



Instituto Politécnico Nacional
Escuela Superior de Física y Matemáticas

MODELACIÓN DEL PRECIO DE ACCIONES USANDO UN MODELO ARIMA PARA LOS RENDIMIENTOS

TESIS

Para obtener el título de
Licenciada en Ingeniería Matemática

PRESENTA

Nancy Rocío Martínez Martínez

DIRECTOR DE TESIS

Dr. Ramón Sebastián Salat Figols

Ciudad de México

Agosto, 2019

CARTA CESIÓN DE DERECHOS AL IPN

Ciudad de México a 2 de Agosto de 2019

El que suscribe:

C. Nancy Rocío Martínez Martínez con número de boleta 2014330373

Egresado del Programa Académico: Licenciatura en Ingeniería Matemática

que se imparte en la Escuela Superior de Física y Matemáticas, manifiesta que es autor intelectual del presente trabajo de tesis titulado:

Modelación del Precio de Acciones usando un modelo ARIMA para los rendimientos

El cual fue desarrollado bajo la dirección de: Dr. Ramón Sebastián Salat Figals

Por lo anterior es mi interés manifestar que **SI** cedo los derechos del trabajo antes mencionado, al Instituto Politécnico Nacional, ello con el propósito de que lo ponga a disposición de la comunidad politécnica que requiera consultarlo con fines académicos y de investigación. Es importante aclarar que los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y/o director del trabajo. Dicho permiso puede ser solicitado a la siguiente dirección de correo electrónico: nancyrocioamm@gmail.com. Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

Atentamente

Nancy Rocío

Nancy Rocío Martínez Martínez
Nombre y firma del alumno

DEDICATORIA

Este proyecto va dedicado

A mis padres por ser el pilar fundamental en todo lo que soy.

A mis hermanas por su apoyo incondicional.

A mis amigos que siempre están para apoyarme en todo momento.

A mi director de tesis por brindarme gratos momentos y hacer que no fuera tan pesado el proceso de elaboración de este proyecto.

AGRADECIMIENTO

Agradezco a mis padres Raúl y Rocío porque han sido mi gran apoyo, quienes me han dado el valor para seguir adelante después de todas las adversidades y gracias a la educación y valores que me han brindado me servirá para el desarrollo de mi vida personal y profesional.

A mis hermanas Jessy e Iris por su apoyo y compañía durante los momentos más importantes y también los difíciles y que me ayudan a ser una mejor persona. Y a mi sobrino porque gracias a sus locuras y juegos me ayudaba a relajarme.

Al profesor Ramón Sebastián Salat Figols, director de tesis por guiarme en este proceso y brindarme más de sus conocimientos.

A mis sinodales, por el apoyo en la elaboración de este proyecto.

A Aquetzalli, Eric, Fran, Jessy, Itzel y Daniel por brindarme su grandiosa amistad y porque siempre estuvieron pendientes y apoyándome para seguir con este proyecto.

A mis demás amigos con quienes compartí gratas experiencias y aprendizajes y que siempre paso agradables momentos junto a ellos.

A mis profesores quienes dieron lo mejor de sí para transmitirme sus conocimientos.

Al Instituto Politécnico Nacional y a la Escuela Superior de Física y Matemáticas por dejarme formar parte de esta hermosa institución y brindarme los conocimientos para desarrollarme profesionalmente.

Agradezco a todos porque sin ellos no hubiera sido posible realizar este proyecto.

ÍNDICE

	Pág.
Introducción.....	1
Resumen.....	3
Capítulo 1. Marco de referencia	
Planteamiento del problema.....	4
Objetivos.....	4
Capítulo 2. Marco teórico	
2.1 Conceptos de probabilidad	
Experimento aleatorio o fenómeno aleatorio.....	5
Espacio muestral.....	5
σ -álgebra de conjuntos.....	5
Evento.....	5
Espacio de probabilidad.....	5
Probabilidad.....	5
Variable aleatoria.....	6
Variable aleatoria discreta.....	6
Variable aleatoria continua.....	6
Función de densidad de probabilidad de una variable aleatoria discreta.....	7
Función de densidad de probabilidad de una variable aleatoria continua	7
Función de distribución.....	8
Función de distribución acumulada de una variable aleatoria discreta...	8
Función de distribución acumulada de una variable aleatoria continua..	8
Valor esperado.....	9
Varianza.....	10
Función generadora de momentos.....	11
Covarianza.....	13
Coeficiente de correlación.....	14
DISTRIBUCIONES	
Distribución normal.....	14
Distribución t-Student.....	16
2.2 Conceptos de estadística	
Teorema del límite central.....	19
Método de máxima verosimilitud.....	19
Hipótesis.....	19
Pruebas de hipótesis.....	20
Serie de tiempo.....	21
Proceso estocástico.....	22
Proceso estocástico estacionario.....	22
Ruido blanco.....	23
Procesos lineales estacionarios.....	23
Procesos Autoregresivos (AR).....	23
Procesos de Medias Móviles (MA).....	23
Proceso autoregresivo de Medias Móviles (ARMA).....	24

	Pág.
2.3 Conceptos financieros	
Acciones.....	25
Rendimiento de una acción.....	25
2.4 Conceptos estocásticos	
Caminata aleatoria.....	26
Movimiento browniano.....	27
Lema de Itô.....	28
Precio de una acción.....	30
Proyección del precio de una acción siguiendo una normal.....	33
2.5 Conceptos de programación	
Lenguaje de programación R.....	35
RStudio.....	37
Prueba de Shapiro-Wilk.....	37
Prueba de Kolmogorov-Smirnov.....	38
Prueba de Ljung-Box.....	39
Prueba de Dickey-Fuller.....	39
Función auto.arima.....	41
Criterio de selección del modelo.....	42
Akaike Information Criterion (AIC).....	42
Bayesian Information Criterion (BIC).....	43
Pronóstico, Función forecast.....	44
Capítulo 3 Marco metodológico	
Procedimiento para la realización del proyecto.....	47
Capítulo 4 Resultados	
4.1 Introducción.....	52
4.2 Resultados numéricos.....	54
4.3 Gráfica de los rendimientos de las acciones.....	66
4.4 Histograma de los rendimientos de las acciones.....	68
4.5 Predicciones de las acciones.....	70
Histograma de las predicciones usando la t-Student.....	70
Predicción usando la función forecast.....	73
4.6 Comparación de precios reales vs precios simulados utilizando la t-Student y la normal.....	75
Conclusiones.....	77
Referencias.....	78

ÍNDICE DE TABLAS Y FIGURAS

Figuras

	Pág.
Figura 1. Concepto de variable aleatoria.....	6
Figura 2. Área bajo la curva de la función de densidad de una variable aleatoria continua.....	8
Figura 3. Curva de la distribución normal.....	15
Figura 4. Función de densidad de las distribuciones Normal y la t de Student.....	17
Figura 5. Diferencia entre las funciones de densidad de las distribuciones Normal y la t de Student.....	18

Tablas

	Pág.
Tabla 1. Tipos de errores.....	20
Tabla 2. Resultados prueba de normalidad.....	52
Tabla 3. Resultados prueba de independencia.....	53

INTRODUCCIÓN

Uno de los principales temas dentro de las matemáticas financieras es el de modelar y realizar pronósticos de precios de acciones; para realizar estos pronósticos se necesita tener en primera instancia, un modelo para estudiarlo y observar el comportamiento del precio y así tratar de lograr un pronóstico lo más preciso posible.

En el presente trabajo se realizará un análisis de algunos precios de las acciones elegidas al azar que cotizan en la Bolsa Mexicana de Valores, para ver si siguen un movimiento Browniano geométrico; de no ser así, se plantearán otros modelos similares que se ajusten mejor a los datos, y ya que se tengan el modelo modificado, se podrá realizar un pronóstico mejor.

Primero, se elegirán 4 acciones al azar, de ahí se analizará cada una para ver si los rendimientos siguen o no una distribución normal, si son independientes y si son estacionarios; esto, para analizar si siguen un movimiento Browniano geométrico; si lo siguen, se procede a realizar un pronóstico del precio de la acción en el futuro; en caso contrario, se plantea otro modelo el cual se ajuste mejor a los precios; con la finalidad de que se puedan hacer mejores proyecciones y reducir los errores de predicción, es decir, disminuirlos.

Para saber si los rendimientos siguen una distribución normal se utilizan las pruebas de Shapiro-Wilk y la de Kolmogorov-Smirnov, con el fin de ver si se puede rechazar la hipótesis nula de que siga una normal con un nivel de significancia de 0.05; para saber si son independientes y que no existe autocorrelaciones entre los diferentes plazos se utiliza la prueba de Ljung-Box con un nivel de significancia de 0.05 y para ver si son estacionarios se utiliza la prueba de Dickey-Fuller donde aquí el p-value tiene que ser menor a 0.05 o con un nivel de significancia del 5% para poder rechazar que haya una raíz unitaria y que así sea estacionaria; si los rendimientos indican con estas 3 pruebas que siguen un movimiento Browniano geométrico, se prosigue a realizar un pronóstico de los precios; en caso contrario, se encuentra un modelo que se ajuste a los rendimientos y después se usa la función auto.arima sobre los rendimientos para saber qué ARIMA se ajusta mejor a los datos, y por último se analizan los residuos del auto.arima para saber si siguen una distribución normal o una t-Student y de ahí realizar los pronósticos.

Se realizará un histograma de los rendimientos y se comparará con la distribución normal y la distribución t-Student para ver cuál de las dos distribuciones se ajusta mejor a los rendimientos.

De igual manera se hará un diagnóstico sobre los residuos de un modelo ARIMA ajustado, esto para ver si los residuos son independientes y estacionarios; también en el diagnóstico se verá la función de autocorrelación, para verificar si existe

autocorrelación entre los rendimientos, realizando un análisis de los p-value de la prueba de Ljung-Box.

Ya una vez que se sabe si los rendimientos siguen más una distribución normal o una distribución t-Student, se realizará un pronóstico de precios, según la distribución que siga cada acción.

Por último, se hace una comparación de los precios reales contra los simulados, esto para ver qué tan preciso es el modelo planteado.

RESUMEN

Los modelos econométricos son importantes hoy en día para elaborar modelos estadísticos, ya que en muchas instituciones se requiere hacer pronósticos de sus acciones o para ver el rendimiento que tiene dicha institución.

En este proyecto se ve una metodología para determinar si el comportamiento de los rendimientos sigue una caminata aleatoria; para esto, se utiliza la metodología de movimiento Browniano geométrico y la de la distribución t-Student. Se utilizarán algunas librerías y funciones que contiene el lenguaje de programación R, entre las librerías están *stats* para estadística, *tseries* para series de tiempo, *LambertW* para estimar los parámetros de la distribución t-Student y *forecast* para realizar pronósticos. Entre algunas funciones que se utilizaron están: *auto.arima* para ver qué ARIMA es mejor para cada modelo; *shapiro.test* para ver la normalidad de los rendimientos; y *Box.test* para que no tengan autocorrelación. Todo esto para ver si siguen un movimiento Browniano geométrico; de no ser así, se plantearán otros modelos similares que se ajusten mejor a los datos y se realizará un pronóstico de cada acción, con el mejor modelo.

CAPÍTULO 1

MARCO DE REFERENCIA

PLANTEAMIENTO DEL PROBLEMA

La mayor parte de los empresarios quieren hacer proyecciones en el tiempo de algunos precios de acciones para ver la rentabilidad financiera que tienen durante su vida útil y así saber cómo podrán invertir en el mercado.

El problema de la proyección en cualquier campo es una tarea compleja que exige en algunas ocasiones la utilización de las técnicas estadístico-económicas complicadas.

Para hacer las proyecciones lo que se usa con frecuencia es el movimiento Browniano geométrico para modelar el precio de una acción; pero muchas veces los rendimientos no siguen dicho movimiento o este no modela bien los precios de las acciones, entonces se necesita una metodología para ver qué modelo siguen y aplicarlo.

OBJETIVOS

En esta sección se definirán los objetivos del proyecto

- Utilizar la función `auto.arima` para ver si los rendimientos siguen un $ARIMA(0,0,0)$. Si se rechaza la hipótesis nula de que siga un $ARIMA(0,0,0)$ pensamos que no lo sigue, por el contrario, si no podemos rechazar no hay razones suficientes para rechazar que siguen un movimiento browniano geométrico.
- Plantear un modelo alternativo cuando el precio de la acción no sigue un movimiento browniano geométrico.
- Ver si los errores siguen una distribución normal o no y se harán los cambios al modelo de acuerdo con la distribución que sigan los errores.

CAPÍTULO 2

MARCO TEÓRICO

2.1 CONCEPTOS DE PROBABILIDAD

Experimento aleatorio o fenómeno aleatorio

Se llamará experimento aleatorio o fenómeno aleatorio a una observación repetida en un conjunto específico de condiciones donde no siempre conduce al mismo resultado, un ejemplo de esto es al lanzar una moneda 100 veces.

Espacio muestral

Es el conjunto de todos los posibles resultados de un experimento aleatorio y está denotado por Ω .

σ -álgebra de conjuntos

Una colección no vacía de subconjuntos \mathcal{A} de un conjunto Ω es llamada σ -álgebra de subconjuntos de Ω siempre que se cumplan las siguientes 3 propiedades:

1. $\Omega \in \mathcal{A}$

2. Si $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

3. Si $A_n, n = 1, 2, \dots$ es una sucesión de elementos de $\mathcal{A} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$

Evento

Se le llama evento a un conjunto A en \mathcal{A} , donde \mathcal{A} es una colección no vacía de subconjuntos de Ω .

Espacio de probabilidad

Un espacio de probabilidad, denotado por (Ω, \mathcal{A}, P) , es un conjunto Ω , una σ -álgebra de subconjuntos de \mathcal{A} y una medida de probabilidad P definido en \mathcal{A} .

Probabilidad

Una función de probabilidad P en una σ -álgebra de subconjuntos de \mathcal{A} de un conjunto Ω es una función de valores reales con dominio \mathcal{A} que satisface las siguientes propiedades:

i) $P(\Omega) = 1$.

ii) $P(A) \geq 0$ para todo $A \in \mathcal{A}$.

iii) $0 \leq P(A) \leq 1$.

iv) Si $A_n, n = 1, 2, 3, \dots$ son conjuntos mutuamente disjuntos en \mathcal{A} , entonces

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Variable aleatoria

Una variable aleatoria X sobre un espacio de probabilidad (Ω, \mathcal{A}, P) es una función real valuada $X(\omega), \omega \in \Omega$, tal que para $-\infty < x < \infty$, $\{\omega | X(\omega) \leq x\}$ es un evento.

Se denotan comúnmente por las letras mayúsculas, y los valores que toman se denotan por letras minúsculas. Puede ser discreta o continua; la variable aleatoria discreta toma un conjunto numerable de valores (valores enteros finitos o contables) y la continua toma un conjunto no numerable de valores (valores reales). Se ilustra de manera gráfica la definición de variable aleatoria:

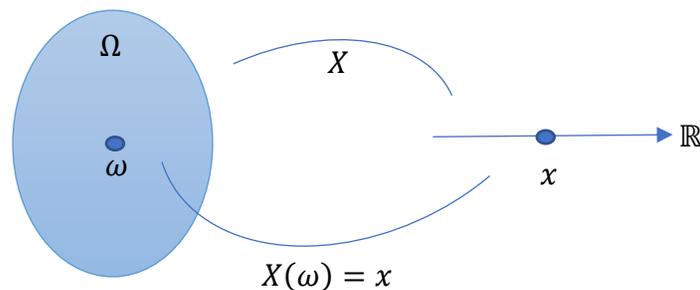


Figura 1. Concepto de variable aleatoria. Ω es el espacio muestral de ε y x es el número real.

Variable aleatoria discreta

Una variable aleatoria discreta X de valores reales en un espacio de probabilidad (Ω, \mathcal{A}, P) , es una función X con dominio Ω y rango un subconjunto finito o contable infinito $\{x_1, x_2, \dots\}$ de los números reales \mathbb{R} , de manera que $\{\omega : X(\omega) = x_i\}$ es un evento para todo i .

Usualmente se escribe el evento $\{\omega : X(\omega) = x_i\}$ como $\{X = x_i\}$ y se denota la probabilidad de este evento como $P(X = x_i)$ en lugar de $P(\{\omega : X(\omega) = x_i\})$.

Variable aleatoria continua

Se puede definir una variable aleatoria continua X sobre un espacio de probabilidad Ω como una función $X(\omega), \omega \in \Omega$, tal que:

$$P(\{\omega \mid X(\omega) = x\}) = 0, \quad \text{o también } P(X = x) = 0, \quad -\infty < x < \infty$$

Es decir, de tal manera que X toma cualquier valor específico x con probabilidad cero.

Denotar mediciones de cantidades físicas tales como peso, tiempo, temperatura se describen como variables aleatorias continuas.

Función de densidad de probabilidad de una variable aleatoria discreta

La función de valor real f definida en \mathbb{R} por $f(x) = P(X = x)$ se denomina función de densidad discreta de X . Un número x se llama un valor posible de X si $f(x) > 0$.

Si es necesario, se denota la función de densidad de X por f_X para mostrar que es la función de densidad para la variable aleatoria X .

La densidad f de una variable aleatoria discreta X tiene las siguientes tres importantes propiedades:

i) $f(x) \geq 0, x \in \mathbb{R}$.

ii) $\{x: f(x) \neq 0\}$ es un subconjunto finito o infinito contable de \mathbb{R} .

Deje $\{x_1, x_2, \dots\}$ denotar este conjunto. Entonces

iii)
$$\sum_{i=1}^{\infty} f(x_i) = 1$$

Función de densidad de probabilidad de una variable aleatoria continua

Una función de densidad (con respecto a la integración) es una función no negativa f tal que:

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Note que si f es una función de densidad, entonces

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

donde $P(a \leq x \leq b)$ significa la probabilidad de que X se encuentre en el intervalo a a b .

O de forma más general,

$$P(X \in A) = \int_A f(x)dx$$

Si A es una unión finita o infinita contable de intervalos disjuntos.

Por lo tanto $P(X \in A)$ se puede representar como el área bajo la curva f como x rangos sobre el conjunto A , la gráfica se muestra en la figura 2

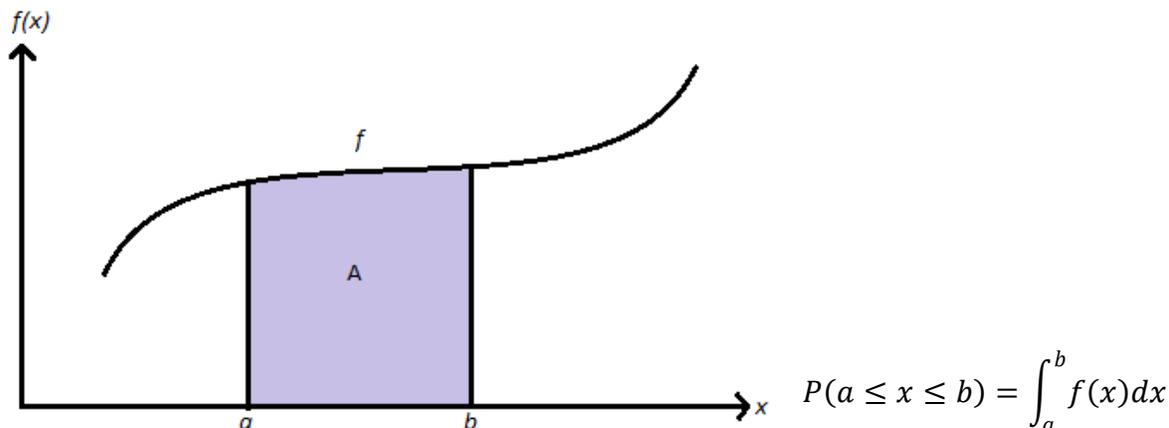


Figura 2. Área bajo la curva de la función de densidad de una variable aleatoria continua

Función de distribución

La función de distribución F de una variable aleatoria X es la función

$$F(x) = P(X \leq x), \quad -\infty < x < \infty$$

Función de distribución de una variable aleatoria discreta

La función $F(x)$, $-\infty < x < \infty$, definida por:

$$F(x) = P(X \leq x) = \sum_{y \leq x} f_X(y) \quad \text{para } -\infty < x < \infty$$

Se denomina la función de distribución de la variable aleatoria X o de la densidad f . Se deduce inmediatamente de la definición de la función de distribución que

$$P(a \leq x \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$

Función de distribución de una variable aleatoria continua

Sea X una variable aleatoria continua, la función

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt \quad \text{para } -\infty < x < \infty$$

Donde $f(t)$ representa el valor de la función de densidad de probabilidad de X en t , se denomina la función acumulada de X .

Una función de distribución es cualquier función F satisfaciendo las siguientes propiedades:

i) $0 \leq F(x) \leq 1$ para todo x .

ii) F es una función no decreciente de x . Si $a < b$, entonces $F(a) \leq F(b) \quad \forall a, b \in R$.

iii) $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$.

iv) $F(\infty) = \lim_{x \rightarrow \infty} F(x) = 1$.

v) $F(x+) = F(x)$ para todo x .

Note que $F'(x) = f(x)$

Valor esperado o media de una variable aleatoria

El valor esperado de una variable aleatoria discreta X , denotado por $E(X)$, se define

$$E(X) = \sum_x xf(x)$$

Donde \sum_x significa la suma sobre todos los valores de X y $f(x)$ es la función de densidad de probabilidad (discreta) de X .

Para una variable aleatoria continua, el valor esperado se define como

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx$$

Propiedades de los valores esperados

1. El valor esperado de una constante es la constante misma. Así, si c es una constante, $E(c) = c$.
2. Sea c una constante, $E(cX) = cE(X)$
3. Si a y b son constantes,

$$E(aX + b) = aE(X) + b$$

4. Sean X y Y variables aleatorias,

$$E(X + Y) = E(X) + E(Y)$$

5. Si X y Y son variables aleatorias independientes,

$$E(XY) = E(X)E(Y)$$

Es decir, la esperanza del producto XY es el producto de las esperanzas individuales de X y Y .

