MODELO CONCEPTUAL PARA EL DISEÑO DE ALGORITMO DE VALIDACIÓN DE PRECIOS

Teorema del límite central

Sea X_1, X_2, ..., X_n un conjunto de variables aleatorias, independientes e idénticamente distribuidas con media **μ** y varianza $\boldsymbol{\sigma^2}$ distinta de cero. Sea

$$S_n = X_1 + \cdots + X_n$$

Entonces

$$\lim_{n\to\infty} \Pr\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le z\right) = \Phi(z)$$

Propiedad

El teorema del límite central garantiza una distribución normal cuando n es suficientemente grande.

Premisas Básicas

Una serie de tiempo de precios, por el teorema del límite central, sigue una distribución normal de media **μ** y de varianza $\boldsymbol{\sigma^2}$. Cuando una serie de precios es demasiado larga, se debe "tratar" los datos a fin de eliminar factores exógenos que modifican los datos como por ejemplo la inflación (componente de tendencia); no hacerlo podríamos obtener distribuciones del tipo chi-cuadrado.

En nuestro caso, asumiremos que los precios siguen una distribución normal, dado que la frecuencia de los datos es, en la mayoría de los casos, diaria. Por lo tanto el factor inflacionario (considerado como componente de tendencia) no es un factor que incidirá significativamente en la distribución dado que se tomarán series no mayor a tres años en un país con niveles de inflación de un dígito.

Diseño de variables de control

La media y la desviación estándar se constituyen en herramientas útiles para discriminar información "extraña"; la misma que pudiera explicarse por errores cometidos en la digitación. La bondad de utilizar la media y la desviación estándar radica en:

a) Que los parámetros obtenidos a partir de ellos son objetivos y,
b) Son de fácil comprensión y disponible en la mayoría de los lenguajes de programación.

La probabilidad de que una observación caiga dentro de un rango puede ser estimada tomando en consideración dos alternativas:

1. Si la variable precios no sigue una distribución normal

Hacer uso del teorema de Chevyshev: La probabilidad de que cualquier variable aleatoria X tome un valor dentro de K desviaciones estándar de la media es al menos de $1-1/k^2$.

$$\boldsymbol{P(\mu - k\sigma} < x < \mu + k\sigma) \geq \mathbf{1 - 1/k^2}$$

Donde,

- **μ** es la media
- **σ** es la desviación estándar
- k>1

Si k=2, entonces la probabilidad de que un valor caiga dentro de rango es del 75%.

2. Si los precios siguen una distribución normal

Se aplica las propiedades de la distribución normal y se obtiene valores de probabilidad para los K dada por su propia función.

a) Una desviación estándar (k=1) la probabilidad es: 68.1%
b) Dos desviaciones (k=2), la probabilidad es: 95.4%
c) Tres desviaciones (k=3), la probabilidad es 99.6%

Los valores de probabilidad para cada valor de K se muestran en el gráfico para una mayor claridad.