



BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE
PUEBLA

FACULTAD DE CIENCIAS FÍSICO-MATEMÁTICAS

**Análisis y construcción de modelos para
pronosticar los precios de las acciones de Grupo
Financiero Inbursa**

Tesis presentada para obtener el título de

LICENCIADA EN ACTUARÍA

P R E S E N T A:

K A R L A C R U Z M O N T I E L

Directoras de tesis:

Dra. Daniela Cortés Toto

Dra. Hortensia Josefina Reyes Cervantes

Febrero, 2019
Puebla, Puebla

*Dedicado a
la memoria de
mis abuelos
y abuelita,
de quienes
aprendí mucho.*

Agradecimientos

Agradezco mucho a todas las personas que hicieron posible esta meta. A las personas cercanas amí, en especial a mi mamá y papá, gracias por ser quienes me inspiraron a perseguir los sueños que un día decidí alcanzar.

A mis hermanas, quienes son mis buenos momentos cuando todo se pone difícil y también cuando no lo son.

A mi novio, porque aparte de ser mi mejor amigo, me aconseja y me apoya en cualquier meta.

A mis amigos quienes a pesar de la distancia o falta de tiempo, siempre me brindaron los mejores lazos de amistad durante todo este trayecto.

A mis pequeños mentores, porque me enseñaron la razón, motivación y el gusto por seguir aprendiendo.

A todos los profesores de la facultad de quienes tuve la fortuna de recibir una pequeña semilla de su experiencia y conocimiento, también por su entusiasmo, visión y paciencia.

A mis directoras de tesis, por haber aceptado dirigir este trabajo, por todo el tiempo dedicado, consejos, comprensión y perseverancia, porque sin todo eso no lo hubiera logrado.

A mis sinodales por unirse a participar en esta recta final, el tiempo dedicado a la revisión y sugerencias para realizar un mejor trabajo.

Por último, quiero agradecer a V.I.E.P. por el apoyo que me brindo en la impresión de está tesis, en el proyecto Mínimos Cuadrados en la estimación de parámetros en procesos $ARX(p, q)$ con ruido correlacionado de orden uno.

A todos ustedes mil gracias.

Índice general

1. Introducción	9
1.1. Marco Contextual	9
1.2. Antecedentes	10
1.3. Planteamiento del problema	11
1.4. Justificación	11
1.5. Objetivo General	11
1.6. Breve descripción del contenido	11
2. Conceptos Básicos de Series de Tiempo	13
2.1. Elementos de Series de Tiempo	13
2.1.1. Componentes de una Serie de Tiempo	14
2.1.2. Modelos paramétricos para datos con tendencia	15
2.1.3. Autocorrelación en series de tiempo	16
2.1.4. Series Estacionarias	17
2.2. Operadores y Polinomio de Retraso	19
3. Metodología Box-Jenkins	23
3.1. Funciones de autocovarianza y autocorrelación	23
3.1.1. Matriz de covarianza de un proceso estacionario	27
3.1.2. Funciones Muestrales	28
3.2. Modelos Lineales Estacionarios	29
3.2.1. Modelo lineal general	29
3.2.2. Procesos Autorregresivos (AR)	31
3.2.3. Proceso de Medias Móviles (MA)	33
3.2.4. Procesos autorregresivos de medias móviles o AR- MA(p,q)	34
3.3. Modelos no estacionarios	35
3.3.1. Modelos ARIMA(p, d, q)	36
3.4. Modelación ARIMA	38
3.4.1. Identificación	38
3.4.2. Estimación	41
3.4.3. Validación	42
3.4.4. Predicción	44