6. Si X es una variable aleatoria con función de densidad de probabilidad $f(x)$ y si $g(X)$ es cualquier función de X , entonces

$$\begin{aligned} E[g(X)] &= \sum_x g(X)f(x) && \text{si } X \text{ es discreta.} \\ &= \int_{-\infty}^{\infty} g(X)f(x)dx && \text{si } X \text{ es continua} \end{aligned}$$

Por tanto, si $g(X) = X^k$ con $k = 1, 2, \dots$,

$$\begin{aligned} E[X^k] &= \sum_x x^k f(X) && \text{si } X \text{ es discreta.} \\ &= \int_{-\infty}^{\infty} x^k f(X)dx && \text{si } X \text{ es continua} \end{aligned}$$

Estos son llamados el k -ésimo momento de X .

Varianza

Sea X una variable aleatoria y sea $E(X) = \mu$. La distribución o dispersión de los valores de X alrededor del valor esperado se mide por la varianza, la cual se define como

$$\text{var}(X) = \sigma_X^2 = E[(X - \mu)^2]$$

La raíz cuadrada positiva de σ_X^2 , σ_X , se define como **desviación estándar** de X . La varianza o la desviación estándar da una indicación de qué tan cercanos o dispersos están los valores individuales de X respecto del valor de su media.

La varianza definida anteriormente se calcula de la siguiente forma:

$$\begin{aligned} \text{var}(X) &= \sum_x (X - \mu)^2 f(x) && \text{si } X \text{ es una variable aleatoria discreta} \\ &= \int_{-\infty}^{\infty} (X - \mu)^2 f(x)dx && \text{si } X \text{ es una variable aleatoria continua} \end{aligned}$$

Propiedades de la varianza

1. $E(X - \mu)^2 = E(X^2) - \mu^2$,

$$\begin{aligned} \text{var}(X) &= \sigma_X^2 = E(X - \mu)^2 \\ &= E(X^2) - \mu^2 \\ &= E(X^2) - [E(X)]^2 \end{aligned}$$

2. $\text{var}(c) = 0$

3. Si a y b son constantes,

$$\text{var}(aX + b) = a^2 \text{var}(X)$$

4. Si X y Y son variables aleatorias independientes,

$$\begin{aligned} \text{var}(X + Y) &= \text{var}(X) + \text{var}(Y) \\ \text{var}(X - Y) &= \text{var}(X) + \text{var}(Y) \end{aligned}$$

Esto puede generalizarse a más de dos variables.

5. Si X y Y son variables aleatorias independientes y a y b son constantes,

$$\text{var}(aX + bY) = a^2 \text{var}(X) + b^2 \text{var}(Y)$$

Función generadora de momentos

La función generadora de momentos $M_X(t)$ de una variable aleatoria X está definida por:

$$M_X(t) = E(e^{tX})$$

Donde la función generadora en t se define como la esperanza de e^{tX}

El dominio de M_X son todos los números reales t tal que e^{tX} tiene una esperanza finita.

Si X es una variable aleatoria discreta con distribución de probabilidades $p(x_i) = P(X = x_i)$, $i = 1, 2, \dots$, la función generadora de momentos de X está definida por:

$$M_X(t) = \sum_{j=1}^{\infty} e^{tx_j} p(x_j)$$

Si X es una variable aleatoria continua con función de densidad $f_X(x)$, la función generadora de momentos es

$$M_X(t) = \int_{-\infty}^{+\infty} e^{tx} f_X(x) dx$$

Ejemplo: Sea $f_X(x)$ la densidad de una normal estándar y se quiere saber cuál es la función generadora de momentos, entonces

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$M_X(t) = \int_{-\infty}^{+\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2-2tx}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(x-t)^2-t^2}{2}} dx$$

$$M_X(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{t^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{(x-t)^2}{2}} dx$$

Pero el $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(x-t)^2}{2}} dx$ tiene que ser uno porque esta es la función de densidad de una normal con media t por lo que queda

$$M_X(t) = e^{\frac{t^2}{2}}$$

La cual es la función generadora de la normal estándar.

Ahora suponga $Y = \sigma X + \mu$ la cual es una normal con media μ y varianza σ^2 porque la esperanza de Y va a ser

$$E(Y) = \sigma E(X) + \mu = \mu \text{ porque la esperanza de } X \text{ es cero}$$

$$V(Y) = \sigma^2 V(X) = \sigma^2 \text{ porque la varianza de } X \text{ es } 1$$

Ahora para calcular la función generadora de momento de $M_{\sigma X + \mu}(t)$

Primero se va a calcular la función generadora de una variable aleatoria por una constante

$$M_{\sigma X}(t) = E(e^{\sigma X t}) = E(e^{(\sigma t) X}) = M_X(\sigma t) \text{ que es la función generadora de } X \text{ valuada en } \sigma t$$

Entonces $M_{\sigma X}(t) = M_X(\sigma t) = e^{\frac{\sigma^2 t^2}{2}}$ la función generadora de σX en t es la función generadora de X valuada en σt .

$$\text{Ahora } M_{X+\mu}(t) = E(e^{(X+\mu)t}) = E(e^{Xt} e^{\mu t}) = e^{\mu t} E(e^{Xt}) = e^{\mu t} M_X(t)$$

Por lo tanto, $M_{X+\mu}(t) = e^{\mu t} M_X(t)$ la función generadora de $X + \mu$ en t es

$$M_{\sigma X + \mu}(t) = e^{\mu t} M_{\sigma X}(t) = e^{\mu t} e^{\frac{\sigma^2 t^2}{2}} = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

Esta es la función generadora de una normal con media μ y varianza σ^2

La otra cosa que hay que tomar en cuenta es

$$M_X(t) = M_X(0) + M'_X(0)t + \frac{M''_X(0)}{2!}t^2 + \dots \quad (1)$$

Como esto es e^{Xt} entonces se puede escribir como

$$e^{Xt} = 1 + Xt + \frac{X^2 t^2}{2!} + \dots$$

Entonces la esperanza de e^{Xt} es

$$E(e^{Xt}) = 1 + E(X)t + \frac{E(X^2)t^2}{2!} + \dots \quad (2)$$

Como la fórmula 1 es igual a la 2 y como son series de potencias en t y la igualdad se cumple para toda t , entonces de ahí se desprende que

$$M_X(0) = 1, \quad M'_X(0) = E(X), \quad M''_X(0) = E(X^2), \quad \dots$$

Entonces si se tiene la función generadora se puede calcular fácilmente la esperanza de X, X^2, X^3, \dots calculando las derivadas sucesivas de la función generadora y evaluándolas en cero.

De ahí se puede ver que la varianza es

$$\text{var}(X) = E(X^2) - [E(X)]^2 = M''_X(0) - [M'_X(0)]^2$$

Covarianza

Sean X y Y dos variables aleatorias con medias $\mu_x = E(X)$ y $\mu_y = E(Y)$, respectivamente. Entonces, la covarianza entre las dos variables se define como

$$\text{cov}(X, Y) = \sigma_{XY} = E\left((X - \mu_x)(Y - \mu_y)\right) = E(XY) - \mu_x\mu_y$$

Se observa con facilidad que la varianza de una variable es la covarianza de dicha variable con ella misma.

La covarianza se calcula de la siguiente manera:

$$\begin{aligned} \text{cov}(X, Y) &= \sum_y \sum_x (x - \mu_x)(y - \mu_y)f(x, y) \\ &= \sum_y \sum_x xyf(x, y) - \mu_x\mu_y \end{aligned}$$

Si X y Y son variables aleatorias discretas y

$$\begin{aligned} \text{cov}(X, Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y)f(x, y)dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y)dx dy - \mu_x\mu_y \end{aligned}$$

Si X y Y son variables aleatorias continuas.

Propiedades de la covarianza

1. Si X y Y son independientes, su covarianza es cero, pues

$$\text{cov}(X, Y) = E(XY) - \mu_x\mu_y$$

$$= \mu_x\mu_y - \mu_x\mu_y \quad \text{porque } E(XY) = E(X)E(Y) = \mu_x\mu_y \text{ cuando } X \text{ y } Y \text{ son independientes}$$

$$= 0$$

2. $\text{cov}(a + bX, c + dY) = bdcov(X, Y)$ Donde a, b, c y d son constantes.

Coeficiente de correlación

Sean X y Y dos variables aleatorias que tienen varianzas finitas distintas de cero. Una medida del grado de dependencia entre las dos variables aleatorias es el coeficiente de correlación.

El coeficiente de correlación (poblacional) ρ se define como

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{\{\text{var}(X)\text{var}(Y)\}}} = \frac{\sigma_{XY}}{\sigma_x \sigma_y}$$

Se dice que estas variables aleatorias no están correlacionadas si $\rho_{XY} = 0$. Puesto que $\text{cov}(X, Y) = 0$ si X y Y son independientes, vemos a la vez que las variables aleatorias independientes no están correlacionadas. También es posible que las variables aleatorias dependientes no estén correlacionadas.

Así definido, ρ_{XY} es una medida de la asociación lineal entre dos variables, y su valor se sitúa entre $-1 \leq \rho_{XY} \leq 1$, donde -1 indica una perfecta asociación negativa y $+1$ indica una perfecta asociación positiva, y que $|\rho_{XY}| = 1$ sí y sólo si $P(X = aY) = 1$ para cualquier constante a .

De la fórmula del coeficiente de correlación se ve que

$$\text{cov}(X, Y) = \rho_{XY} \sigma_x \sigma_y$$

DISTRIBUCIONES

Distribución normal

Se dice que una variable aleatoria continua X está normalmente distribuida si su función de distribución de probabilidad tiene la siguiente forma:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)} \quad -\infty < x < \infty$$

Donde μ y σ^2 , se conocen como *parámetros de la distribución*, son la media y la varianza de la distribución, representativamente. Las propiedades de esta distribución son las siguientes:

1. Es simétrica alrededor de su valor medio.
2. Aproximadamente 68% del área bajo la curva normal se encuentra entre los valores de $\mu \pm \sigma$, alrededor de 95% del área se encuentra entre $\mu \pm 2\sigma$ y alrededor de 99.7% del área se encuentra entre $\mu \pm 3\sigma$.

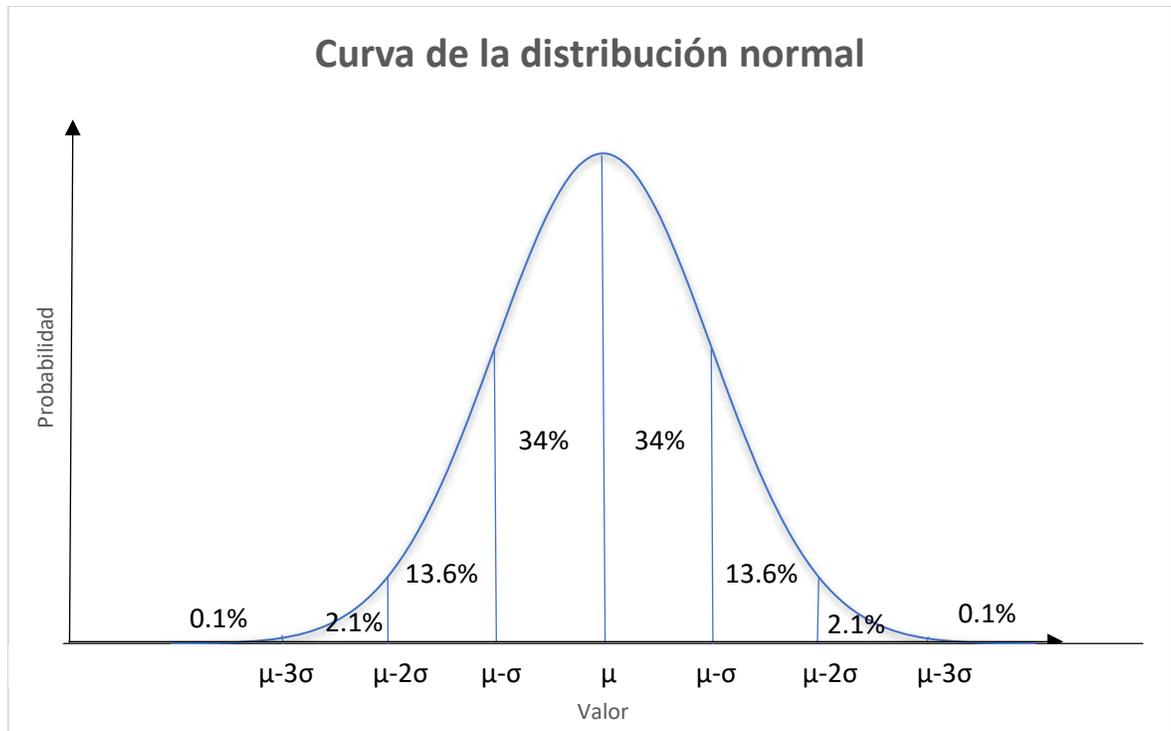


Figura 3. Curva de la distribución normal

3. La distribución normal depende de dos parámetros μ y σ^2 . Por tanto, una vez especificados, podemos encontrar la probabilidad de que X esté dentro de cierto intervalo mediante la función de distribución de probabilidad de la distribución normal. Pero esto se aligera al consultar la tabla de áreas debajo de la distribución normal estandarizada. Para utilizar esta tabla, convertimos la variable dada X normalmente distribuida con media μ y σ^2 en una variable Z normal estandarizada mediante la siguiente transformación:

$$Z = \frac{x - \mu}{\sigma}$$

Una propiedad importante de toda variable estandarizada es que su valor medio es cero y su varianza es la unidad. Así, Z tiene media cero y varianza unitaria. Al sustituir z en la función de distribución de probabilidad anterior obtenemos

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{1}{2}Z^2\right)} \quad Z \sim (0,1)$$

Que es la función de distribución de probabilidad de la variable normal estandarizada. Las probabilidades en la tabla de áreas debajo de la distribución normal estandarizada se basan en esta variable normal estandarizada. Por convención, denotamos una variable distribuida normalmente como

$$X \sim N(\mu, \sigma^2)$$

Donde \sim significa “distribuido como”, N significa distribución normal y las cantidades en los paréntesis son los dos parámetros de la distribución normal, a saber: la media y la varianza.

Según esta convención,

$$X \sim N(0,1)$$

Significa que X es una variable normalmente distribuida con media cero y varianza unitaria. En otras palabras, es una variable Z normalmente estandarizada.

Distribución de t-Student

Una variable aleatoria se distribuye según el modelo de probabilidad t-Student con ν grados de libertad y su función de densidad es:

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \quad -\infty < x < \infty$$

El parámetro de la distribución t es $\nu = n - 1$, su número de grados de libertad.

Si se define otra variable aleatoria Y

$$Y = \sigma X + \mu$$

Donde

σ es la desviación estándar y μ es la media

Entonces su función de distribución de probabilidad de Y es:

$$F_Y(y) = P(Y \leq y) = P(\sigma X + \mu \leq y) = P\left(X \leq \frac{y - \mu}{\sigma}\right) = \int_{-\infty}^{\frac{y - \mu}{\sigma}} h dh$$

La función de densidad de Y es (usando el teorema fundamental del cálculo):

$$f(X) = F'(X) = F'\left(\frac{y - \mu}{\sigma}\right) \frac{1}{\sigma} = \left(\frac{1}{\sigma}\right) \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{\left(\frac{y - \mu}{\sigma}\right)^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

La cual se le conoce como **distribución de t-student no estandarizada** que se generaliza a 3 parámetros, introduciendo un parámetro de localización μ y otro de escala σ .

La esperanza y varianza de esta versión de la distribución t son:

$$E(X) = \mu \quad \text{para } \nu > 1$$

$$Var(X) = \sigma^2 \frac{\nu}{\nu - 2} \quad \text{para } \nu > 2$$

Propiedades de la distribución t

1. Cada curva t tiene forma de campana con centro en 0.
2. Cada curva t, está más dispersa que la curva normal estándar.
3. A medida que ν aumenta, la dispersión de la curva t correspondiente disminuye.
4. A medida que $\nu \rightarrow \infty$, la secuencia de curvas t se aproxima a la curva normal estándar.

Los parámetros se calculan por la función de máxima verosimilitud, sin embargo, no es fácil dar una fórmula explícita.

Para simular los valores de la t-student en el programa R se utiliza la fórmula:

$$rt(N, df = df) * s + m$$

donde rt es la función de la distribución t-Student que se utiliza en R, N es el número de datos, df son los grados de libertad, s es el parámetro escala y m es el parámetro de localización.

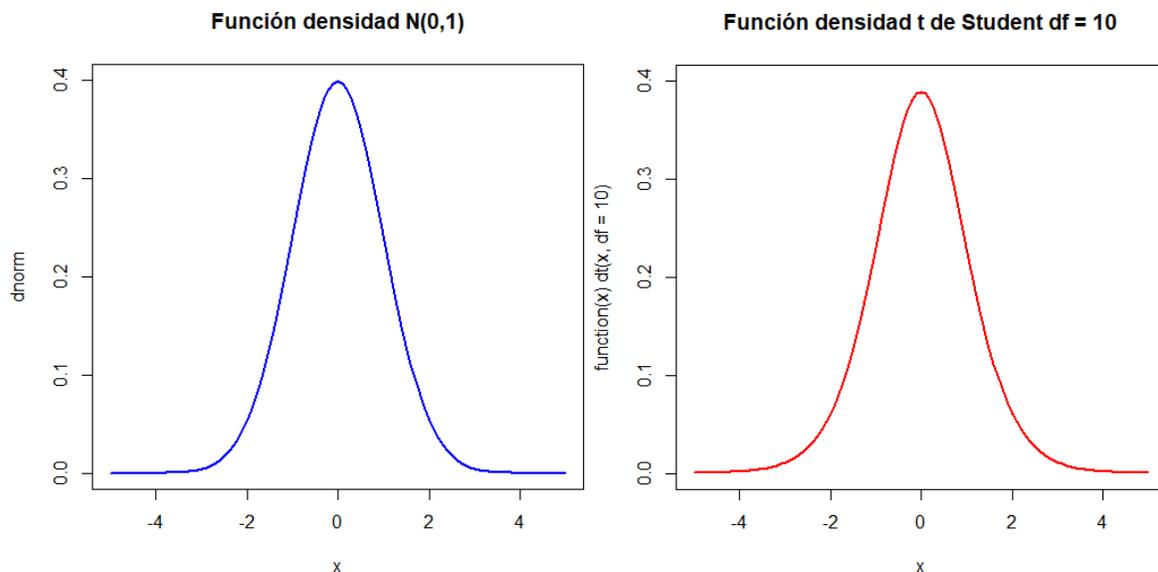


Figura 4. Función de densidad de las distribuciones Normal y la t de Student

Diferencia entre las distribuciones Normal y t de Student

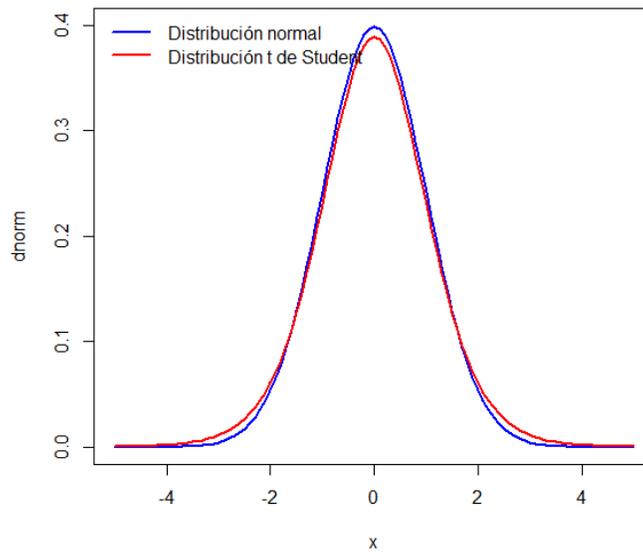


Figura 5. Diferencia entre las funciones de densidad de las distribuciones Normal y la t de Student

2.2 CONCEPTOS DE ESTADISTICA

Teorema del límite central

Sean X_1, \dots, X_n variables aleatorias independientes idénticamente distribuidas con $E(X) = \mu$ y $Var(X) = \sigma^2 < \infty$. Definamos

$$U_n = \sqrt{n} \left(\frac{\bar{X} - \mu}{\sigma} \right) \quad \text{donde } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Entonces, la función de distribución de U_n converge a la función de distribución normal estándar conforme $n \rightarrow \infty$.

En otras palabras, la conclusión del teorema indica que

$$P(a \leq U_n \leq b) \rightarrow \int_a^b \left(\frac{1}{\sqrt{2\pi}} \right) e^{-u^2/2} du$$

Método de máxima verosimilitud.

Suponga que la función de verosimilitud depende de k parámetros $\theta_1, \theta_2, \dots, \theta_k$. Elija como estimaciones aquellos valores de los parámetros que maximizan la verosimilitud $L(x_1, x_2, \dots, x_n | \theta_1, \theta_2, \dots, \theta_k)$.

La técnica, denominada *método de máxima verosimilitud*, elige como estimaciones los valores de los parámetros que maximizan la verosimilitud (la función de probabilidad conjunta o la función de densidad conjunta) de la muestra observada de tal forma que los datos observados tengan máxima probabilidad de ocurrir.

Para destacar el hecho de que la función de verosimilitud es una función de los parámetros $\theta_1, \theta_2, \dots, \theta_k$, a veces expresamos la función de verosimilitud como $L(\theta_1, \theta_2, \dots, \theta_k)$.

Hipótesis

Una hipótesis o hipótesis estadística es una afirmación o conjetura acerca de la distribución de una o más variables aleatorias.

Una hipótesis es simple si especifica por completo la distribución de probabilidad en cuestión. Es compuesta cuando no especifica por completo la distribución de probabilidad en cuestión.

