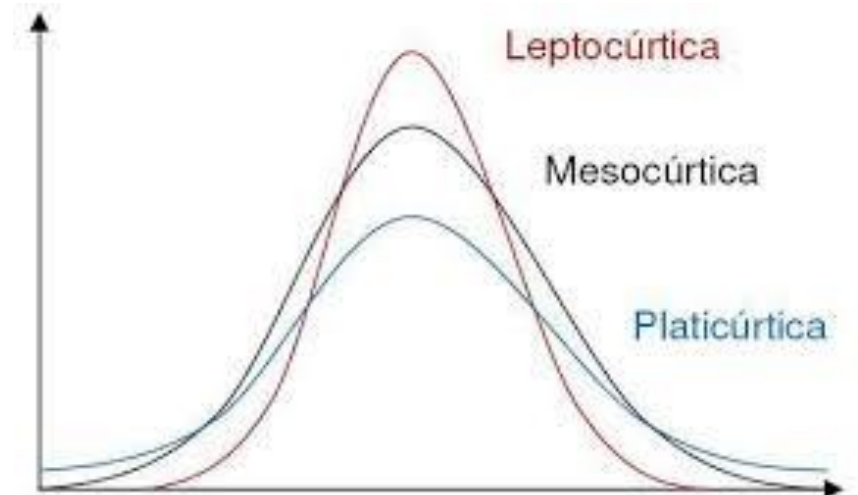
Muchos valores observados en la cola genera una mayor ponderación y por lo tanto crea alta curtosis.

$$K = \frac{\sum_{i=1}^n (x_i - \mu)^4}{\sigma^4}$$

DISTRIBUCIONES



Alta curtosis: Alta o baja probabilidad de valores extremos.



Covarianza y correlación

Son medidas de relación lineal entre dos variables aleatorias describiendo el movimiento conjunto entre éstas.

Mide el grado de movimiento conjunto entre dos variables o la relación lineal entre ambas en un rango entre -1 y +1

$$COV(R_i, R_j) = \frac{1}{n} \sum_{i=1}^n (R_i - \mu_i)(R_j - \mu_j)$$

$COV(R_i, R_j)$: Covarianza entre los activos i y j.

$$Corr(R_i, R_j) = \rho_{ij} = \frac{COV(R_i, R_j)}{\sigma_i \sigma_j}$$

ρ_{ij} : Correlación entre los activos i y j.

σ_i : Desviación estándar del activo i.

σ_j : Desviación estándar del activo j.



$Corr > 0$ las dos variables se mueven en la misma dirección, mientras más cercano a la unidad, mayor el grado de dependencia mutua.



$Corr < 0$ indica que las dos variables se mueven en sentidos opuestos.



Mientras más cercano a 0 el coeficiente de correlación, mayor será la independencia de las variables.

Conceptos básicos de estadística

GRACIAS!!!

Teoría y Estructuración de Portafolios