4. Pronóstico del precio de las acciones de Grupo Financiero In-	
bursa	49
4.1. Aspectos Generales	49
4.2. Análisis estadístico	49
4.3. Análisis de los datos con metodología Box–Jenkins	49
4.3.1. Identificación del modelo	49
4.3.2. Estimación de parámetros	57
4.3.3. Verificación	58
4.3.4. Predicción	67
5. Resultados	73
5.0.1. Conclusión	74
5.0.2. Trabajos futuros	74
A. Código en R	75
B. Prueba Chi-Cuadrada	81
C. Estadístico de Box-Pierce y Ljung-Box	83
D. Criterio de información de Akaike	85
E. Criterio de información Bayesiano	87
F. FAC y FACP para diferentes modelos	89
Bibliografía	91

Capítulo 1

Introducción

1.1. Marco Contextual

La Bolsa de Valores es una organización privada que brinda las facilidades necesarias para que sus miembros, atendiendo los mandatos de sus clientes, introduzcan órdenes y realicen negociaciones de compra y venta de valores, tales como acciones de sociedades o compañías anónimas, bonos públicos y privados, certificados, títulos de participación y una amplia variedad de instrumentos de inversión.

Hoy en día, los sistemas de valores funcionan con métodos de pronóstico que permiten a las corporaciones y a los inversores tener un marco de referencia sobre cómo se comportará el mercado en el futuro y por ende tomar buenas decisiones de cartera. En ocasiones, estos sistemas funcionan a base de datos históricos que permiten construir mejores modelos matemáticos (Navidi, 2006).

Es importante mencionar la siguiente definición de las acciones, son las partes iguales en las que se divide el capital social de una empresa. Estas partes iguales del capital social de la empresa son adquiridas por una persona, que recibe el nombre de accionista, y representan la propiedad que la persona tiene sobre la empresa, es decir, el porcentaje de la empresa que le pertenece al accionista. En resumidas cuentas, una acción de una empresa es una parte de una empresa, para conseguir financiación y continuar con sus negocios. Los dueños de la empresa pueden reservarse una parte importante de la empresa para conservar el control de la propiedad si así lo desean (Rankia Chile, 2016).

La experiencia ha mostrado que la mayoría de los fenómenos reales estudiados en las finanzas, o en economía, poseen una naturaleza compleja, es decir, mientras más cercano a la realidad se desea realizar un modelo éste se volverá más complejo. Cuando uno intenta explicar, modelar o pronosticar alguna variable económico-financiera, comúnmente, se enfrentará a una serie de inconvenientes.

1.2. Antecedentes

La teoría y práctica del análisis y pronóstico de series de tiempo se han desarrollado rápidamente en los últimos años. Uno de los métodos de pronóstico a corto plazo desarrollado a finales de los sesenta del siglo XX y más conocido se denomina análisis ARIMA o análisis univariado de Box-Jenkins, se usa esta etiqueta porque George E. P. Box y Gwilym M. Jenkins son las dos personas responsables de formalizar el procedimiento (Alan Pankratz, 1983) y (Box, 2008).

En la disciplina económica, la línea de conocimiento Box–Jenkins, existe literatura que desarrolla modelos empíricos útiles para cuestiones financieras sin embargo dichos trabajos fueron adecuados a las circunstancias para las que se trabajó. A continuación se presentan algunos de los más relevantes.

1. Un caso de estudio para desarrollar un modelo acerca del precio de cierre semanal de la acción AT&T del año 1979, tomado de las acciones que se cotizan en la bolsa de Nueva York, usando la metodología Box–Jenkins. (Alan, 1983).
2. Parody (2016) desarrollo la predicción del precio de las acciones del sector bancario que cotizaron del índice general de la Bolsa de Valores de Colombia (IGBC), utilizando un modelo Log–normal complementado con simulaciones de Monte Carlo, a fin de determinar pruebas de bondad de ajuste del modelo mediante la raíz del error métrico cuadrado (RMSE). Los resultados encontrados indican que el modelo es eficiente para tener una aproximación a los posibles valores mínimos y máximos que pueden tomar las acciones, sin embargo los resultados carecen de la suficiente precisión para inducir compra certera de este tipo de activo financiero, razón por la cual se recomienda en próximas investigaciones la aplicación de modelos con promedios móviles de suavizamiento exponencial y modelos de la familia ARCH y GARCH que generan mayor capacidad de predicción.
3. Servín (2011) realiza un análisis para estimar la volatilidad de las acciones en la Bolsa de Valores de México, sustentado en trabajos como el de Alizadeh (1999), los cuáles valoran la utilidad práctica de los estimadores de Cuasi–Máxima Verosimilitud (QMLE); Magdoni (2003), presenta estimaciones del rango basados en los precios máximos y mínimos de la acción de IBM y haciendo las aplicaciones que desarrollaron Chou y Wang (2005), quienes describen los rasgos característicos del Modelo Autorregresivo de Rango Condicional (CARR); encontrando que los mejores resultados se obtienen mediante la utilización de los “modelos ARMA y GARCHX, con independencia de los rangos de valores extremos que se obtengan a partir