En general se contrastan dos hipótesis de la siguiente manera:

$$H_0: (\text{hipótesis nula}) \quad \text{vs} \quad H_1: (\text{hipótesis alternativa})$$

Tanto la hipótesis nula como la alternativa pueden ser simple o compuesta, es así que se tienen cuatro diferentes tipos de contraste de hipótesis: simple vs simple, simple vs compuesta, compuesta vs simple y compuesta vs compuesta.

Pruebas de hipótesis

Una prueba de hipótesis es una regla para decidir si se acepta la hipótesis nula o se rechaza en favor de la hipótesis alternativa. En el supuesto de que el modelo ajustado sea una aproximación razonablemente buena de la realidad, tenemos que establecer criterios apropiados para comprobar si los valores estimados obtenidos de una ecuación concuerdan con las expectativas de la teoría que estamos probando.

Al tomar una decisión de este tipo se corre el riesgo de cometer errores. Al rechazo de la hipótesis nula cuando ésta es verdadera se le conoce como *error tipo I*, y a la probabilidad de cometer este primer tipo de error se le denota por α . En cambio, a la aceptación de la hipótesis nula cuando ésta es falsa recibe el nombre de *error tipo II*, y a la probabilidad de cometer este segundo tipo de error se le denota por β . A las probabilidades α y β se les conoce con el nombre de **nivel de significancia**. Estas definiciones de errores se resumen en la siguiente tabla:

Tabla 1. Tipos de errores

	H_0 cierta	H_0 falsa
Rechazar H_0	Error tipo I Con probabilidad α	Decisión Correcta
No rechazar H_0	Decisión Correcta	Error tipo II Con probabilidad β

El procedimiento consta de 5 pasos esenciales:

Paso 1: Plantear hipótesis nula (H_0) e hipótesis alternativa (H_1).

- La hipótesis alternativa plantea matemáticamente lo que queremos demostrar.
- La hipótesis nula plantea exactamente lo contrario.

Paso 2: Determinar nivel de significancia (Rango de aceptación de hipótesis alternativa) α .

Se considera:

- 0.05 para proyectos de investigación.
- 0.01 para aseguramiento de calidad.
- 0.10 para encuestas de mercadotecnia y políticas.

Paso 3: Evidencia muestral. Se calcula la media y la desviación estándar a partir de la muestra.

Paso 4: Se aplica la distribución la cual se esté utilizando para calcular la probabilidad de error (P)

Paso 5: En base a la evidencia disponible se acepta o se rechaza la hipótesis alternativa.

- Si la probabilidad de error (P) es mayor que el nivel de significancia: se rechaza hipótesis alternativa.
- Si la probabilidad de error (P) es menor que el nivel de significancia: se acepta hipótesis alternativa.

Serie de tiempo

Una serie de tiempo es una secuencia de observaciones, medidos en determinados momentos del tiempo, ordenados cronológicamente y, espaciados entre sí de manera uniforme, así los datos usualmente son dependientes entre sí. El principal objetivo de una serie de tiempo X_t , donde $t = 1, 2, \dots, n$ es su análisis para hacer pronóstico.

Algunos ejemplos en los que surgen series de tiempo:

- Economía y Finanzas
- Modelado ambiental
- Demografía
- Medicina
- Ingeniería
- Control de calidad

Los datos de series temporales consisten en un único individuo observado durante múltiples periodos de tiempo.

Objetivos del análisis de series de tiempo:

- Descripción: resumen estadístico, gráficos
- Análisis e interpretación: encuentre un modelo para describir la dependencia del tiempo en los datos, ¿podemos interpretar el modelo?
- Pronóstico o predicción: dada una muestra de la serie, pronostique el siguiente valor o los siguientes valores
- Control: ajusta varios parámetros de control para que la serie se ajuste más a un objetivo
- Ajuste: en un modelo lineal, los errores podrían formar una serie temporal de observaciones correlacionadas, y es posible que queramos ajustar las variaciones estimadas para permitir esto.

Suponga que la serie X_t se ejecuta a lo largo del tiempo, es decir $(X_t)_{t=0, \pm 1, \pm 2, \dots}$, pero solo se observa en los tiempos $t = 1, 2, \dots, n$.

La teoría de las series temporales se basa en el supuesto de "estacionariedad de segundo orden". Los datos de la vida real a menudo no son estacionarios: p. exhiben una tendencia lineal a lo largo del tiempo o tienen un efecto estacional. Por lo tanto, los supuestos de estacionariedad a continuación se aplican después de que se hayan eliminado las tendencias / efectos estacionales.

Una serie de tiempo tiene cuatro componentes, los cuales son:

- **Tendencia (T):** Representa el comportamiento predominante de la serie. Se puede definir como el cambio de la media a lo largo de un periodo de tiempo.
- **Ciclo (C):** Se caracterizan por tener oscilaciones alrededor de la tendencia con una larga duración, y sus factores no son claros.
- **Estacionalidad (E):** Es un movimiento periódico que se producen dentro de un periodo corto y conocido.
- **Aleatorio (A):** Son movimientos erráticos que no siguen un patrón específico. Este componente es impredecible.

Proceso estocástico

Un proceso estocástico es una colección de variables aleatorias $\{X_t, t \in T\}$, donde t es el tiempo. O también se puede definir como una sucesión de variables aleatorias cuyas características pueden variar a lo largo del tiempo.

Se le conoce como estados a los posibles valores que puede tomar la variable aleatoria, por lo que se puede tener un espacio de estados discreto y un espacio de estados continuo. También, la variable t puede ser de tipo discreto o de tipo continuo.

Proceso estocástico estacionario

Un proceso estocástico se dice que es estacionario si tiene un comportamiento constante a lo largo del tiempo, es decir, si su media y su varianza son constantes en el tiempo y si el valor de la covarianza entre dos periodos depende solamente de la distancia o rezago entre estos dos periodos de tiempo y no del tiempo en el cual se ha calculado la covarianza.

Hay dos definiciones de estacionariedad: estacionariedad débil y estacionariedad estricta.

Estacionariedad débil: La serie de tiempo $\{X_t, t \in \mathbb{Z}\}$ (donde \mathbb{Z} es un conjunto entero) se dice que es estacionario si

$$i. E(X_t^2) < \infty \quad \forall t \in \mathbb{Z}$$

$$ii. E(X_t) = \mu \quad \forall t \in \mathbb{Z}$$

$$iii. \gamma_X(s, t) = \gamma_X(s + h, t + h) \quad \forall s, t, h \in \mathbb{Z}$$

Sea X_t una serie de tiempo entonces con estas propiedades se cumple la estacionariedad débil:

$$\text{Media } E(X_t) = E(X_{t+k}) = \mu$$

$$\text{Varianza } V(X_t) = V(X_{t+k}) = \sigma^2$$

$$\text{Covarianza } \gamma_k = E[(X_t - \mu)(X_{t+k} - \mu)] = \gamma_{k+l} \text{ para todo } l$$

Donde γ_k , la covarianza (o autocovarianza) al rezago k , es la covarianza entre los valores de X_t y X_{t+k} , que están separados k periodos.

Estacionariedad estricta: La serie de tiempo $\{X_t, t \in Z\}$ se dice que es estacionario estricto si al realizar un mismo desplazamiento en el tiempo de todas las variables de cualquier distribución conjunta finita de $(X_{t_1}, X_{t_2}, \dots, X_{t_k})$ es lo mismo que $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h})$, es decir:

$$F(X_{t_1}, X_{t_2}, \dots, X_{t_k}) = F(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h})$$

Para todo conjunto de índices (t_1, t_2, \dots, t_k) y todo h .

Ruido blanco

Un ruido blanco, denotado por ε_t es un proceso estocástico, el cual sus valores son independientes e idénticamente distribuidos a lo largo del tiempo con media cero, varianza constante y sin autocorrelación. No se puede predecir a partir de su pasado ya que el proceso es aleatorio y, por consiguiente, carece de memoria.

$$\varepsilon_t \sim N(0, \sigma^2) \quad \text{cov}(\varepsilon_{t_i}, \varepsilon_{t_j}) = 0 \quad \forall t_i, t_j$$

Procesos Lineales Estacionarios

Procesos Autoregresivos $AR(p)$

Los modelos autoregresivos se basan en la idea de que el valor actual de la serie, X_t , puede explicarse en función de p valores pasados $X_{t-1}, X_{t-2}, \dots, X_{t-p}$, donde p determina el número de rezagos necesarios para pronosticar un valor actual.

El modelo autoregresivo de orden p está dado por:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad \dots (1)$$

Expresado en términos del operador de retardos,

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) X_t = \varepsilon_t$$

$$\phi_p(L) X_t = \varepsilon_t$$

Donde ε_t es un proceso de ruido blanco y $\phi_0, \phi_1, \phi_2, \dots, \phi_p$ son los parámetros del modelo.

Proceso de Medias Móviles $MA(q)$

Modelo "determinados por una fuente externa". Estos modelos suponen linealidad, el valor actual de la serie, X_t , está influenciado por los valores de la fuente externa.

El modelo de promedio móviles de orden q está dado por:

$$X_t = \theta_0 - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} + \varepsilon_t \quad \dots (2)$$

Expresado en términos del polinomio operador de retardos se tiene:

$$X_t = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) \varepsilon_t$$

$$X_t = \theta_q(L) \varepsilon_t$$

donde ε_t es un proceso de ruido blanco, y $\theta_1, \theta_2, \dots, \theta_q$ son los parámetros del modelo.

Proceso Autoregresivo de Medias Móviles ARMA(p,q)

Existen muchos casos en donde una serie de tiempo X_t , tenga características de AR y de MA a la vez, o de igual forma, sea ARMA. Así X_t sigue un proceso $ARMA(p,q)$, en este proceso habrá p términos autoregresivos y q términos de media móvil.

$$X_t = c + \underbrace{\phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p}}_{AR(p)} + \underbrace{\theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}}_{MA(q)} + \varepsilon_t$$

donde ε_t es un proceso de ruido blanco, y $c, \phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$ son los parámetros del modelo.

Para un proceso $ARMA(p,q)$ una condición de estacionariedad es la misma que para un proceso $AR(p)$, del mismo modo una condición de invertibilidad es la misma que para el proceso $MA(q)$.

El modelo $ARMA(p,q)$ se puede escribir en términos del operador de retardos de la siguiente manera:

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) X_t = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q) \varepsilon_t$$

$$\phi_p(L) X_t = \theta_q(L) \varepsilon_t$$

Donde:

$\phi_p(L)$: es el polinomio autoregresivo.

$\theta_q(L)$: es el polinomio de medias móviles.

Los modelos $ARMA(p,q)$ siempre van a compartir las características de los modelos $AR(p)$ y $MA(q)$, esto es porque contiene a ambas estructuras a la vez. El modelo $ARMA(p,q)$ tiene media cero, varianza constante y finita y una función de autocorrelación infinita, esta última es infinita decreciendo rápidamente hacia cero.

2.3 CONCEPTOS FINANCIEROS

Acciones

Son títulos emitidos por una sociedad que representa una parte en la que se divide de igual manera el capital social una empresa. Por ejemplo, si una empresa tiene un capital social de cincuenta mil pesos, y se divide en cinco partes, cada acción tendrá un valor de diez mil pesos.

El hecho de poseer acciones de una sociedad da derechos y obligaciones a su titular. Sin embargo, entre mayor número de acciones se tenga, se tiene una mayor influencia y un nivel más alto de beneficios y/o pérdidas. Entre los derechos están: el derecho a voto en las juntas de accionistas, exigir información sobre la empresa, recibir dividendos o vender las acciones de las que disponga. Entre las obligaciones están que el accionista deberá soportar pérdidas si la empresa no obtiene buenos resultados.

Rendimiento de una acción

El rendimiento de una acción es una medida de la rentabilidad que se genera debido al pago de dividendos al accionista y del rendimiento del capital durante un período de tiempo. El período durante el cual se miden los rendimientos de las acciones se suele medir en bases diarias, semanales, mensuales y anuales.

El rendimiento de una acción se divide en dos componentes: las ganancias o las pérdidas de capital y los dividendos. Las ganancias o pérdidas de capital son el resultado de los movimientos en los precios de las acciones; una ganancia es lo que resulta de un aumento de precios, mientras que una pérdida resulta de una disminución en estos. Los dividendos suelen ser pagados a los accionistas de una empresa, basados en los ingresos de la compañía. Cuando estos dos componentes se suman juntos, forman el retorno total para la acción.

$$R_t = \frac{D_t}{P_{t-1}} + \frac{P_t - P_{t-1}}{P_{t-1}}$$

R_t : Rendimiento total de la acción en el período t

D_t : Monto de dividendos en efectivo pagado al accionista en el periodo t

P_t : Es el precio de la acción en el tiempo t

P_{t-1} : Es el precio de la acción en el tiempo t-1

$P_t - P_{t-1}$: Ganancia o pérdida de capital, dada por la diferencia entre el precio de venta al final del periodo t y el precio de compra al principio del período.

En el caso en que no haya dividendos:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

En este proyecto se considera el caso donde no hay dividendos.

2.4 CONCEPTOS ESTOCÁSTICOS

Caminata aleatoria

Una caminata aleatoria es un proceso estocástico en el que avanza una unidad hacia adelante y una unidad hacia atrás con probabilidad p y $1 - p$ respectivamente

$$P(X = 1) = p \text{ y } P(X = -1) = 1 - p$$

Donde X_1, X_2, \dots, X_n son independientes y con la misma distribución de probabilidad de X .

Y sea

$$S_n = X_1 + X_2 + \dots + X_n$$

Donde S_1, S_2, \dots, S_n es un proceso estocástico que representa la posición de la partícula después de $1, 2, \dots, n$ movimientos.

La esperanza y varianza de X son:

$$\begin{aligned} E(X) &= 1 * P(X = 1) + (-1) * P(X = -1) \\ &= p + (-1) * (1 - p) = 2p - 1 \end{aligned} \quad \dots(1)$$

si $p = \frac{1}{2}$, entonces $E(X) = 0$.

$$\begin{aligned} V(X) &= E(X^2) - (E(X))^2 \\ &= 1^2 * P(X = 1) + (-1)^2 * P(X = -1) - (2p - 1)^2 \\ &= p + (1 - p) - (2p - 1)^2 \\ &= 1 - (2p - 1)^2 \\ &= 1 - (4p^2 - 4p + 1) \\ &= -4p^2 + 4p = 4p(1 - p) \end{aligned} \quad \dots(2)$$

Sea $Y = \frac{X+1}{2}$ una variable de Bernoulli

$$P(Y = 1) = P(X = 1) = p \quad \text{y} \quad P(Y = 0) = P(X = -1) = 1 - p$$

$$X = 2Y - 1$$

$$S_n = X_1 + X_2 + \dots + X_n = 2(Y_1 + Y_2 + \dots + Y_n) - n$$

$$\begin{aligned} E(S_n) &= 2(E(Y_1 + Y_2 + \dots + Y_n)) - n \\ &= 2(E(Y_1) + E(Y_2) + \dots + E(Y_n)) - n \end{aligned} \quad \text{de (1)}$$

$$\begin{aligned}
&= n(2p - 1) \\
V(S_n) &= 2(V(Y_1 + Y_2 + \dots + Y_n)) \\
&= 4(V(Y_1) + V(Y_2) + \dots + V(Y_n)) \quad \text{de (2)} \\
&= 4np(1 - p)
\end{aligned}$$

Por lo tanto, $S_n \sim N(n(2p - 1), 4np(1 - p))$ que sigue una distribución binomial.

Movimiento browniano

El **movimiento Browniano** en (Ω, \mathcal{A}, P) es un proceso aleatorio, $W = (W_t)_{\{t \geq 0\}}$ tal que

- a) Sus trayectorias son continuas.
- b) Sus incrementos son independientes. Si $0 \leq t_1 \leq \dots \leq t_n$, entonces
$$W_{t_1}, W_{t_2} - W_{t_1}, W_{t_n} - W_{t_{n-1}}$$
son variables aleatorias independientes.
- c) $W_0 = 0$,
- d) $W_t - W_s$ es una variable gaussiana, que sigue una distribución normal con media cero y varianza $t - s$, es decir

$$W_t - W_s \sim N(0, t - s).$$

Se sabe que X es gaussiana o normal ($X \sim N(\mu, \sigma^2)$) cuando su distribución de probabilidad es

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du$$

La densidad es la campana de Gauss

$$\phi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Lema de Itô

Sea V_t una función de W_t y de t , digamos que $V_t = (W_t, t)$

Por la fórmula de Taylor para una función de dos variables:

$$\begin{aligned} f(x_0 + h, y_0 + k) &= f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k \\ &+ \frac{1}{2} \left(\frac{\partial^2 f}{\partial x^2}(x_0, y_0)h^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)hk + \frac{\partial^2 f}{\partial y^2}(x_0, y_0)k^2 \right) + R \end{aligned}$$

$$\begin{aligned} f(x_0 + h, y_0 + k) &= f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(x_0, y_0)h^2 \\ &+ \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)hk + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(x_0, y_0)k^2 + R \end{aligned}$$

Sea $h = dx$ y $k = dy$

$$df(x, y) = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dx)^2 + \frac{1}{2} \frac{\partial^2 f}{\partial y^2} (dy)^2 + \frac{\partial^2 f}{\partial x \partial y} dx dy + R$$

Pero $\lim_{(dx, dy) \rightarrow (0, 0)} \frac{R}{dx^2 + dy^2} = 0$, por lo tanto, R se puede despreciar

Si $V_t = (W_t, t)$ es una función que depende del movimiento browniano en el tiempo t y del tiempo, se puede escribir como:

$$dV_t(W_t, t) = \frac{\partial V_t}{\partial W_t} dW_t + \frac{\partial V_t}{\partial t} dt + \frac{1}{2} \frac{\partial^2 V_t}{\partial W_t^2} (dW_t)^2 + \frac{1}{2} \frac{\partial^2 V_t}{\partial t^2} (dt)^2 + \frac{\partial^2 V_t}{\partial W_t \partial t} dW_t dt \quad \dots (3)$$

Es una forma de tratar de entenderlo intuitivamente.

La idea es eliminar todos los términos T , que converjan a 0 más rápido que dt , ya que,

$$\lim_{dt \rightarrow 0} \frac{T}{dt} = 0$$

Los términos que cumplan con la fórmula anterior se despreciarán porque tienden a 0 más rápido que dt .

Cambio en el movimiento browniano

Si se supone $dW_t = W_{t+dt} - W_t$

Sea $X = W_{t+dt} - W_t \sim N(0, dt)$ por la definición del movimiento browniano

El valor esperado es

$E(dW_t^2) = E(X^2) = V(X) + (E(X))^2$, pero X tiene una distribución normal con media 0, por lo tanto, su esperanza es 0.

$$= V(X) = dt, \quad \dots(4)$$

ya que sigue $N(0, dt)$

Como dW_t^2 es una variable aleatoria y su valor esperado es dt , entonces cuando dt tiende a 0, dW_t^2 va a tender a 0 con la misma rapidez que dt .

La varianza es:

$$V(dW_t^2) = V(X^2) = E(X^4) - (E(X^2))^2 \quad \dots(5)$$

Se sabe de (4) que $E(X^2) = dt$ por lo que $(E(X^2))^2 = dt^2$

Para calcular $E(X^4)$ se usará la función generadora de momentos de una variable aleatoria X .

$$M_X(s) = E(e^{sX}) = E\left(1 + \frac{sX}{1!} + \frac{s^2X^2}{2!} + \frac{s^3X^3}{3!} + \dots\right) \text{ desarrollado en serie de potencias}$$

$$= 1 + sE(X) + \frac{s^2}{2!}E(X^2) + \frac{s^3}{3!}E(X^3) + \dots \quad \dots(6)$$

Por otro lado, la M_X se puede desarrollar también en serie de Taylor

$$M_X(s) = M_X(0) + sM_X'(0) + \frac{s^2}{2!}M_X''(0) + \frac{s^3}{3!}M_X'''(0) + \dots \quad \dots(7)$$

Comparando (6) y (7), se desprende que

$$M_X(0) = 1, M_X'(0) = E(X), M_X''(0) = E(X^2), M_X'''(0) = E(X^3) \dots M_X^{(n)}(0) = E(X^n)$$

Así, se puede ver que una forma de calcular la $E(X^n)$, si se conoce la función generadora de momentos de la variable aleatoria X basta con derivarla n veces y evaluarla en 0.

De la función generadora de momentos de la distribución normal

$$M_X(s) = e^{\mu s + \frac{\sigma^2 s^2}{2}}$$

Para calcular $E(X^4)$, se deriva 4 veces la fórmula anterior utilizando el programa Maxima y evaluando en 0 resulta:

$$M_X^{(4)}(0) = \mu^4 + 6\sigma^2\mu^2 + 3\sigma^4$$

Pero se sabe que $\mu = 0$, por lo que de la fórmula anterior se tiene

$$M_X^{(4)}(0) = E(X^4) = 3\sigma^4 = 3(\sqrt{dt})^4 = 3dt^2$$

Donde $\sigma = \sqrt{dt}$

Por lo que la varianza del cambio del movimiento Browniano es

$$V(dW_t^2) = E(X^4) - (E(X^2))^2 = 3dt^2 - dt^2 = 2dt^2 \quad \dots(8)$$

Por lo tanto

$$E(dW_t^2) = dt$$

$$V(dW_t^2) = 2dt^2$$

Esto quiere decir que cuando $dt \rightarrow 0$, la varianza de dW_t^2 va a tender más rápido a cero que dt y cuando la varianza es muy cercana a 0, la variable aleatoria es casi constante que es el valor esperado.

Lo anterior justifica que en la fórmula (3) se puede sustituir $(dW_t)^2$ por dt , eliminar $dW_t dt$ y $(dt)^2$ porque tienden a 0 más rápido que dt .

$$dW_t dt = (W_{t+dt} - W_t)dt \sim N(0, (dt)^3)$$

$$\frac{dW_t dt}{dt} = dW_t, \quad \text{si } dt \rightarrow 0 \text{ entonces } dW_t = 0$$

Por lo tanto, la ecuación queda:

$$dV_t = \frac{\partial V_t}{\partial W_t} dW_t + \frac{\partial V_t}{\partial t} dt + \frac{1}{2} \frac{\partial^2 V_t}{\partial W_t^2} (dW_t)^2 + \frac{1}{2} \frac{\partial^2 V_t}{\partial t^2} (dt)^2 + \frac{\partial^2 V_t}{\partial W_t \partial t} dW_t dt$$

$$dV_t = \frac{\partial V_t}{\partial W_t} dW_t + \frac{\partial V_t}{\partial t} dt + \frac{1}{2} \frac{\partial^2 V_t}{\partial W_t^2} dt$$

Asociando dt

$$dV_t = \left(\frac{\partial V_t}{\partial t} + \frac{1}{2} \frac{\partial^2 V_t}{\partial W_t^2} \right) dt + \frac{\partial V_t}{\partial W_t} dW_t \quad \dots(8)$$

La fórmula anterior es la llamada fórmula de Itô.

Precio de una acción

El modelo para el precio de una acción es el siguiente:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \quad \dots(9)$$

Un cambio pequeño en el precio de la acción se puede desglosar en 2 cambios, uno determinista y otro aleatorio. Donde:

$\mu S_t dt$ es la componente determinista, que es directamente proporcional al precio de la acción y al mismo tiempo por el dt

$\sigma S_t dW_t$ es la componente aleatoria, es directamente proporcional al precio de la acción por un cambio en el movimiento browniano.

Para comenzar, aplicando el mismo orden de ideas por la fórmula de Taylor se obtiene:

$$d \ln S_t(t, S_t) = \frac{\partial \ln S_t}{\partial t} dt + \frac{\partial \ln S_t}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 \ln S_t}{\partial t^2} (dt)^2 + \frac{1}{2} \frac{\partial^2 \ln S_t}{\partial S_t^2} (dS_t)^2 + \frac{\partial^2 \ln S_t}{\partial t \partial S_t} dt dS_t$$

Como (9) no depende de t , entonces las parciales con respecto a t son cero.

$$\begin{aligned} d \ln S_t &= \frac{\partial \ln S_t}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 \ln S_t}{\partial S_t^2} (dS_t)^2 \\ &= \frac{1}{S_t} dS_t - \frac{1}{2} \frac{1}{S_t^2} (dS_t)^2 \end{aligned} \quad \dots(10)$$

Sustituyendo dS_t (9) en (10)

$$\begin{aligned} &= \frac{1}{S_t} (\mu S_t dt + \sigma S_t dW_t) - \frac{1}{2} \frac{1}{S_t^2} (\mu S_t dt + \sigma S_t dW_t)^2 \\ &= \mu dt + \sigma dW_t - \frac{1}{2} (\mu^2 (dt)^2 + 2\mu\sigma dt dW_t + \sigma^2 (dW_t)^2) \end{aligned} \quad \dots(11)$$

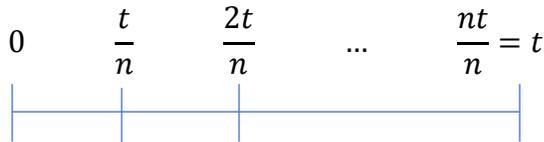
Por lo tanto, en (11) se puede sustituir $(dW_t)^2$ por dt , eliminar $dW_t dt$ y $(dt)^2$ porque tienden a 0 más rápido que dt .

$$\begin{aligned} &= \mu dt + \sigma dW_t - \frac{1}{2} \sigma^2 dt \\ &= \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dW_t \end{aligned}$$

Integrando

$$\begin{aligned} \int_0^t d \ln S_t dt &= \int_0^t \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma \int_0^t dW_t \\ \ln S_t - \ln S_0 &= \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma \int_0^t dW_t \end{aligned}$$

Para calcular $\int_0^t dW_t$



$$\begin{aligned} \int_0^t dW_t &= W_t - W_{(n-1)\frac{t}{n}} + W_{(n-1)\frac{t}{n}} - W_{(n-2)\frac{t}{n}} + W_{(n-2)\frac{t}{n}} + \dots + W_{\frac{t}{n}} - W_0 \\ &= W_t - W_0 \end{aligned}$$

Como por definición $W_0 = 0$,

$$\int_0^t dW_t = W_t$$

Por lo que

$$\ln S_t - \ln S_0 = \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t$$

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t} \quad \dots(12)$$

La ecuación (12) es la fórmula del precio de una acción.

Para calcular el **valor esperado del precio de una acción** se utilizará la función generadora de momentos de una variable aleatoria en s

$$M_X(s) = E(e^{sX})$$

$$\text{Si } s = 1, \quad M_X(1) = E(e^X)$$

La función generadora de momentos de la distribución normal está dada por:

$$M_X(s) = e^{\mu s + \frac{\sigma^2 s^2}{2}}$$

Donde $X \sim N(\mu, \sigma^2)$

Entonces al evaluar en 1 resulta: $M_X(1) = e^{\mu + \frac{\sigma^2}{2}}$

Por lo que

$$E(S_t) = E(S_0 e^{Z_t}) = S_0 E(e^{Z_t}) = S_0 M_{Z_t}(1) = S_0 e^{\mu t + \frac{\sigma^2 t}{2}} = S_0 e^{\left(\mu + \frac{\sigma^2}{2}\right)t}$$

La varianza es

$$\begin{aligned} V(S_t) &= V(S_0 e^{Z_t}) = S_0^2 V(e^{Z_t}) = S_0^2 \{E[(e^{Z_t})^2] - [E(e^{Z_t})]^2\} \\ &= S_0^2 \left\{ E(e^{2Z_t}) - \left[e^{\left(\mu + \frac{\sigma^2}{2}\right)t} \right]^2 \right\} \\ &= S_0^2 \left\{ M_{Z_t}(2) - \left[e^{\left(\mu + \frac{\sigma^2}{2}\right)t} \right]^2 \right\} \\ &= S_0^2 \left(e^{2\mu t + \frac{4\sigma^2 t}{2}} - e^{2\left(\mu + \frac{\sigma^2}{2}\right)t} \right) \\ &= S_0^2 (e^{2\mu t + 2\sigma^2 t} - e^{2\mu t + \sigma^2 t}) \\ &= S_0^2 e^{\sigma^2 t} \end{aligned}$$

La distribución de probabilidad del logaritmo del rendimiento del precio de la acción es

$$\ln \frac{S_t}{S_0} \sim N(\mu t, \sigma^2 t)$$

La función de distribución del precio de la acción es

$$\begin{aligned} F_{S_t}(s_t) &= P(S_t \leq s_t) = P\left(\frac{S_t}{S_0} \leq \frac{s_t}{S_0}\right) = P\left(\ln \frac{S_t}{S_0} \leq \ln \frac{s_t}{S_0}\right) \\ &= P\left(\frac{\ln \frac{S_t}{S_0} - \mu t}{\sigma \sqrt{t}} \leq \frac{\ln \frac{s_t}{S_0} - \mu t}{\sigma \sqrt{t}}\right) = \Phi\left(\frac{\ln \frac{s_t}{S_0} - \mu t}{\sigma \sqrt{t}}\right) \end{aligned}$$

La función de densidad de probabilidad del precio

$$\begin{aligned} f_{S_t}(s_t) &= F'_{S_t}(s_t) = \Phi'\left(\frac{\ln \frac{s_t}{S_0} - \mu t}{\sigma \sqrt{t}}\right) \frac{1}{\sigma \sqrt{t}} \frac{1}{S_t} \\ &= \frac{1}{\sqrt{2\pi} \sigma \sqrt{t} S_t} e^{-\frac{(\ln \frac{s_t}{S_0} - \mu t)^2}{2\sigma^2 t}} \end{aligned} \quad \dots (13)$$

A la ecuación (13) se le denomina log-normal.

Proyección del precio de una acción siguiendo una normal

Cuando los rendimientos siguen una distribución normal, el precio de la acción en $t + 1$ es

$$P_{t+1} = P_t e^{\mu + \sigma W_{[t,t+1]}}$$

Donde $W_{[t,t+1]} \sim N(0,1)$

Si se tiene dado P_0 y se quieren proyectar precios hacia adelante, entonces se elabora una proyección para P_1

$$P_1 = P_0 e^{\mu + \sigma Z_1}, \quad \text{donde } Z_1 \sim N(0,1)$$

Si se quiere generar P_2 dado que ya se generó P_1 , entonces,

$$P_2 = P_1 e^{\mu + \sigma Z_2}, \quad \text{donde } Z_2 \sim N(0,1)$$

Si se está interesado en generar la proyección hasta $t + k$, entonces

$$P_{t+k} = P_{t+k-1} e^{\mu + \sigma Z_k}$$

$$\begin{aligned}
&= P_{t+k-2}e^{\mu+\sigma Z_2} \\
&= P_{t+k-3}e^{\mu+\sigma Z_3} \\
&\quad \vdots \\
&= P_t e^{\mu k + \sigma(Z_1 + Z_2 + \dots + Z_k)} \\
&= P_t e^{\mu k + \sigma Z}
\end{aligned}$$

Donde $Z \sim N(0, \sqrt{k}\sigma)^*$

$$*V(\sigma Z) = \sigma^2 V(Z) = \sigma^2 k$$

$$*s = \sqrt{V(\sigma Z)} = \sqrt{\sigma^2 k} = \sqrt{k}\sigma$$

Por lo que, el precio en $t + k$ se puede generar conociendo el precio en t

$$P_{t+k} = P_t e^{\mu k + \sigma Z}, \quad \text{donde } Z \sim N(0, \sqrt{k}\sigma)$$

Esto quiere decir que conforme el valor de k aumenta, la varianza va a crecer ya que depende de \sqrt{k} , entonces cuando se quiera una proyección más hacia adelante, más puede fallar.

2.5 CONCEPTOS DE PROGRAMACIÓN

Lenguaje de programación R

R es un lenguaje de programación y un entorno de software libre para computación y gráficos estadísticos. R proporciona una amplia variedad de técnicas estadísticas (modelado lineal y no lineal, pruebas estadísticas clásicas, análisis de series de tiempo, clasificación, agrupamiento, ...) y gráficos, y es altamente extensible. El lenguaje R es ampliamente utilizado entre estadísticos y mineros de datos para desarrollar software estadístico y análisis de datos. Se compila y se ejecuta en una amplia variedad de plataformas UNIX, Windows y MacOS.

R es un paquete de GNU. El código fuente para el entorno de software R está escrito principalmente en C, Fortran y R. R está disponible libremente bajo la Licencia pública general de GNU, y se proporcionan versiones binarias precompiladas para varios sistemas operativos. Aunque R tiene una interfaz de línea de comando, hay varios front-ends gráficos, sobre todo RStudio. Entornos de desarrollo integrados están disponibles.

Funciones estadísticas

R y sus bibliotecas implementan una amplia variedad de técnicas estadísticas y gráficas, que incluyen modelos lineales y no lineales, pruebas estadísticas clásicas, análisis de series de tiempo, clasificación, agrupamiento y otros. R es fácilmente extensible a través de funciones y extensiones. Muchas de las funciones estándar de R están escritas en R, lo que facilita a los usuarios seguir las elecciones algorítmicas realizadas. R es altamente extensible mediante el uso de paquetes enviados por el usuario para funciones específicas o áreas de estudio específicas.

Otra fortaleza de R son los gráficos estáticos, que pueden producir gráficos con calidad de publicación, incluidos los símbolos matemáticos. Los gráficos dinámicos e interactivos están disponibles a través de paquetes adicionales.

R tiene Rd, su propio formato de documentación similar a LaTeX, que se utiliza para suministrar documentación completa, tanto en línea en una variedad de formatos como en copia impresa.

Funciones de programación

R es un lenguaje interpretado; los usuarios suelen acceder a él a través de un intérprete de línea de comandos. Si un usuario escribe $2+2$ en el símbolo del sistema R y presiona enter, la computadora responde con 4, como se muestra a continuación:

```
> 2 + 2
[1] 4
```

Este cálculo se interpreta como la suma de dos vectores de elemento único, lo que da como resultado un vector de elemento único. El prefijo [1] indica que la lista de elementos que lo siguen en la misma línea comienza con el primer elemento del vector (una característica que es útil cuando la salida se extiende sobre múltiples líneas).

Las estructuras de datos de R incluyen vectores, matrices, marcos de datos (similares a tablas en una base de datos relacional) y listas. El sistema de objetos extensibles de R incluye objetos para (entre otros) modelos de regresión, series de tiempo y coordenadas geoespaciales. El tipo de datos escalares nunca fue una estructura de datos de R. En cambio, un escalar se representa como un vector de longitud uno.

Uno de los puntos fuertes de R es la facilidad con la que se pueden producir parcelas de calidad de publicación bien diseñadas, que incluyen símbolos matemáticos y fórmulas cuando es necesario. Se ha tenido mucho cuidado con los valores predeterminados para las opciones menores de diseño en los gráficos, pero el usuario conserva el control total.

El entorno R

R es un conjunto integrado de instalaciones de software para manipulación de datos, cálculo y visualización gráfica. Incluye

- una instalación efectiva de manejo y almacenamiento de datos,
- un conjunto de operadores para cálculos en matrices, en particular matrices,
- una colección grande, coherente e integrada de herramientas intermedias para el análisis de datos,
- instalaciones gráficas para el análisis de datos y visualización en pantalla o en copia impresa, y
- un lenguaje de programación bien desarrollado, simple y efectivo que incluye condicionales, bucles, funciones recursivas definidas por el usuario e instalaciones de entrada y salida.

El término "medio ambiente" tiene la intención de caracterizarlo como un sistema completamente planificado y coherente, en lugar de una acumulación incremental de herramientas muy específicas e inflexibles, como es frecuentemente el caso con otro software de análisis de datos.

Muchos usuarios piensan en R como un sistema de estadísticas. R se puede extender (fácilmente) a través de paquetes. Hay alrededor de ocho paquetes suministrados con la distribución R y muchos más están disponibles a través de la familia de sitios de Internet de CRAN que cubren una amplia gama de estadísticas modernas.

RStudio

RStudio es un entorno de desarrollo integrado (IDE) para R. Incluye una consola, editor de resaltado de sintaxis que admite la ejecución directa de código, así como herramientas para trazado, historial, depuración y gestión del espacio de trabajo.

RStudio está disponible en ediciones comerciales y de código abierto y se ejecuta en el escritorio (Windows, Mac y Linux) o en un navegador conectado a RStudio Server o RStudio Server Pro (Debian / Ubuntu, RedHat / CentOS y SUSE Linux).

Funciones de RStudio IDE

Un IDE que fue construido solo para R

- Resaltado de sintaxis, finalización de código e indentación inteligente
- Ejecutar código R directamente desde el editor de origen
- Salte rápidamente a las definiciones de funciones

Junte su flujo de trabajo

- Ayuda y documentación integradas de R
- Gestione fácilmente múltiples directorios de trabajo usando proyectos
- Navegador de área de trabajo y visor de datos

Poderosa creación y depuración

- Depurador interactivo para diagnosticar y corregir errores rápidamente
- Extensas herramientas de desarrollo de paquetes
- Creación con Sweave y R Markdown

PAQUETES Y FUNCIONES UTILIZADOS EN EL PROYECTO

Utilización de paquetes en R para la realización del código

Prueba de Shapiro-Wilk

Dada una muestra aleatoria simple de tamaño n , (x_1, x_2, \dots, x_n) , se quiere saber si procede de una población con distribución normal. Este problema es muy frecuente, ya que son muchas las pruebas de inferencia estadística que exigen como condición imprescindible que la población de procedencia de la muestra sea normal.

El contraste que se desarrolla en esta sección recibe el nombre de Shapiro-Wilk. El método consiste en comenzar ordenando la muestra de menor a mayor, obteniendo el nuevo vector muestral $(x_{(1)}, x_{(2)}, \dots, x_{(n)})$, siendo $x_{(j)}$ el j -ésimo valor muestral tras la ordenación; a continuación, se calcula el estadístico de contraste

$$W = \frac{1}{nS^2} \left(\sum_{j=1}^h a_{jn} (x_{n-i+1} - x_i) \right)^2$$

Siendo s^2 la varianza muestral,

$$h = \begin{cases} \frac{n}{2} & , \text{si } n \text{ es par} \\ \frac{n-1}{2} & , \text{si } n \text{ es impar} \end{cases}$$

Obtenido el valor de W , su distribución permite calcular el valor crítico de la prueba que permite tomar una decisión sobre la normalidad de la muestra.

El programa devuelve el valor del estadístico W , así como su probabilidad crítica, p_c , la cual da la clave para aceptar o rechazar la hipótesis nula de normalidad, que representamos por H_0 .

H_0 : Si $p_c < 0.05$, se rechazará la hipótesis con un nivel de significación del 5%, si $p_c < 0.01$, se rechazará con un nivel del 1%.

La secuencia para realizar los contrastes de normalidad es:

1. Analiza
2. Estadísticos Descriptivos
3. Explorar

Prueba de Kolmogórov-Smirnov

La prueba de Kolmogórov-Smirnov es una prueba de bondad de ajuste que se utiliza para contrastar si un conjunto de datos se ajusta o no a una distribución, pero en este caso se utilizará para la distribución normal y para la de t-Student.

Cuando la prueba Kolmogórov-Smirnov se aplica para contrastar la hipótesis de normalidad de la población, el estadístico de prueba es supremo:

$$D = \sup |F_n(x) - F_0(x)|$$

Siendo $F_n(x)$ la función de distribución muestral y $F_0(x)$ la función teórica o correspondiente a la población normal especificada en la hipótesis nula.

Por lo tanto, el criterio para la toma de decisión entre las dos hipótesis será de la forma

Si $D \leq D_\alpha \Rightarrow$ No rechazar H_0

Si $D > D_\alpha \Rightarrow$ Rechazar H_0

Donde el valor D_α se elige de tal manera que:

$P(\text{Rechazar } H_0 | H_0 \text{ es cierta}) = P(D > D_\alpha | \text{Los datos siguen una distribución normal}) = \alpha$

Un caso específico de ajuste a una distribución teórica es la correspondiente a la distribución normal. Este contraste se realiza para comprobar si se verifica la hipótesis de normalidad necesaria para que el resultado de algunos análisis sea fiable.

Para comprobar la hipótesis nula de que la muestra ha sido extraída de una población con distribución de probabilidad normal se puede realizar un estudio gráfico y también un estudio analítico.

Prueba de Ljung-Box

Esta prueba permite probar en forma conjunta de que todos los coeficientes de autocorrelación son simultáneamente iguales a cero, esto es que son independientes, está definida como

$$LB = n(n + 2) \sum_{k=1}^m \left(\frac{\hat{\rho}_k^2}{n - k} \right) \sim X_{(m)}^2$$

donde n es el tamaño de la muestra, m la longitud del rezago.

H_0 : Las autocorrelaciones son independientes.

H_1 : Las autocorrelaciones no son independientes.

En una aplicación, si el Q calculada excede el valor Q crítico de la tabla ji cuadrada al nivel de significancia seleccionado, no se acepta la hipótesis nula de que todos los coeficientes de autocorrelación son iguales a cero; por lo menos algunos de ellos deben ser diferentes de cero.

Prueba de Dickey-Fuller

La prueba de Dickey-Fuller busca determinar la existencia o no de raíces unitarias en una serie de tiempo, es decir, si hay estacionariedad.

H_0 : $p < 0.05$ la serie de tiempo es estacionaria

H_1 : $p > 0.05$ existe una raíz unitaria en la serie $\rho = 1$

Expresándolo matemáticamente:

Sea $Y_t = \rho Y_{t-1} + e_t$

Donde Y_t es la serie de tiempo

ρ es el coeficiente

e_t es un ruido blanco, donde es no correlacionados, tienen varianza constante y valor esperado 0

Entonces,

$$\begin{aligned} Y_t &= \rho(\rho Y_{t-2} + e_{t-1}) + e_t = \rho^2 Y_{t-2} + \rho e_{t-1} + e_t \\ &= \rho^2(\rho Y_{t-3} + e_{t-2}) + \rho e_{t-1} + e_t = \rho^3 Y_{t-3} + \rho^2 e_{t-2} + \rho e_{t-1} + e_t \\ &\quad \vdots \end{aligned}$$

$$= \rho^t Y_0 + e_t + \rho e_{t-1} + \rho^2 e_{t-2} + \dots + \rho^{t-1} e_1$$

Su valor esperado es

$$E(Y_t) = \rho^t Y_0 \text{ ya que } E(e_t) = 0$$

$$\text{Si } \rho > 1 \Rightarrow E(Y_t) \xrightarrow[t \rightarrow \infty]{} +\infty$$

$$\text{Si } |\rho| < 1 \Rightarrow E(Y_t) \xrightarrow[t \rightarrow \infty]{} 0$$

Para calcular su varianza se pone $\sigma^2 = V(e_t) \forall t$ suponiendo que son homogéneos

$$\begin{aligned} V(Y_t) &= \sigma^2 + \rho\sigma^2 + \rho^2\sigma^2 + \dots + \rho^{t-1}\sigma^2 \\ &= \sigma^2(1 + \rho + \rho^2 + \dots + \rho^{t-1}) && \dots (*) \\ &= \sigma^2 \frac{1 - \rho^t}{1 - \rho} && \dots (+) \end{aligned}$$

$$\text{Si } \rho = 1 \Rightarrow V(Y_t) = t\sigma^2 \text{ si } t \rightarrow \infty \Rightarrow V(Y_t) \rightarrow \infty, \text{ (para (*))}$$

$$\text{Si } |\rho| < 1 \Rightarrow V(Y_t) = \frac{\sigma^2}{1 - \rho}, \text{ (para (+))}$$

$$\text{Si } \rho > 1 \Rightarrow V(Y_t) \text{ no converge, (para (+))}$$

Como $\rho^t Y_0$ es una constante, entonces $V(\rho^t Y_0) = 0$

Planteándolo de otra forma,

$$\text{Sea } Y_t = \rho Y_{t-1} + e_t$$

Donde Y_t es la serie de tiempo

ρ es el coeficiente

e_t es el error

Aplicando el operador backshift para el caso de que $|\rho| < 1$

$$BY_t = Y_{t-1}$$

$$IY_t - \rho BY_t = e_t$$

$$(I - \rho B)Y_t = e_t$$

Pensando que el operador B tiene inversa, entonces Y_t lo expresamos como

$$Y_t = (I - \rho B)^{-1} e_t$$

$$Y_t = (I + \rho B + \rho^2 B^2 + \rho^3 B^3 + \dots) e_t$$

$$Y_t = e_t + \rho e_{t-1} + \rho^2 e_{t-2} + \rho^3 e_{t-3} + \dots$$

Su esperanza y varianza son

$$E(Y_t) = 0$$

Para la varianza se pone $\sigma^2 = V(e_t) \forall t$

$$\begin{aligned} V(Y_t) &= \sigma^2 + \rho\sigma^2 + \rho^2\sigma^2 + \dots \\ &= \sigma^2(1 + \rho + \rho^2 + \dots) \\ &= \frac{\sigma^2}{1 - \rho} \quad \text{si } |\rho| < 1 \end{aligned}$$

Función auto.arima()

La función `auto.arima()` en R se encuentra en la librería `forecast` y su autor es Rob J. Hyndman, proporciona una opción rápida para construir pronósticos con series temporales, debido a que evalúa entre todos los posibles modelos, al mejor modelo considerando diversos criterios: estacionariedad, estacionalidad, diferencias, entre otras. Devuelve el mejor modelo ARIMA tomando en cuenta el valor AIC, AICc o BIC. Esta función combina pruebas de raíz unitaria, minimización de AICc y MLE para obtener un modelo ARIMA.

Procedimiento de modelación

Al ajustar un modelo ARIMA a un conjunto de datos de series de tiempo (no estacionales), se realizará lo siguiente:

1. Graficar los datos e identificar cualquier observación inusual.
2. Si es necesario, transformar los datos para estabilizar la varianza.
3. Si los datos no son estacionarios, se toman las primeras diferencias de los datos hasta que los datos sean estacionarios.
4. Examinar el ACF/PACF.
5. Probar los modelos elegidos y usar el AICc para buscar un modelo mejor.
6. Verificar los residuos del modelo elegido trazando el ACF de los residuos. Si los residuos no se ven como ruido blanco, entonces se probará un modelo modificado.
7. Una vez que los residuos parezcan ruido blanco, calcular los pronósticos.

Criterio de selección del modelo

¿Qué criterio se puede utilizar para seleccionar el mejor modelo? En la elección de un criterio para la selección de modelos, se acepta el hecho de que los modelos solo se aproximan a la realidad. Dado un conjunto de datos, el objetivo es determinar cuál de los modelos candidatos se aproxima mejor a los datos. Esto implica tratar de minimizar la pérdida de información. Debido a que el campo de la teoría de la

información se usa para cuantificar o medir el valor esperado de la información, el enfoque de la teoría de la información se usa para derivar los dos criterios más comúnmente utilizados en la selección de modelos: el criterio de información de Akaike y el criterio de información bayesiano. Los criterios, se pueden utilizar para la selección de modelos econométricos.

Akaike Information Criterion (AIC)

En 1951, Kullback y Leibler desarrollaron una medida para capturar la información que se pierde cuando se aproxima a la realidad; es decir, la medida de Kullback y Leibler es un criterio para un buen modelo que minimiza la pérdida de información. Dos décadas más tarde, Akaike estableció una relación entre la medida de Kullback-Leibler y el método de estimación de máxima probabilidad, un método de estimación utilizado en muchos análisis estadísticos, para derivar un criterio (es decir, una fórmula) para la selección del modelo. Este criterio, conocido como el criterio de información de Akaike (Akaike Information Criterion), generalmente se considera el primer criterio de selección de modelo que debe usarse en la práctica. El AIC es:

$$AIC = -2 \log L(\hat{\theta}) + 2k$$

Donde: θ es el conjunto (vector) de los parámetros del modelo

$L(\hat{\theta})$ es la probabilidad del modelo candidato dados los datos cuando se evalúa en la estimación de la probabilidad máxima de θ

k es el número de parámetros estimados en el modelo candidato.

El AIC en el aislamiento no tiene sentido. Más bien, este valor se calcula para cada modelo candidato y el modelo "mejor" es el modelo candidato con el AIC más pequeño. Veamos los dos componentes del AIC. El primer componente, $-2 \log L(\hat{\theta})$, es el valor de la función de probabilidad, $\log L(\hat{\theta})$, que es la probabilidad de obtener los datos según el modelo candidato. Como el valor de la función de verosimilitud se multiplica por -2 , ignorando el segundo componente, el modelo con el AIC mínimo es el que tiene el valor más alto para la función de verosimilitud. Sin embargo, a este primer componente agregamos un ajuste basado en el número de parámetros estimados. Cuantos más parámetros, mayor será la cantidad agregada al primer componente, lo que aumenta el valor para el AIC y penaliza el modelo. Por lo tanto, hay un compromiso: el mejor ajuste, creado al hacer que un modelo sea más complejo al requerir más parámetros, debe considerarse a la luz de la penalización impuesta al agregar más parámetros. Esta es la razón por la cual el segundo componente del AIC está pensado en términos de una penalización.

Para tamaños de muestra pequeños, el criterio de información de Akaike de segundo orden (AICc) debe utilizarse en lugar del AIC descrito anteriormente. El AICc es

$$AICc = AIC + \frac{2k(k+1)}{n-k-1} = -2 \log L(\hat{\theta}) + 2k + \frac{2k(k+1)}{n-k-1}$$

Donde n es el número de observaciones. Un pequeño tamaño de muestra es cuando n/k es menor que 40. Note que a medida que n aumenta, el tercer término en AICc se acerca a cero y, por lo tanto, dará el mismo resultado que AIC. También se ha sugerido que se use AICc en lugar de AIC cuando n es pequeño o k es grande. Se ha sugerido, por ejemplo, que, al seleccionar las órdenes de un ARMA, se utilice el AICc.

Bayesian Information Criterion (BIC)

El criterio de información bayesiano (Bayesian Information Criterion), propuesto por Schwarz y, por lo tanto, también denominado criterio de información de Schwarz y criterio de información bayesiano de Schwarz es otro criterio de selección de modelo basado en la teoría de la información, pero establecido dentro de un contexto bayesiano. La diferencia entre el BIC y el AIC es la mayor penalización impuesta para el número de parámetros por el BIC que por el AIC. Burnham y Anderson proporcionan argumentos teóricos a favor del AIC, particularmente el AICc sobre el BIC.

Además, en el caso del análisis de regresión multivariante, Yang explica por qué el AIC es mejor que el BIC en la selección del modelo.

El BIC se calcula de la siguiente manera:

$$BIC = -2 \log L(\hat{\theta}) + k \log(n)$$

Donde: θ es el conjunto (vector) de los parámetros del modelo

$L(\hat{\theta})$ es la probabilidad del modelo candidato dados los datos cuando se evalúa en la estimación de la probabilidad máxima de θ

k es el número de parámetros estimados en el modelo candidato.

n es el número de observaciones

Pronóstico, Función forecast

Cuando lo que se quiere predecir no son eventos cotidianos, sino el comportamiento de un precio en donde se está pensando invertir o las ventas de una empresa para el próximo año, no se puede dejar guiar por la intuición. Es aquí donde entra en juego el pronóstico.

Pronóstico es el proceso de hacer predicciones sobre el futuro mediante el análisis estadístico de tendencias observadas en datos históricos. Dicho de otra manera: utilizar el pasado para conocer el futuro.

Se pueden encontrar técnicas de pronóstico muy simples, y otras extremadamente complejas. Obviamente, cuanto más simple sea nuestro modelo, más sencillo será obtener unas predicciones.

Los principales rasgos que ha de identificar un buen pronóstico son:

- Tendencias: incrementos/disminuciones a largo término en nuestros datos.
- Estacionalidad: cómo periodos concretos (vacaciones, día de la semana) afectan a nuestros datos.
- Puntos de inflexión: cambios acelerados en los valores de nuestros datos. Por ejemplo, puede ocurrir cuando una compañía se expande a nuevos mercados.

En R existen un par de paquetes –forecast, prophet– que permiten predecir el futuro (en un cierto intervalo de confianza).

Pero nos enfocaremos en el paquete forecast, desarrollado por Rob J. Hyndman, y que es la principal referencia de los usuarios de R a la hora de adentrarse en el mundo del pronóstico. Este paquete es muy completo, y para predecir series temporales estacionarias tenemos a disposición los modelos ARIMA (p,d,q). En forecast podemos emplear la función.

Este tipo de herramientas combinan técnicas de regresión y series temporales para encontrar los patrones que caracterizan al modelo. Las regresiones buscan deducir las relaciones entre la variable objeto del pronóstico y otras variables explicativas. Las series temporales en cambio usan los datos históricos recogidos a intervalos regulares de la variable objeto para deducir sus futuros valores.

Los pronósticos de puntos se pueden calcular usando los siguientes pasos:

1. Expandir la ecuación de ARIMA para que y_t esté en el lado izquierdo y todos los demás términos estén en el derecho.
2. Reescribir la ecuación reemplazando t con $T + h$.
3. En el lado derecho de la ecuación, reemplazar las observaciones futuras con sus pronósticos, los errores futuros con cero y los errores pasados con los residuos correspondientes.

Comenzando con $h = 1$, estos pasos se repiten para $h = 2, 3, \dots$ hasta que se hayan calculado todos los pronósticos.

Se ejemplificará con un modelo ARIMA(3,1,1)

$$(1 - \widehat{\phi}_1 B - \widehat{\phi}_2 B^2 - \widehat{\phi}_3 B^3)(1 - B)y_t = (1 + \widehat{\theta}_1 B)\varepsilon_t,$$

Donde $\widehat{\phi}_1 = 0.0044, \widehat{\phi}_2 = 0.0916, \widehat{\phi}_3 = 0.3698$ y $\widehat{\theta}_1 = -0.3921$

Entonces ampliamos el lado izquierdo para obtener

$$[1 - (1 + \widehat{\phi}_1)B + (\widehat{\phi}_1 - \widehat{\phi}_2)B^2 + (\widehat{\phi}_2 - \widehat{\phi}_3)B^3 + \widehat{\phi}_3B^4]y_t = (1 + \widehat{\theta}_1B)\varepsilon_t$$

y aplicando el operador backshift da

$$y_t - (1 + \widehat{\phi}_1)y_{t-1} + (\widehat{\phi}_1 - \widehat{\phi}_2)y_{t-2} + (\widehat{\phi}_2 - \widehat{\phi}_3)y_{t-3} + \widehat{\phi}_3y_{t-4} = \varepsilon_t + \widehat{\theta}_1\varepsilon_{t-1}$$

Finalmente, nos movemos todos los términos distintos de y_t a la derecha:

$$y_t = (1 + \widehat{\phi}_1)y_{t-1} - (\widehat{\phi}_1 - \widehat{\phi}_2)y_{t-2} - (\widehat{\phi}_2 - \widehat{\phi}_3)y_{t-3} - \widehat{\phi}_3y_{t-4} + \varepsilon_t + \widehat{\theta}_1\varepsilon_{t-1}$$

Esto completa el primer paso. Si bien la ecuación ahora parece un ARIMA (4,0,1), sigue siendo el mismo modelo ARIMA (3,1,1) con el que comenzamos. No puede considerarse un ARIMA (4,0,1) porque los coeficientes no satisfacen las condiciones de estacionariedad.

Para el segundo paso, reemplazamos $T + 1$ en t

$$y_{T+1} = (1 + \widehat{\phi}_1)y_T - (\widehat{\phi}_1 - \widehat{\phi}_2)y_{T-1} - (\widehat{\phi}_2 - \widehat{\phi}_3)y_{T-2} - \widehat{\phi}_3y_{T-3} + \varepsilon_{T+1} + \widehat{\theta}_1\varepsilon_T$$

$$\widehat{y}_{T+1|T} = (1 + \widehat{\phi}_1)y_T - (\widehat{\phi}_1 - \widehat{\phi}_2)y_{T-1} - (\widehat{\phi}_2 - \widehat{\phi}_3)y_{T-2} - \widehat{\phi}_3y_{T-3} + \widehat{\theta}_1\varepsilon_T$$

Un pronóstico de y_{T+2}

$$\widehat{y}_{T+2|T} = (1 + \widehat{\phi}_1)y_{T+1|T} - (\widehat{\phi}_1 - \widehat{\phi}_2)y_T - (\widehat{\phi}_2 - \widehat{\phi}_3)y_{T-1} - \widehat{\phi}_3y_{T-2}$$

El proceso continúa de esta manera para todos los períodos de tiempo futuros. De esta manera, se puede obtener cualquier número de pronósticos de puntos.

Intervalos de predicción

El cálculo de los intervalos de predicción de ARIMA es difícil. Sin embargo, el primer intervalo de predicción es fácil de calcular. Si $\widehat{\sigma}$ es la desviación estándar de los residuales, entonces un intervalo de predicción del 95% viene dado por $\widehat{y}_{T+1|T} \pm 1.96\widehat{\sigma}$. Este resultado es cierto para todos los modelos ARIMA, independientemente de sus parámetros y órdenes.

Los intervalos de predicción de varios pasos para los modelos ARIMA(0,0,q) son fáciles de calcular. Se puede escribir el modelo como:

$$y_t = \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i}$$

Luego, la variación estimada del pronóstico se puede escribir como:

$$\hat{\sigma}_h = \hat{\sigma}^2 \left[1 + \sum_{i=1}^{h-1} \hat{\theta}_i^2 \right], \quad \text{para } h = 2, 3, \dots$$

Y un intervalo de predicción del 95% viene dado por

$$\hat{y}_{T+1|T} \pm 1.96\sqrt{\hat{\sigma}_h}$$

Los intervalos de predicción para los modelos ARIMA se basan en suposiciones de que los residuos no están correlacionados y que se distribuyen normalmente. Si alguna de estas suposiciones no se cumple, entonces los intervalos de predicción pueden ser incorrectos. Por este motivo, siempre se tiene que trazar el ACF y el histograma de los residuos para verificar los supuestos antes de producir intervalos de predicción.

En general, los intervalos de predicción de los modelos ARIMA aumenta a medida que aumenta el horizonte del pronóstico. Para los modelos estacionarios (es decir, con $d=0$) convergerán, de modo que los intervalos de predicción para horizontes largos son esencialmente iguales. Para $d>1$, los intervalos de predicción continuarán creciendo en un futuro.

Al igual que con la mayoría de los cálculos de intervalo de predicción, los intervalos basados en ARIMA tienden a ser demasiados estrechos. Esto ocurre porque solo se ha tenido en cuenta la variación de los errores. Además, el cálculo supone que los patrones históricos que se han modelado continuarán en el período de pronóstico.

CAPÍTULO 3

MARCO METODOLÓGICO

El procedimiento que se utilizó a la hora de la realización este proyecto fue: Primero que nada, elegir 4 distintas acciones mexicanas de la Bolsa Mexicana de Valores para analizar sus precios durante 2 años, para esto, se buscaron las acciones en Yahoo! Finance y se descargaron los archivos de cada acción desde el 30 de abril del 2014 al 30 de abril del 2017 y pasarla a un solo archivo de Excel y guardarlo en formato csv; de ahí, se comenzó a realizar el código en el lenguaje de programación R empezando por las librerías que se utilizarían próximamente para el código que fueron la “stats”, “tseries”, “LambertW” y “forecast”. Después se realizó el código donde se pueda leer el archivo con la función read.csv de R y leer cada precio de cada acción columna por columna para luego calcular el rendimiento de cada precio

```
x<-read.csv(file="precio.csv",header=TRUE,sep=",")
n<-750
for(col in 1:4){
  print("Acción")
  print(col)
  precio<-1:n
  for(i in 1:n){
    precio[i]<-x[i,col]
  }
  t<-1:n
  nl<-n-1
  Rt<-1:nl
  for(i in 1:nl)
  {
    Rt[i]<-log(precio[i+1]/precio[i],base=exp(1))
  }
}
```

Se calculó la media y la desviación estándar de cada rendimiento con las funciones “mean” y “sd” respectivamente, luego se analizan los rendimientos; para esto, hay distintas pruebas que se les pueden aplicar como son para la normalidad se utiliza la prueba de Shapiro Wilk (shapiro.test) o la de Kolmogórov-Smirnov (ks.test). Para la independencia se utilizó la prueba de Ljung-Box (Box.test) y, para la estacionariedad se utilizó la prueba de Dickey-Fuller (adf.test) esto para ver si siguen un movimiento browniano geométrico y la prueba de t-Student (ks.test.t) para verificar que siguen una distribución de t-Student.

Así se podrán comparar los resultados de la prueba para la normal y para la t-Student (comparando los p-value), y decidir si los rendimientos siguen una normal, o si se ajustan mejor, por ejemplo, a una t-Student.

```

print("*****Pruebas sobre los rendimientos*****")

print("*****Prueba de normalidad*****")
sh<-shapiro.test(Rt)
print(sh)
kol<-ks.test(Rt,"pnorm",mean=mu, sd=0)
print(kol)

print("*****Prueba de independencia*****")
bx<-Box.test(Rt, lag = 1, type = c("Box-Pierce","Ljung-Box"))
print(bx)

print("*****Prueba de estacionaridad*****")
est<-adf.test(Rt)
print(est)

```

Se analizan las pruebas y se comprueba que los rendimientos siguen más una t-Student que una normal. Ahora se trata de investigar cómo sacar los parámetros de la distribución t-Student con la función “fitdistr(Rt,t)”.

```

#Como el rendimiento no sigue una normal, entonces hacemos una t-sudent
print("*****Prueba de t-student*****")
kss<-ks.test.t(Rt,param=NULL)
print(kss)

#Fit.distr se usa para Ajuste de maxima probabilidad de distribuciones ur
#mantengan fijos si se desea.
fi<-fitdistr(Rt,"t")
print(fi)
m<-fi$estimate[1]
print(m)
s<-fi$estimate[2]
print(s)
df<-fi$estimate[3]
print(df)

```

Para dibujar las gráficas de los rendimientos se utilizó la función “plot”, luego se realizó un histograma de los rendimientos con la función que se escribió anteriormente utilizando los parámetros que se calcularon, esto para ver si siguen la distribución t-Student, posteriormente, se graficó la t-Student y la normal sobrepuestas en las gráficas del histograma de los rendimientos para notar la diferencia y ver cual se aproxima mejor.

```

print("*****Grafica de los rendimientos*****")
x11()
plot(Rt,type="lw",col="black",xlab="Tiempo",main=paste("Grafica de rendimientos, acción",toString(col)))
x11()
hist(Rt,ylim=c(0,35),xlim=c(-.15,.12),breaks=50,freq=FALSE,main=paste("Histograma de los rendimientos, acción",toString(col)))
#x11()
h<-seq(-0.12,0.12,l=50) #limite de la x
vt<-1:50 #los tamanos de puntos donde vt son los valores de t
for(i in 1:50){
  vt[i]<-tst(h[i],df,m,s)
}

curve(dnorm(x,mean(Rt),sd(Rt)),col="blue",lwd=2,add=T) #grafica de la distri normal
lines(h,vt,col="red",type="lw",lwd=2) #grafica de la distri t-student
legend("topleft",lty=1,lwd=2,col=c("blue","red"),legend=c("Distribución normal","Distribución t-Student"),bty="n")

```

Al tener los parámetros, se podrán utilizar para simular valores que serán los rendimientos generados con un tamaño de muestra de 100,000 y de ahí poder realizar

una predicción de los precios de las acciones, ya sea el valor del día siguiente, de un mes, o tantos como uno quiera.

```
print("*****Prediccion de precios usando la t-student*****")
tammuestra<-100000
muestra<-rt(tammuestra,df=df)*s+m #rt(N, df=df)*s + m
print("Precio inicial")
pinicial<-precio[n]
print(pinicial)
prediccion<-pinicial*exp(muestra)
x11()
hist(prediccion,freq=FALSE,breaks=50,main=paste("Histograma de predicciones, acción",toString(col)))
```

Ya que se sabe que se aproxima mejor con la t-Student, se le aplica el modelo ARIMA utilizando la función "auto.arima" para ver qué ARIMA se ajusta a los rendimientos; en este caso como en la acción 3 dio un ARIMA(1,0,2) se tuvo que ver si los residuos del modelo ARIMA también seguían una distribución t-Student. Ya que se vio qué ARIMA se ajustó mejor a cada acción se prosiguió a realizar el pronóstico a tantos periodos como se quiera con la función "forecast".

```
print("***Proyeccion de los rendimientos***")
if(col==3){
  res50<-arima(Rt,order=c(1,0,2),method="CSS")
  resi<-residuals(res50)
  kss<-ks.test.t(resi,param=NULL)
  print(kss)
  sh<-shapiro.test(resi)
  print(sh)
  Rtarima<-Arima(Rt,order=c(1,0,2),seasonal=c(1,0,2),method=c("CSS-ML","ML","CSS"))
  x11()
  tsvdiag(Rtarima)
  print(Rtarima)
  forecastRt<-forecast(Rtarima,h=20,level=c(99.5),main=paste("Pronostico de los rendimientos, acción",toString(col)))
  x11()
  plot(forecastRt)
}
else{
  Rtarima<-Arima(Rt,order=c(0,0,0),seasonal=c(0,0,0),method=c("CSS-ML","ML","CSS"))
  x11()
  tsvdiag(Rtarima)
  print(Rtarima)
  forecastRt<-forecast(Rtarima,h=20,level=c(99.5),main=paste("Pronostico de los rendimientos, acción",toString(col)))
  x11()
  plot(forecastRt)
}
}
```

Luego para saber si los pronósticos se aproximan a los precios reales se realiza una gráfica donde se comparan los precios reales y los precios simulados tanto con la normal como para la t-Student, para ver qué tan preciso es el modelo presentado.

```

print("****Proyección de precios****")
print("****Precios reales****")
precio741<-precio[n-9]
precio742<-precio[n-8]
precio743<-precio[n-7]
precio744<-precio[n-6]
precio745<-precio[n-5]
precio746<-precio[n-4]
precio747<-precio[n-3]
precio748<-precio[n-2]
precio749<-precio[n-1]
precio750<-precio[n]
preal<-c(precio741,precio742,precio743,precio744,precio745,precio746,precio747,precio748,precio749,precio750)
print(preal)

print("****Precios simulados con la t-Student****")
s0<-precio[n-9]
ss<-c(0,0,0,0,0,0,0,0,0,0)
nr<-10000
for(j in 1:nr){
  s0<-precio[n-9]
  tamañomuestra<-10
  predict<-rt(tamañomuestra,df=df)*s+m
  ss[1]=ss[1]+s0
  for(i in 2:10){
    s0<-precio[n-10+i]*exp(predict[i])
    ss[i]=ss[i]+s0
  }
}
suma<-0
for(i in 1:10){
  ss[i]=ss[i]/nr
  suma<-suma+(s0^2-precio[n-9+i]^2)
}
print(ss)

print("****Precios simulados con la normal****")
ssl<-c(0,0,0,0,0,0,0,0,0,0)
nr<-10000
for(j in 1:nr){
  s0<-precio[n-9]
  tamañomuestra<-10
  predinorm<-rnorm(tamañomuestra,mu,o)
  ssl[1]=ssl[1]+s0
  for(i in 2:10){
    s0<-precio[n-10+i]*exp(predinorm[i])#s0 es el precio simulado para [n-6+i] osea uno hacia adelante
    ssl[i]=ssl[i]+s0
  }
}
sumal<-0
for(i in 1:10){
  ssl[i]=ssl[i]/nr
  sumal<-sumal+(s0^2-precio[n-9+i]^2)
}
print(ssl)

limsup<-c(17,13.9,111,99)
liminf<-c(15.5,13.1,105,94)
x11()
plot(preal,ylim=c(liminf[col],limsup[col]),xlab="Tiempo",ylab="Precio",type="b",col="black",lwd=2,main=paste(
lines(ss,ylim=c(liminf[col],limsup[col]),type="b",col="green",lwd=2)
lines(ssl,ylim=c(liminf[col],limsup[col]),type="b",col="pink",lwd=2)

legend("bottomleft",lty=4,lwd=2,col=c("black","green","pink"),legend=c("Precios reales","Precios simulados t-

```

Y por último se realizó una comparación de las varianzas de las distribuciones normal y t-Student.

```
print("***Comparación de las varianzas***")
varn<-o^2
print("Varianza normal")
print(varn)
vart<-s^2*(df/(df-2))
print("Varianza t-Student")
print(vart)
```

En el capítulo 4, se muestra la sección de resultados donde se encuentra el análisis de los mismos y de las gráficas.

CAPÍTULO 4

RESULTADOS

4.1. INTRODUCCIÓN

En la siguiente tabla se comparan los p-value que se obtienen al aplicar la prueba de bondad de ajuste para una normal de Shapiro Wilk y Kolmogórov-Smirnov a los rendimientos de las acciones y la prueba de bondad de ajuste para la t-Student generalizada que es la prueba de Kolmogórov-Smirnov.

Tabla 2. Resultados prueba de normalidad

	p-value con la normal (Shapiro-Wilk)	p-value con la normal (Kolmogorov-Smirnov)	p-value con la t-student (Kolmogorov)
Acción 1	6.844e-16	0.0002105	0.9101
Acción 2	2.2e-16	0.08605	0.907
Acción 3	4.488e-13	0.08088	0.9694
Acción 4	9.65e-12	0.1519	0.992

Para que se pueda rechazar una hipótesis nula en una prueba de hipótesis con nivel de significancia de 0.05, el p-value tiene que ser menor que 0.05 y como se puede ver en la tabla anterior al hacer la prueba con Shapiro Wilk todas arrojan que se puede rechazar que sigan una normal ya que son mucho menores que 0.05, sin embargo, con la de Kolmogórov-Smirnov no se podría rechazar a un 5% en las 3 últimas acciones, debido a que son mayores a 0.05, con esto, se puede concluir que la prueba de Shapiro Wilk es mejor ya que los p-value obtenidos son más pequeños que en el caso de Kolmogórov-Smirnov, un ejemplo de esto es que en la acción 2 con la primer prueba da el p-value 2.2e-16 y en la segunda prueba da 0.08605 lo que quiere decir que en este proyecto es mejor la prueba de Shapiro Wilk que la de Kolmogórov-Smirnov para ver si los datos siguen una normal, sin embargo, la prueba de Shapiro Wilk se utiliza más para probar la bondad de ajuste a una distribución normal y la de Kolmogórov-Smirnov se puede aplicar a cualquier distribución.

En cambio, al aplicar la prueba de bondad de ajuste de Kolmogórov-Smirnov para probar la hipótesis nula de que los rendimientos siguen una t-Student, no se puede rechazar que sigan una t-Student debido a que los p-value están muy altos, por ejemplo, en la acción 1 el p-value es de 0.9101, el de la acción 2 es 0.907, el de la acción 3 es 0.9694 y el de la acción 4 es 0.992, como se puede ver todas son cercana a 1 por lo que es muy difícil poder rechazar que los rendimientos siguen una distribución de t-student.

En resumen, se comprobó que no siguen una normal y que parecen ajustarse razonablemente bien a una t-Student.

También, en los resultados de la prueba de independencia quedaron los valores de la siguiente forma:

Tabla 3. Resultados prueba de independencia

	Acción 1	Acción 2	Acción 3	Acción 4
p-value	0.3636	0.7382	0.858	0.6191

Para que se pueda rechazar una hipótesis nula en una prueba de hipótesis con nivel de significancia de 0.05, el p-value tiene que ser menor que 0.05 y como se puede ver en la tabla anterior al hacer la prueba con Ljung-Box todas arrojan que se puede rechazar que sean independientes ya que son mucho mayores que 0.05.

Para los resultados de la prueba de estacionariedad en todos los casos los p-value dieron 0.01 lo que quiere decir que es difícil de rechazar que no exista estacionariedad.

A continuación, se mostrarán los resultados numéricos y las representaciones gráficas de los histogramas de los rendimientos para constatar que efectivamente los rendimientos parecen ajustarse a una t-Student que a una normal.

Primero se verá el comportamiento de los rendimientos de cada una de las acciones, luego se verán las comparaciones entre las 2 distribuciones y finalmente se mostrarán las predicciones de los precios utilizando la t-Student, así como la comparación entre los precios reales y los simulados.

4.2. RESULTADOS NUMÉRICOS

En la siguiente sección se pueden ver los resultados de las pruebas sobre los rendimientos de cada acción, como son la prueba de normalidad de Shapiro-Wilk y la de Kolmogorov-Smirnov, la prueba de independencia de Ljung-Box, la prueba de estacionariedad de Dickey-Fuller, la prueba de la t-Student no estandarizada con sus respectivos parámetros, los cuales se utilizaron para poder calcular las predicciones a futuro de acuerdo con el mejor ARIMA que se ajusta para cada acción.

ACCIÓN 1

En la primer acción se puede ver que los rendimientos tienen una media de 9.912841×10^{-5} y una varianza de 0.2134086 y en las pruebas que se les realizaron a los rendimientos, en la prueba de Shapiro Wilk el p-value dio 6.844×10^{-16} , como es un valor muy por debajo de 0.05 se puede rechazar la hipótesis nula de que sigan una distribución normal a un nivel de significancia del 5%; en la prueba de Kolmogórov-Smirnov dio 0.0002105 por lo cual también podemos rechazar la hipótesis nula de que sigan una normal.

Para ver si los rendimientos son independientes, se utilizó la prueba de Ljung-Box la cual dio el p-value de 0.3636, y para ver la estacionariedad se utilizó la prueba de Dickey-Fuller en la cual el p-value dio 0.01.

Como se vio que los rendimientos no siguen una normal, entonces se procede a ver si siguen una distribución t-Student por lo cual se puede ver que sí la siguen, ya que dio un valor de 0.9101. Como para este proyecto se utilizó la t-Student no estandarizada, se puede observar que los parámetros son: los grados de libertad ($df=3.150793$), el de localización ($m=0.0002109824$) y el de escala ($s=0.0139967071$). Estos parámetros se estiman con la función `fitdist()` de R. Ya que se tienen los parámetros se puede realizar una simulación de la predicción a un periodo hacia adelante utilizando la distribución t-Student no estandarizada, para realizar un histograma y ver si realmente sigue (se comportan como) tal distribución.

Luego se tiene que ver qué modelo ARIMA se ajusta mejor a los rendimientos con la cual se utiliza la función `auto.arima`, la que arrojó (0,0,0) como el mejor ARIMA lo que significa que sigue el modelo $\ln \frac{P_t}{P_{t-1}} = \mu + e_{t-1}$ donde e_{t-1} sigue una t-Student, para así después realizar las predicciones a 5 periodos hacia adelante y compararlos con los precios reales, pudiendo ver que es mínima la diferencia entre ambos.

Estos comentarios se sustentan en los resultados que se muestran a continuación.

```

[1] "Acción"
[1] 1
[1] "****mu y sigma****"
[1] 9.912841e-05
[1] 0.02134086
[1] "*****Pruebas sobre los rendimientos*****"
[1] "*****Prueba de normalidad*****"

      Shapiro-wilk normality test

data: Rt
w = 0.94602, p-value = 6.844e-16

      One-sample Kolmogorov-Smirnov test

data: Rt
D = 0.078194, p-value = 0.0002105
alternative hypothesis: two-sided

[1] "*****Prueba de independencia*****"

      Box-Pierce test

data: Rt
X-squared = 0.82555, df = 1, p-value = 0.3636

[1] "*****Prueba de estacionaridad*****"

      Augmented Dickey-Fuller Test

data: Rt
Dickey-Fuller = -8.7223, Lag order = 9, p-value = 0.01
alternative hypothesis: stationary

[1] "*****Prueba de t-student*****"

      One-sample Kolmogorov-Smirnov test student-t with df=3.15, location=0,
      scale=0.01

data: Rt
D = 0.020539, p-value = 0.9101
alternative hypothesis: two-sided

      m          s          df
0.0002109824  0.0139967071  3.1507931029
(0.0006209729) (0.0006863297) (0.4203054233)
      m
0.0002109824
      s
0.01399671
      df
3.150793
[1] "*****Grafica de los rendimientos*****"
[1] "*****prediccion de precios usando la t-student*****"
[1] "Precio inicial"
[1] 16.0192

```

```

[1] "****Modelo arima****"
Series: Rt
ARIMA(0,0,0) with zero mean

sigma^2 estimated as 0.0004548: log likelihood=1819.21
AIC=-3636.42 AICc=-3636.41 BIC=-3631.8

Training set error measures:
              ME          RMSE          MAE MPE MAPE          MASE          ACF1
Training set 9.912841e-05 0.02132684 0.01489847 100 100 0.7065131 0.03319953
[1] "****Proyeccion de los rendimientos****"
Series: Rt
ARIMA(0,0,0) with non-zero mean

Coefficients:
      mean
      1e-04
s.e. 8e-04

sigma^2 estimated as 0.0004554: log likelihood=1819.22
AIC=-3634.43 AICc=-3634.42 BIC=-3625.2
[1] "****Proyección de precios****"
[1] "****Precios reales****"
[1] 16.7500 16.7981 16.7789 16.6635 16.2211 16.0577 15.9038 15.7692 16.0385 16
.0192
[1] "****Precios simulados con la t-student****"
[1] 16.75000 16.80460 16.77931 16.67660 16.23492 16.06202 15.90700 15.77861
[9] 16.03970 16.02380
[1] "****Precios simulados con la normal****"
[1] 16.75000 16.80609 16.78646 16.66726 16.22222 16.06105 15.90870 15.77149
[9] 16.04630 16.02647
[1] "****Comparación de las varianzas****"
[1] "Varianza normal"
[1] 0.0004554325
[1] "Varianza t-Student"
      s
0.0005363822

```

ACCIÓN 2

En la segunda acción se puede ver que los rendimientos tienen una media de 5.758259×10^{-5} y una varianza de 0.01590887 y en las pruebas que se les realizaron a los rendimientos, por ejemplo, en la prueba de Shapiro Wilk el p-value dio 2.2×10^{-16} , como es un valor muy por debajo de 0.05 se puede rechazar la hipótesis nula de que sigan una distribución normal a un nivel de significancia del 5%; en la prueba de Kolmogórov-Smirnov dio 0.08605 por lo cual también podemos rechazar la hipótesis nula de que sigan una normal.

Para ver si los rendimientos son independientes, se utilizó la prueba de Ljung-Box la cual dio el p-value de 0.7382, y para ver la estacionariedad se utilizó la prueba de Dickey-Fuller en la cual el p-value dio 0.01.

Como se vio que los rendimientos no siguen una normal, entonces se procede a ver si siguen una distribución t-Student por lo cual se puede ver que sí la siguen, ya que dio un valor de 0.907. Como para este proyecto se utilizó la t-Student no estandarizada, se puede observar que los parámetros son: los grados de libertad ($df=6.025723$), el de localización ($m=-4.53872 \times 10^{-5}$) y el de escala ($s=0.01254725$). Estos parámetros se estiman con la función `fitdist()` de R. Ya que se tienen los parámetros se puede realizar una simulación de la predicción a un periodo hacia adelante utilizando la distribución t-Student no estandarizada, para realizar un histograma y ver si realmente sigue tal distribución.

Luego se tiene que ver qué modelo ARIMA se ajusta mejor a los rendimientos con la cual se utiliza la función `auto.arima`, la que arrojó (0,0,0) como el mejor ARIMA lo que significa que sigue el modelo $\ln \frac{P_t}{P_{t-1}} = \mu + e_{t-1}$ donde e_{t-1} sigue una t-Student, para así después realizar las predicciones a 5 periodos hacia adelante y compararlos con los precios reales, pudiendo ver que es mínima la diferencia entre ambos.

Estos comentarios se sustentan en los siguientes resultados.

```

[1] "Acción"
[1] 2
[1] "****mu y sigma****"
[1] 5.758259e-05
[1] 0.01590887
[1] "*****Pruebas sobre los rendimientos*****"
[1] "*****Prueba de normalidad*****"

```

Shapiro-wilk normality test

```

data: Rt
w = 0.93957, p-value < 2.2e-16

```

One-sample Kolmogorov-Smirnov test

```

data: Rt
D = 0.045827, p-value = 0.08605
alternative hypothesis: two-sided

```

```
[1] "*****Prueba de independencia*****"
```

Box-Pierce test

```

data: Rt
X-squared = 0.11169, df = 1, p-value = 0.7382

```

```
[1] "*****Prueba de estacionaridad*****"
```

Augmented Dickey-Fuller Test

```

data: Rt
Dickey-Fuller = -9.3037, Lag order = 9, p-value = 0.01
alternative hypothesis: stationary

```

```
[1] "*****Prueba de t-student*****"
```

One-sample Kolmogorov-Smirnov test student-t with df=6.03, location=0, scale=0.01

```

data: Rt
D = 0.020645, p-value = 0.907
alternative hypothesis: two-sided

```

m	s	df
-0.0000453872	0.0125472509	6.0257229038
(0.0005205550)	(0.0005022484)	(1.1515191936)

```

      m
-4.53872e-05

```

```

      s
0.01254725

```

```

      df
6.025723

```

```

[1] "*****Grafica de los rendimientos*****"
[1] "*****Prediccion de precios usando la t-student*****"
[1] "Precio inicial"
[1] 13.74

```

```

[1] "****Modelo arima****"
Series: Rt
ARIMA(0,0,0) with zero mean

sigma^2 estimated as 0.0002528: log likelihood=2039.23
AIC=-4076.46 AICc=-4076.45 BIC=-4071.84

Training set error measures:
              ME          RMSE          MAE MPE MAPE          MASE          ACF1
Training set 5.758259e-05 0.01589835 0.0115686 100 100 0.7052349 0.01221123
[1] "****Proyeccion de los rendimientos****"
Series: Rt
ARIMA(0,0,0) with non-zero mean

Coefficients:
      mean
      1e-04
s.e. 6e-04

sigma^2 estimated as 0.0002531: log likelihood=2039.23
AIC=-4074.47 AICc=-4074.45 BIC=-4065.23
[1] "****Proyección de precios****"
[1] "****Precios reales****"
[1] 13.30 13.45 13.76 13.76 13.47 13.54 13.58 13.65 13.79 13.74
[1] "****Precios simulados con la t-Student****"
[1] 13.30000 13.45322 13.76280 13.76309 13.46813 13.54052 13.58182 13.64907
[9] 13.79351 13.74322
[1] "****Precios simulados con la normal****"
[1] 13.30000 13.45391 13.76240 13.76145 13.47506 13.54028 13.58340 13.65107
[9] 13.79281 13.74026
[1] "****Comparación de las varianzas****"
[1] "Varianza normal"
[1] 0.000253092
[1] "Varianza t-Student"
      5
0.0002356473

```

ACCIÓN 3

En la segunda acción se puede ver que los rendimientos tienen una media de 0.0003198768 y una varianza de 0.01577441 y en las pruebas que se les realizaron a los rendimientos, por ejemplo, en la prueba de Shapiro Wilk el p-value dio 4.488×10^{-13} , como es un valor muy por debajo de 0.05 se puede rechazar la hipótesis nula de que sigan una distribución normal a un nivel de significancia del 5%; en la prueba de Kolmogórov-Smirnov dio 0.08088 por lo cual también podemos rechazar la hipótesis nula de que sigan una normal.

Para ver si los rendimientos son independientes, se utilizó la prueba de Ljung-Box la cual dio el p-value de 0.858, y para ver la estacionariedad se utilizó la prueba de Dickey-Fuller en la cual el p-value dio 0.01.

Como se vio que los rendimientos no siguen una normal, entonces se procede a ver si siguen una distribución t-Student por lo cual se puede ver que sí la siguen, ya que dio un valor de 0.9694. Como para este proyecto se utilizó la t-Student no estandarizada, se puede observar que los parámetros son: los grados de libertad ($df=5.2153603968$), el de localización ($m=0.0002894030$) y el de escala ($s=0.0123185492$). Estos parámetros se estiman con la función `fitdist()` de R. Ya que se tienen los parámetros se puede realizar una simulación de la predicción a un periodo hacia adelante utilizando la distribución t-Student no estandarizada, para realizar un histograma y ver si realmente sigue tal distribución.

Luego se tiene que ver qué modelo ARIMA se ajusta mejor a los rendimientos con la cual se utiliza la función `auto.arima`, la que arrojó (1,0,2) con los coeficientes $ar1=0.5562$, $ma1=-0.5736$ y $ma2=-0.0993$ como el mejor ARIMA lo que significa que sigue el modelo $\ln \frac{P_t}{P_{t-1}} = \mu + 0.5562X_{t-1} + 0.5736\varepsilon_{t-1} + 0.0993\varepsilon_{t-2} + e_{t-1}$ donde $\phi_1 X_{t-1}$ es el polinomio autoregresivo, $\theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$ es el polinomio de medias móviles y e_{t-1} es el error que sigue una t-Student, para así después realizar las predicciones a 5 periodos hacia adelante y compararlos con los precios reales, pudiendo ver que sí varían los valores entre ambos.

Estos comentarios se sustentan en los resultados que se muestran a continuación.

```

[1] "Acción"
[1] 3
[1] "****mu y sigma****"
[1] 0.0003198768
[1] 0.01577441
[1] "*****Pruebas sobre los rendimientos*****"
[1] "*****Prueba de normalidad*****"

```

Shapiro-wilk normality test

```

data: Rt
w = 0.96175, p-value = 4.488e-13

```

One-sample Kolmogorov-Smirnov test

```

data: Rt
D = 0.046276, p-value = 0.08088
alternative hypothesis: two-sided

```

```

[1] "*****Prueba de independencia*****"

```

Box-Pierce test

```

data: Rt
X-squared = 0.031997, df = 1, p-value = 0.858

```

```

[1] "*****Prueba de estacionaridad*****"

```

Augmented Dickey-Fuller Test

```

data: Rt
Dickey-Fuller = -9.4683, Lag order = 9, p-value = 0.01
alternative hypothesis: stationary

```

```

[1] "*****Prueba de t-student*****"

```

One-sample Kolmogorov-Smirnov test student-t with df=5.22, location=0, scale=0.01

```

data: Rt
D = 0.017945, p-value = 0.9694
alternative hypothesis: two-sided

```

m	s	df
0.0002894030	0.0123185492	5.2153603968
(0.0005178118)	(0.0005324497)	(0.9655733294)

```

m
0.000289403

```

```

s
0.01231855

```

```

df
5.21536

```

```

[1] "*****Grafica de los rendimientos*****"
[1] "*****Prediccion de precios usando la t-student*****"
[1] "Precio inicial"
[1] 110.21

```

```

[1] "****Modelo arima****"
Series: Rt
ARIMA(1,0,2) with zero mean

Coefficients:
      ar1      ma1      ma2
 0.5562 -0.5736 -0.0993
s.e. 0.1977 0.1987 0.0439

sigma^2 estimated as 0.0002454: log likelihood=2051.82
AIC=-4095.63 AICc=-4095.58 BIC=-4077.16

Training set error measures:
              ME          RMSE          MAE MPE MAPE          MASE          ACF
1
Training set 0.0004247656 0.01563287 0.01153178 NaN Inf 0.6999347 -0.000729144
7
[1] "****Proyeccion de los rendimientos****"

      One-sample Kolmogorov-Smirnov test student-t with df=5.47, location=0,
      scale=0.01

data: resi
D = 0.014572, p-value = 0.9973
alternative hypothesis: two-sided

      Shapiro-wilk normality test

data: resi
w = 0.96234, p-value = 5.911e-13

Series: Rt
ARIMA(1,0,2) with non-zero mean

Coefficients:
      ar1      ma1      ma2      mean
 0.5702 -0.5882 -0.0981 3e-04
s.e. 0.1957 0.1968 0.0444 4e-04

sigma^2 estimated as 0.0002455: log likelihood=2052.07
AIC=-4094.14 AICc=-4094.06 BIC=-4071.05
[1] "****Proyección de precios****"
[1] "****Precios reales****"
[1] 106.17 109.41 108.47 108.99 107.01 107.49 109.54 109.41 110.10 110.21
[1] "****Precios simulados con la t-Student****"
[1] 106.1700 109.4835 108.5182 109.0518 107.0463 107.5462 109.5812 109.4409
[9] 110.1188 110.2468
[1] "****Precios simulados con la normal****"
[1] 106.1700 109.4859 108.5219 109.0299 107.0529 107.5370 109.6022 109.4535
[9] 110.1175 110.2601
[1] "****Comparación de las varianzas****"
[1] "Varianza normal"
[1] 0.0002488321
[1] "Varianza t-Student"
      S
0.0002461352

```

ACCIÓN 4

En la cuarta acción se puede ver que los rendimientos tienen una media de 0.0001292002 y una varianza de 0.01501262 y en las pruebas que se les realizaron a los rendimientos, por ejemplo, en la prueba de Shapiro Wilk el p-value dio 9.65×10^{-12} , como es un valor muy por debajo de 0.05 se puede rechazar la hipótesis nula de que sigan una distribución normal a un nivel de significancia del 5%; en la prueba de Kolmogórov-Smirnov dio 0.1519 por lo cual también podemos rechazar la hipótesis nula de que sigan una normal.

Para ver si los rendimientos son independientes, se utilizó la prueba de Ljung-Box la cual dio el p-value de 0.6191, y para ver la estacionariedad se utilizó la prueba de Dickey-Fuller en la cual el p-value dio 0.01.

Como se vio que los rendimientos no siguen una normal, entonces se procede a ver si siguen una distribución t-Student por lo cual se puede ver que sí la siguen, ya que dio un valor de 0.992. Como para este proyecto se utilizó la t-Student no estandarizada, se puede observar que los parámetros son: los grados de libertad ($df=5.722476$), el de localización ($m=-4.15846 \times 10^{-5}$) y el de escala ($s=0.01201828$). Estos parámetros se estiman con la función `fitdist()` de R. Ya que se tienen los parámetros se puede realizar una simulación de la predicción a un periodo hacia adelante utilizando la distribución t-Student no estandarizada, para realizar un histograma y ver si realmente sigue tal distribución.

Luego se tiene que ver qué modelo ARIMA se ajusta mejor a los rendimientos con la cual se utiliza la función `auto.arima`, la que arrojó (0,0,0) como el mejor ARIMA lo que significa que sigue el modelo $\ln \frac{P_t}{P_{t-1}} = \mu + e_{t-1}$ donde e_{t-1} sigue una t-Student, para así después realizar las predicciones a 5 periodos hacia adelante y compararlos con los precios reales, pudiendo ver que es mínima la varianza que existe entre ambos.

Estos comentarios se sustentan en los siguientes resultados.

```

[1] "Acción"
[1] 4
[1] "****mu y sigma****"
[1] 0.0001292002
[1] 0.01501262
[1] "*****Pruebas sobre los rendimientos*****"
[1] "*****Prueba de normalidad*****"

```

Shapiro-wilk normality test

```

data: Rt
w = 0.96797, p-value = 9.65e-12

```

One-sample Kolmogorov-Smirnov test

```

data: Rt
D = 0.041478, p-value = 0.1519
alternative hypothesis: two-sided

```

```

[1] "*****Prueba de independencia*****"

```

Box-Pierce test

```

data: Rt
X-squared = 0.24713, df = 1, p-value = 0.6191

```

```

[1] "*****Prueba de estacionaridad*****"

```

Augmented Dickey-Fuller Test

```

data: Rt
Dickey-Fuller = -9.8322, Lag order = 9, p-value = 0.01
alternative hypothesis: stationary

```

```

[1] "*****Prueba de t-student*****"

```

One-sample Kolmogorov-Smirnov test student-t with df=5.72, location=0, scale=0.01

```

data: Rt
D = 0.015808, p-value = 0.992
alternative hypothesis: two-sided

```

m	s	df
0.0000415846	0.0120182815	5.7224763756
(0.0005007917)	(0.0005030210)	(1.1140424611)

```

m
4.15846e-05

```

```

s
0.01201828

```

```

df
5.722476

```

```

[1] "*****Grafica de los rendimientos*****"
[1] "*****Prediccion de precios usando la t-student*****"
[1] "Precio inicial"
[1] 94.54

```

```

[1] "****Modelo arima****"
Series: Rt
ARIMA(0,0,0) with zero mean

sigma^2 estimated as 0.0002251: log likelihood=2082.64
AIC=-4163.27 AICc=-4163.27 BIC=-4158.66

Training set error measures:
              ME          RMSE          MAE MPE MAPE          MASE          ACF1
Training set 0.0001292002 0.01500315 0.01119796 100 100 0.6905778 0.01816431
[1] "****Proyeccion de los rendimientos****"
Series: Rt
ARIMA(0,0,0) with non-zero mean

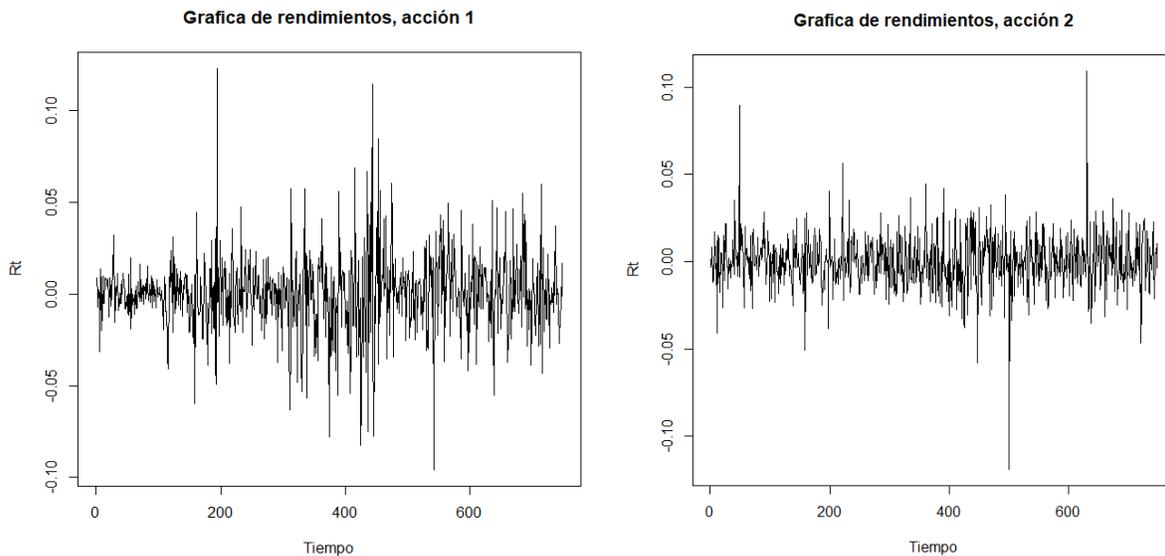
Coefficients:
      mean
      1e-04
s.e. 5e-04

sigma^2 estimated as 0.0002254: log likelihood=2082.66
AIC=-4161.33 AICc=-4161.31 BIC=-4152.09
[1] "****Proyección de precios****"
[1] "****Precios reales****"
[1] 96.53 96.84 97.77 98.42 96.90 96.55 97.10 97.41 95.63 94.54
[1] "****Precios simulados con la t-Student****"
[1] 96.53000 96.88015 97.76782 98.41729 96.89194 96.56223 97.12245 97.43006
[9] 95.66443 94.55409
[1] "****Precios simulados con la normal****"
[1] 96.53000 96.83177 97.76752 98.43381 96.93606 96.58137 97.10870 97.43142
[9] 95.67420 94.58782
[1] "****Comparación de las varianzas****"
[1] "Varianza normal"
[1] 0.0002253786
[1] "Varianza t-Student"
      5
0.0002220429

```

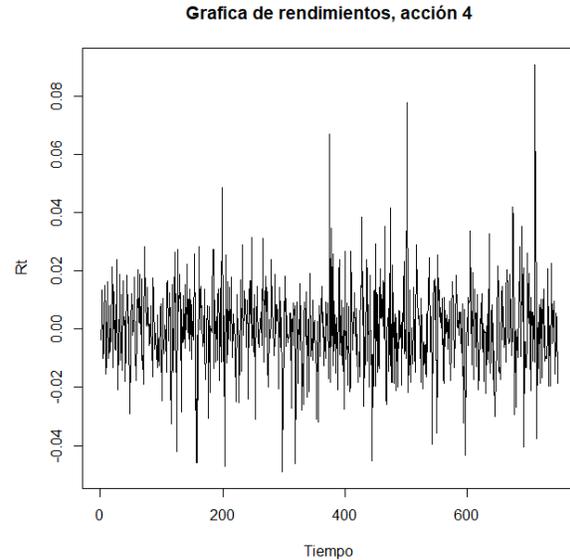
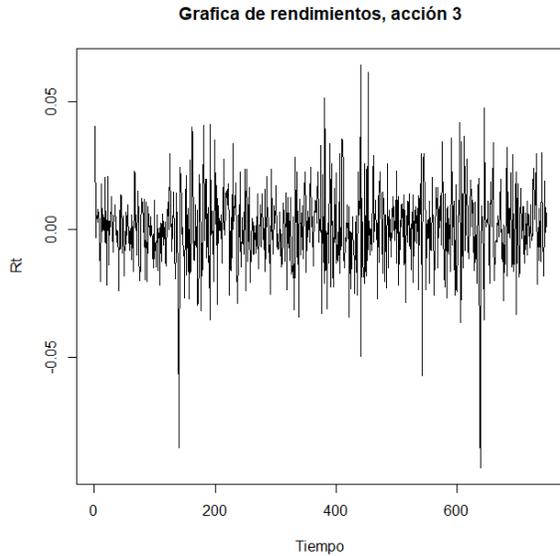
4.3. GRAFICA DE LOS RENDIMIENTOS DE LAS ACCIONES

En esta sección se ve el comportamiento de los rendimientos reales a lo largo del tiempo t , lo que se puede observar en cada una de las gráficas es que en algunos puntos tienen varios picos que sobresale de la gráfica, con esto se hace pensar que sigue una distribución t-Student ya que esta distribución tiene más valores en las colas tanto a la derecha como a la izquierda, en cambio si fuera una normal no tendría los picos tan altos sino que se mantendrían constantes cerca de la media cero. Así también se puede ver que todas parecen que sean estacionarias ya que las gráficas no crecen ni decrecen, sino que se mantienen constantes.



En la gráfica de la acción 1 se puede observar que tiene 3 picos que sobresalen, lo que se puede entender que no sigue una distribución normal, sino más bien una distribución t-Student al tener los picos tan altos y se puede ver también que es estacionaria ya que se mantiene constante en todo el periodo.

En la gráfica de la acción 2 se puede observar que tiene 3 picos que sobresalen, lo que se puede entender que no sigue una distribución normal, sino más bien una distribución t-Student al tener los picos tan altos y se puede ver también que es estacionaria ya que se mantiene constante en todo el periodo.



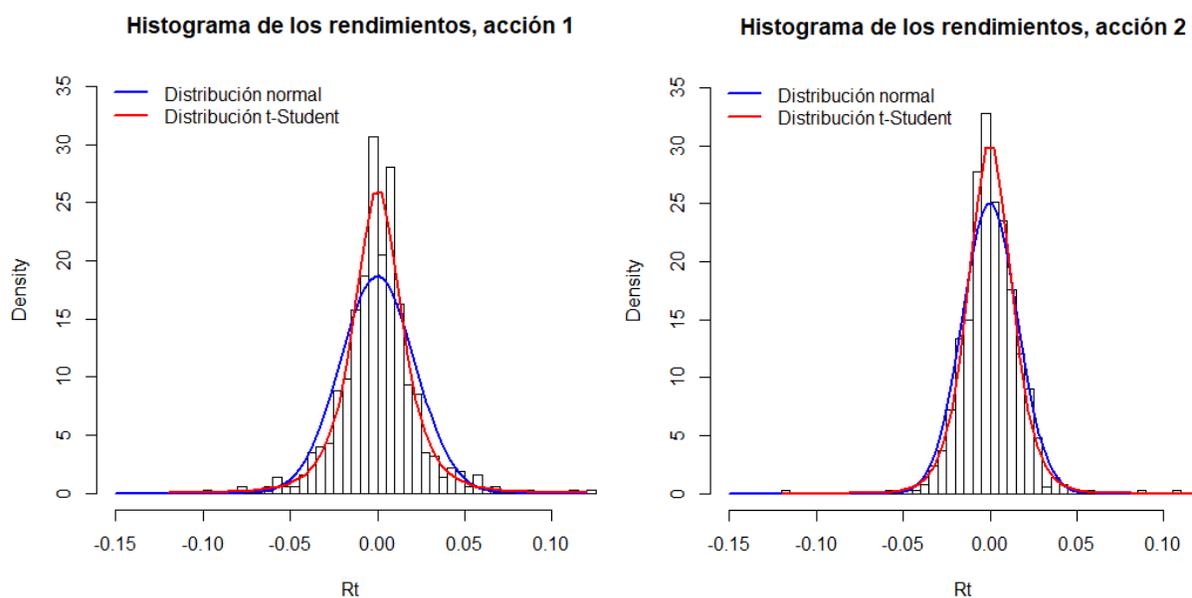
En la gráfica de la acción 3 se puede observar que tiene 4 o 6 picos que sobresalen, lo que se puede entender que no sigue una distribución normal, sino más bien una distribución t-Student al tener los picos tan altos y se puede ver también que es estacionaria ya que se mantiene constante en todo el periodo.

En la gráfica de la acción 4 se puede observar que tiene 4 picos que sobresalen en la parte de arriba y 6 o 7 que sobresalen en la parte de abajo, lo que se puede entender que no sigue una distribución normal, sino más bien una distribución t-Student al tener los picos tan altos y se puede ver también que es estacionaria ya que se mantiene constante en todo el periodo.

4.4. HISTOGRAMA DE LOS RENDIMIENTOS DE LAS ACCIONES

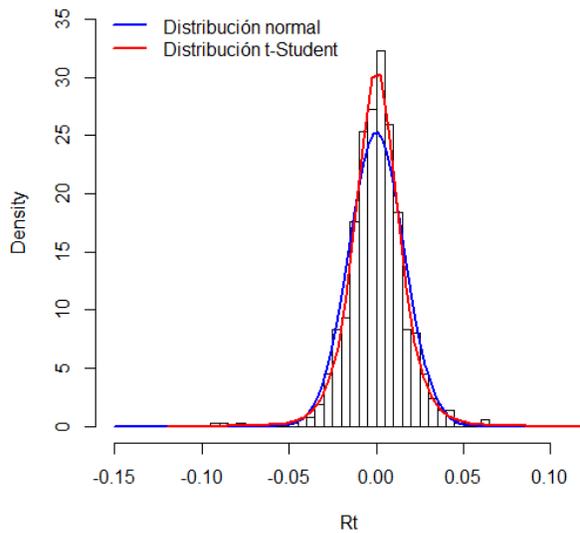
Se realizó un histograma de los rendimientos reales para cada una de las acciones y a este histograma se superpusieron las gráficas de la distribución normal y la distribución de t-Student que se ajustan mejor a los datos. En todos los casos la distribución t-Student se ajusta mejor a los datos que la normal siendo la curva de color rojo la distribución de t-Student y la curva de color azul la distribución normal.

En los casos analizados en este trabajo se puede comprobar que una alternativa para modelar los precios de las acciones es la distribución t-Student.

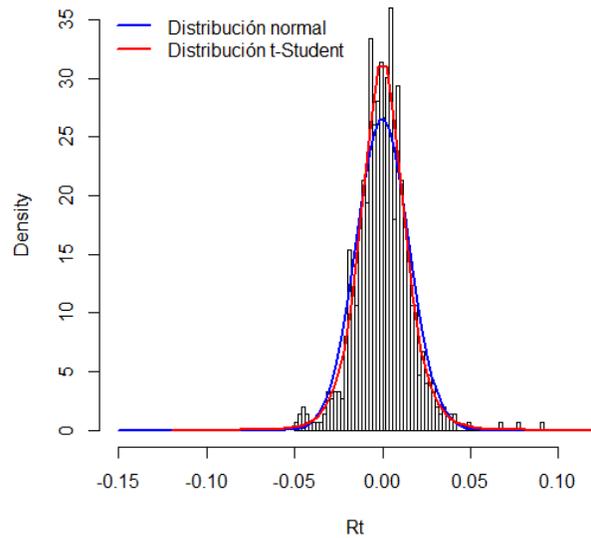


En el histograma de las acciones 1 y 2 se puede observar que al comparar las gráficas de la distribución normal (curva azul) y la distribución t-Student (curva roja), esta última se ajusta mejor al histograma de los rendimientos que la distribución normal, en la acción 1 hay una gran diferencia entre ambas curvas, donde se puede ver que, como se mencionó anteriormente, la curva roja se ajusta mejor a los rendimientos.

Histograma de los rendimientos, acción 3



Histograma de los rendimientos, acción 4



En el histograma de las acciones 3 y 4 se puede observar que al comparar las gráficas de la distribución normal (curva azul) y la distribución t-Student (curva roja), esta última se ajusta mejor al histograma de los rendimientos, aunque hay mínima diferencia entre ambas curvas de ambas gráficas.

4.5. PREDICCIONES DE LAS ACCIONES

HISTOGRAMA DE PREDICCIONES USANDO LA T-STUDENT

En esta sección se desea saber cómo se comportan las gráficas de los histogramas de las predicciones que se realizaron a los precios de las acciones, por lo cual primero se realiza una simulación de los precios usando la función de simulación de una muestra donde se utilizaron los parámetros de la distribución de t-Student no estandarizada, en este caso la muestra son los rendimientos, de ahí se toma como precio inicial al último precio que se tiene de cada precio de acción y luego se hace la predicción multiplicando el precio inicial por la muestra para así hacer un histograma confirmando que siga una distribución t-Student. Abajo se dará una explicación más detallada del procedimiento.

Para realizar el histograma de las predicciones de los precios se simularon valores con la t-Student para los rendimientos, que sería de la siguiente forma:

$$\ln \frac{P_t}{P_{t-1}} = \mu + e_{t-1} \text{ donde } e_{t-1} \text{ sigue una t-Student}$$

Y los precios se pueden simular de la siguiente forma

$$P_t = P_{t-1} e^{\mu + e_{t-1}}$$

Se supone que se tiene como precio inicial P_{100}

Entonces para la predicción a un periodo sería

$$P_{101} = P_{100} e^{\mu + e_{101}}$$

Se generan 10,000, para después hacer un histograma.

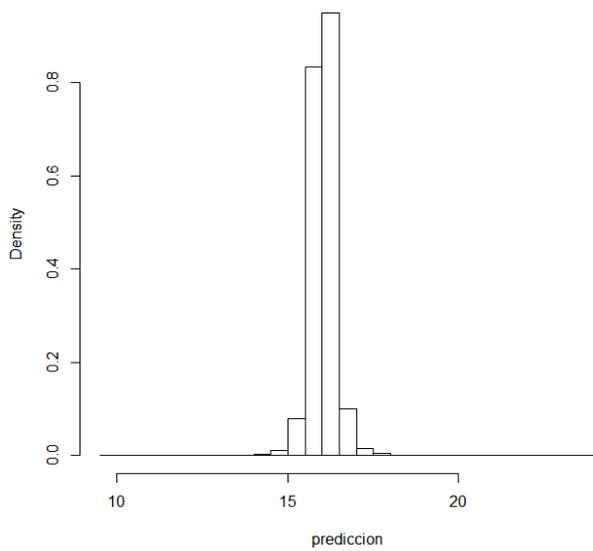
Si se quiere hacer una proyección a 102 sería

$$P_{102} = P_{101} e^{\mu + e_{101}} = P_{100} e^{\mu + e_{100}} e^{\mu + e_{101}} = P_{100} e^{2\mu + e_{100} + e_{101}}$$

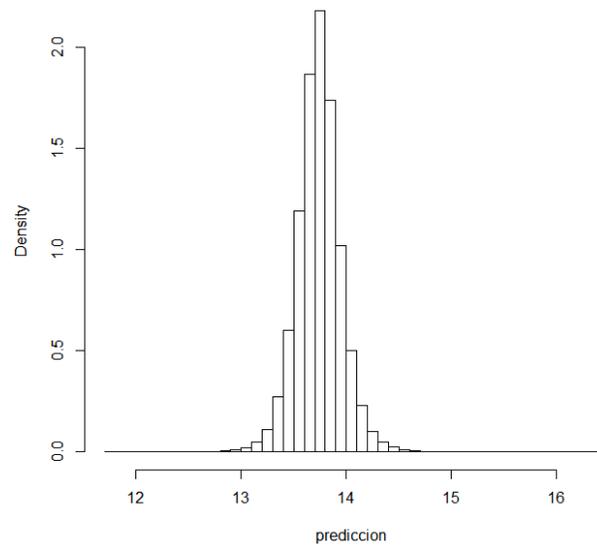
Donde los e_{100} y e_{101} son simulados como una t-Student y son independientes, μ es la media de los rendimientos y P_{100} es el precio en el periodo 100, y así con la fórmula anterior da un valor simulado para 2 periodos hacia adelante, esto es para el caso de que salga un ARIMA(0,0,0). Para el caso de que salga un ARIMA diferente, las predicciones serían

$$P_t = P_{t-1} e^{c + \phi_1 R_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + e_{t-1}}$$

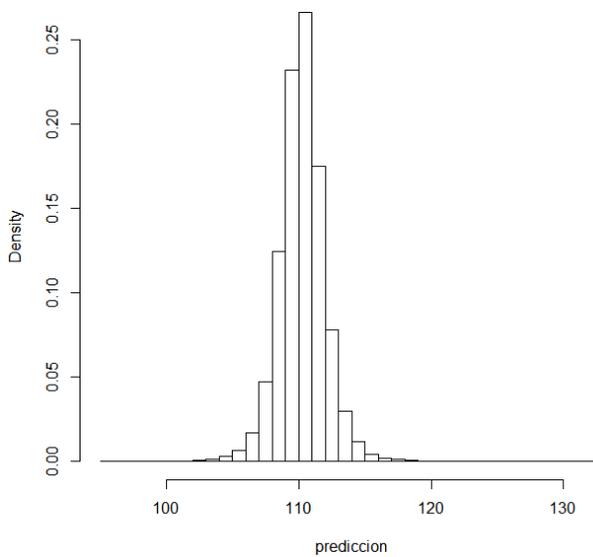
Histograma de predicciones, accion 1



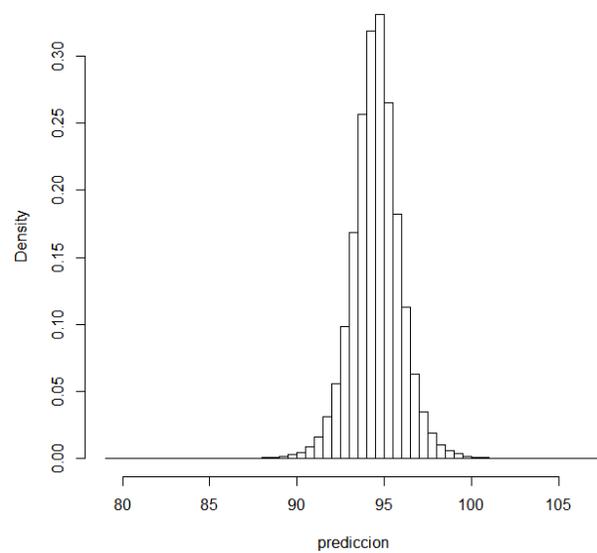
Histograma de predicciones, accion 2



Histograma de predicciones, accion 3



Histograma de predicciones, accion 4



Estos son los histogramas de los precios obtenidos por simulación con una muestra de 100,000 de un periodo hacia adelante que se realizaron empleando la distribución de t-Student.

Si se supone que

$X \sim f_X$ donde f_X es una función de densidad continua y X sigue una distribución t-Student.

Y también se tiene la función

$$Y = e^x$$

¿Cuál es la función de densidad de Y (f_Y)?

Entonces primero se tiene que encontrar la función de distribución de Y

$$F_Y(y) = P(Y \leq y) = P(e^X \leq y) = P(X \leq \ln(y)) = F_X(\ln(y)) \text{ con } y > 0$$

Pero se sabe que la función de densidad de una variable aleatoria es la derivada de la función de distribución, por lo que quedaría

$$f_Y(y) = F'_Y(y) = F'_X(\ln(y)) \frac{1}{y} = \frac{f_X(\ln(y))}{y} \dots (*)$$

Si se tiene que la función de densidad de la distribución t-Student es

$$f_X(x) = \left(\frac{1}{s}\right) \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{\left(\frac{y-\mu}{s}\right)^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

Ahora si se sustituye (*) en la ecuación anterior,

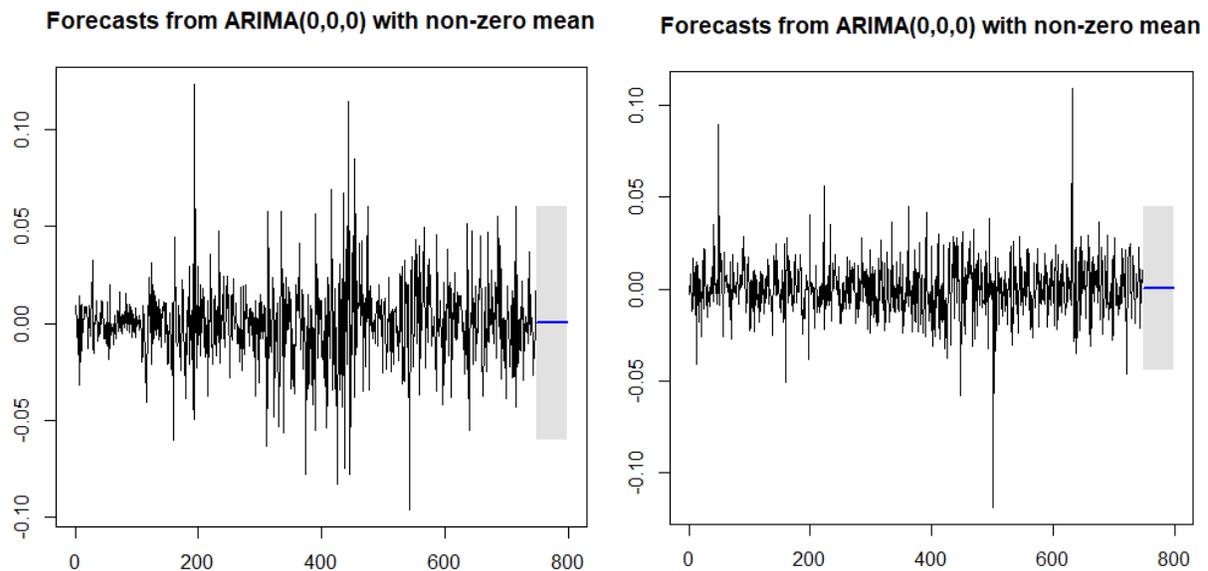
$$f_X(\ln(y)) = \left(\frac{1}{s}\right) \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma\left(\frac{\nu}{2}\right)} \frac{\left(1 + \frac{\left(\frac{\ln(y)-\mu}{s}\right)^2}{\nu}\right)^{-\frac{\nu+1}{2}}}{y}$$

Como bien se sabe, en los precios, en el caso de un movimiento browniano geométrico si la variable aleatoria X sigue una distribución normal, la función e^x sigue una log-normal, sin embargo, en este caso, X sigue una distribución t-Student, por lo que no se puede suponer que la función de densidad de los precios también siga una distribución t-Student.

PREDICCIÓN USANDO LA FUNCIÓN FORECAST

En las siguientes gráficas se puede observar que cuando el ARIMA dio (0,0,0), los pronósticos dan lo mismo, ya que no hay errores, por lo que para los ARIMA(0,0,0) el pronóstico siempre será la media, es decir, son constantes; y como se puede ver en la acción 3, al dar un ARIMA(1,0,2), el pronóstico si cambiará ya que toma los 2 errores anteriores para sacar las proyecciones a un periodo adelante, entonces se da a entender que mientras nos dé un ARIMA distinto al (0,0,0), pronosticarán valores no constantes dependiendo de las Medias Móviles.

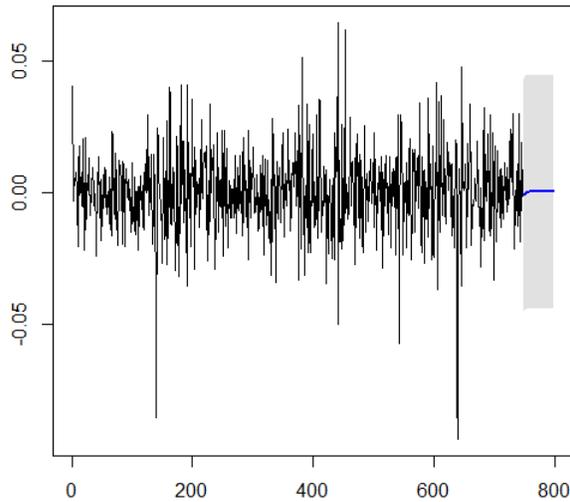
Los criterios que utiliza la función `auto.arima` son el AIC (Akaike Information Criterion), el BIC (Bayesian Information Criterion) y AICc que es lo mismo que el AIC, pero con una corrección para tamaños de muestra pequeños. Aunque todos son criterios de probabilidad penalizada, el criterio BIC tiende a ser mejor que AIC, aunque en ambos criterios se selecciona el que tenga valores pequeños, ya que entre más pequeños sean los valores se considera que el modelo es más eficiente.



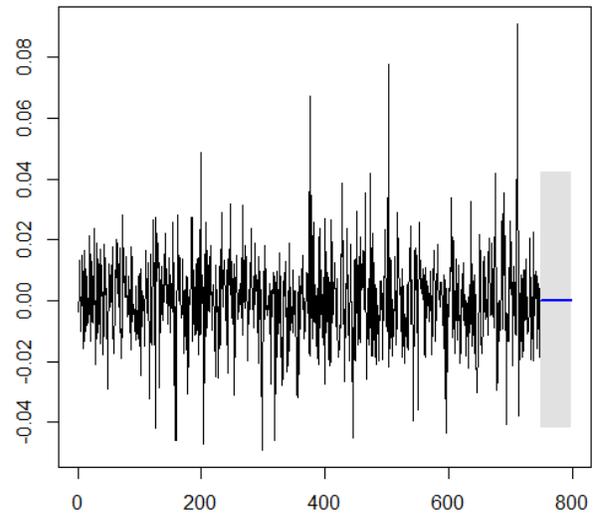
En el caso de la primera acción el modelo que mejor se ajustó a los residuos fue el ARIMA(0,0,0), se utiliza este modelo para hacer proyecciones de los rendimientos, los resultados se muestran en la gráfica de arriba.

Igualmente, para los rendimientos de la acción 2 el modelo que mejor se ajustó fue el ARIMA(0,0,0) y por lo tanto se utilizó este modelo para hacer proyecciones de los rendimientos, los resultados se muestran en la gráfica.

Forecasts from ARIMA(1,0,2) with non-zero mean



Forecasts from ARIMA(0,0,0) with non-zero mean



En la acción 3 el modelo que mejor se ajustó a los rendimientos fue el ARIMA(1,0,2) con media no cero, entonces se utilizó este modelo para hacer las proyecciones de los rendimientos y como se puede ver en esta gráfica los valores de los pronósticos al principio no son constantes, ya después conforme va avanzando el tiempo se van haciendo constantes.

Este es un modelo autoregresivo de orden 1 y de media móvil de orden 2, es decir,

$$R_t = c + \phi_1 R_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t = c + 0.1977 R_{t-1} + 0.1987 \varepsilon_{t-1} + 0.0439 \varepsilon_{t-2} + \varepsilon_t$$

En la acción 4, al igual que en la acción 1 y 2 el modelo que mejor se ajustó fue el ARIMA(0,0,0) con media no cero y éste fue el que se utilizó para realizar las proyecciones.

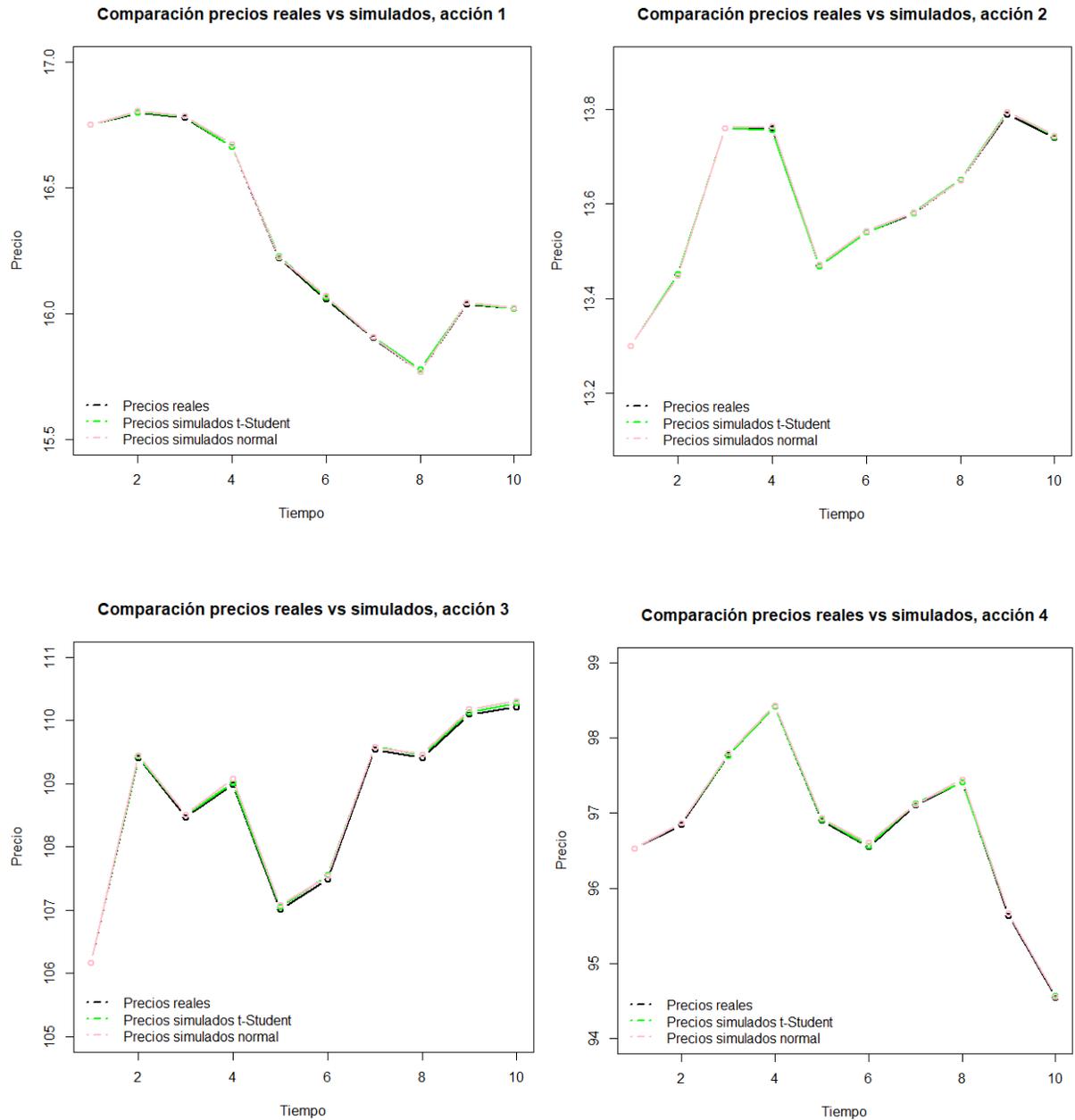
4.6. COMPARACIÓN DE LOS PRECIOS REALES VS PRECIOS SIMULADOS UTILIZANDO LA T-STUDENT Y LA NORMAL

Aquí se hace una comparación entre los precios reales contra los precios simulados utilizando la t-Student y la normal, en donde se puede observar que en la mayoría de los casos es mínima la diferencia entre estos. Para hacer la simulación de los precios primero se hizo una lista nula de 10 elementos, luego se toma un número para hacer las simulaciones (en este proyecto se eligieron 10,000), de ahí se seleccionó como precio inicial al precio del (n-9) elemento de la lista de precios, luego se realizó un ciclo en donde se toma un tamaño de muestra de 10 y se utiliza la función para realizar las proyecciones con la t-Student, después en cada uno se multiplica el precio inicial por la exponencial de la muestra de cada elemento (de 2 a 10) y de ahí se van actualizando los valores dentro de la lista que se creó al principio sumando el valor inicial con el que se calculó. De igual forma, esto se realizó para la normal modificando algunas partes.

A continuación, se muestra una parte del programa utilizando la t-Student, mostrando lo escrito anteriormente:

```
ss<-c (0,0,0,0,0,0,0,0,0,0)
nr<-10000
for (j in 1:nr) {
  s0<-precio[n-9]
  tamañomuestra<-10
  predict<-rt (tamañomuestra, df=df) *s+m
  ss [1] =ss [1] +s0
  for (i in 2:10) {
    s0<-precio[n-10+i] *exp(predict[i])
    ss[i]=ss[i]+s0
  }
}
for (i in 1:10) {
  ss[i]=ss[i]/nr
}
print(ss)
```

Por otra parte, se tomaron los 7 precios reales anteriores suponiendo que no se conocen, esto con el propósito de saber que tanto se ajustan los datos proyectados a los reales. Y finalmente se grafican los precios reales y los simulados en una misma gráfica para compararlos y ver que tanto la predicción se ajusta a los datos reales.



Como se puede ver en las 4 gráficas de las acciones, es mínima la diferencia que existe entre las predicciones utilizando la distribución normal y la distribución t-Student, en las acciones 3 y 4 se puede ver que se aproxima un poco más la distribución t-Student a los precios reales.

CONCLUSIONES

La presente tesis tiene como objetivo comprobar si los rendimientos siguen un movimiento Browniano geométrico. Si es así, entonces se prosigue a realizar un pronóstico de los precios; en caso contrario, se plantea un modelo alternativo, primero, utilizando la función auto.arima sobre los rendimientos y ver si siguen un $ARIMA(0,0,0)$; posteriormente hay que analizar en ambos casos, la distribución de los rendimientos; el siguiente paso es realizar los pronósticos a tantos periodos como se quiera, que en este caso fue a 7.

Así pues, la aportación principal de este trabajo consiste en proponer un modelo alternativo, en caso de que los rendimientos no sigan un movimiento Browniano geométrico, para realizar proyecciones en el tiempo de algunos precios de acciones.

También se comprobó que la prueba de Kolmogorov-Smirnov no es tan efectiva como la prueba de Shapiro Wilk, ya que los p-value obtenidos son más pequeños que en el caso de Kolmogorov-Smirnov, por lo que se concluye que la prueba de Shapiro es más exacta, precisa y efectiva para poder rechazar una prueba de hipótesis.

Se rechaza en todos los casos, a muy buen nivel de significancia (menor del 5%), que los rendimientos siguen una normal, y se mostró que sí son independientes y estacionarios. Por el hecho de no seguir una normal, entonces no siguen un movimiento Browniano geométrico. Sin embargo, si uno supone que los rendimientos siguen una t-Student, en ningún caso se puede rechazar que la sigan, ya que, el p-value da cerca de 1, lo que quiere decir que está muy lejos de poderse rechazar, por lo que el modelo es un éxito: es más preciso y es mucho mejor que el que utiliza la distribución normal para lograr las proyecciones.

Aunque algunos de los rendimientos se alejaron de la distribución normal con una prueba significativa, al realizar las simulaciones utilizando la distribución t-Student y la normal se generan prácticamente los mismos valores, quizá una de las razones es que la diferencia mayor entre las 2 distribuciones es en los extremos de la gráfica, ya que la t-Student puede tener mayor parte de su área en los extremos que la normal, aunque pueden ser ocasionales. Otra razón puede ser la diferencia que existe entre las varianzas de las distribuciones ya que al comparar las varianzas se ve que hay una ligera diferencia, siendo en la mayoría de los casos la varianza de la t-Student más pequeña. En general, suponer el modelo con la normal no es tan malo, en estos casos con ambos se obtiene casi los mismos resultados.

REFERENCIAS

Paul G. Hoel, Sidney C. Port, Charles J. Stone. (1971). Introduction to Probability Theory. Boston, USA: Houghton Mifflin Company.

Paul L. Meyer. (1986). Probabilidad y aplicaciones estadísticas. México, D.F.: SITESA.

Dennis D. Wackerly, William Mendenhall III, Richard L. Scheaffer. (2002) Estadística matemática con aplicaciones, México DF, Thomson Learning,

Gujarati, Damodar N. (2010). Econometría. México, D.F.: McGraw-Hill

Brockwell, P. J., & Davis, R. A. (2016). Introduction to time series and forecasting (3rd ed). New York, USA: Springer. [Amazon]

<https://definicion.mx/acciones/>

<http://www.finanzasparamortales.es/que-son-las-acciones/>

<https://finanzasyproyectos.net/que-son-las-acciones/>

<https://www.cuidatudinero.com/13098614/definicion-del-rendimiento-de-una-accion>

<https://www.bcr.com.ar/Publicaciones/Ediciones%20BCR/Archivos%20de%20cortes%20C3%ADa/Lecturas%2017/Rendimiento%20de%20las%20acciones%20en%20contextos%20inflacionarios.pdf>

http://www.estadisticas.gobierno.pr/iepr/LinkClick.aspx?fileticket=4_BxecUaZmg%3D

<https://www.cimat.mx/Eventos/vpec10/img/RogelioNotasR.pdf>

<https://www.rdocumentation.org/packages/graphics/versions/3.5.1/topics/legend>

<https://www.statmethods.net/advgraphs/parameters.html>

<https://otexts.org/fpp2/arma-r.html>

<https://www.quantinsti.com/blog/forecasting-stock-returns-using-arma-model/>

<http://www.ghahramani.ca/uploads/1/7/0/4/17042208/box-jenkins-r-seminar.pdf>

<https://stat.ethz.ch/R-manual/R-devel/library/graphics/html/legend.html>

<https://cran.r-project.org/doc/contrib/grafi3.pdf>

<https://repository.javeriana.edu.co/bitstream/handle/10554/20408/CalderonCamposJavierAndres2016.pdf?sequence=1>

<http://www.cmat.edu.uy/~mordecki/courses/upae/upae-curso.pdf>

<https://es.slideshare.net/torimatcordova/distribucion-t-de-student-28545004>
http://moodle2.unid.edu.mx/dts_cursos_md/lic/AE/E/AM/12/Distribucion_tStudent.pdf
http://riotorto.users.sourceforge.net/R/noparam_shapiro/
https://www2.ulpgc.es/hege/almacen/download/5/5015/Complemento_3_Prueba_de_Bondad_de_Ajuste_de_Kolmogorov_Smirnov.pdf
<http://biplot.usal.es/problemas/libro/5%20Contrastes.pdf>
<https://www.monografias.com/trabajos89/distribucion-normal-excel/distribucion-normal-excel.shtml>
<https://otexts.com/fpp2/arma-forecasting.html>
<https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118856406.app5>