de datos diarios, semanales o mensuales ”.

1.3. Planteamiento del problema

Los principales mercados financieros a nivel mundial, requieren del conocimiento de modelos cuyo grado de sofisticación a nivel matemático, estadístico y econométrico mejoren con el paso del tiempo. No existe un modelo en el contexto de la metodología Box–Jenkins que obtenga cuál es el comportamiento del precio de cierre mensual de Grupo Financiero Inbursa en el periodo 2017 – 2018.

1.4. Justificación

Se espera que los resultados obtenidos sean utilizados por bancos mexicanos para innovar sus decisiones financieras con el fin de incrementar utilidades y así, otorgar más empleos al crecer como institución.

1.5. Objetivo General

1. Obtener un modelo estadístico que permita ajustarse adecuadamente a los datos para realizar pronósticos.

1.6. Breve descripción del contenido

En este trabajo se analizan los datos correspondientes a las series financieras de las acciones del Grupo Financiero Inbursa, S.A.B. de C.V. (GFINBURO.MX), con el propósito de proponer modelos, usando la metodología de Box–Jenkins, que permita ajustar adecuadamente a dichos datos para desarrollar políticas que innoven sus decisiones financieras.

En el capítulo dos, se hace una presentación básica de la teoría de series de tiempo y sus componentes para amenizar los conceptos manejados más adelante, después tratamos temas acerca de la autocorrelación en series de tiempo y por último los operadores y polinomios de retraso.

En el capítulo tres, vemos conceptos básicos para profundizar más adelante la metodología Box–Jenkins, así como el respectivo desarrollo de sus etapas.

En el capítulo cuatro se afirma cuál es el mejor modelo de acuerdo a el análisis de residuales, la prueba de Akaike, Ljung–Box (las últimas dos anexadas en los apéndices). Finalmente, se realiza el análisis y conclusiones de los resultados

obtenidos.

Capítulo 2

Conceptos Básicos de Series de Tiempo

El objetivo del análisis de una serie de tiempo consiste en encontrar un modelo adecuado para observaciones históricas en el tiempo, las cuales provienen de la realización de un proceso estocástico. Una vez que se ha encontrado el modelo adecuado se pueden tener predicciones de valores futuros. Para abordar este tipo de definiciones, en este caso debemos introducir los conceptos para familiarizarnos con el tema.

2.1. Elementos de Series de Tiempo

Una serie de tiempo se puede escribir como un caso particular de un proceso estocástico, es por eso que es importante definirlo. Los procesos estocásticos se encargan de estudiar fenómenos aleatorios que evolucionan en el tiempo, es decir que van modificándose y cambiando conforme pasa el tiempo (Guerrero, 2009).

Proceso estocástico: Es una colección o familia de variables aleatorias $\{X_t, t \in T\}$ las cuales están asociadas a un conjunto de índices T , de tal forma que a cada elemento del conjunto le corresponda una y sólo una variable aleatoria. Al conjunto de todos los posibles valores que la variable pueda tomar se le llama espacio de estados y se denota por S .

Si T es un intervalo de número reales ya sea cerrado o abierto, se dirá que el proceso estocástico es continuo, por otro lado, si el conjunto de índices es un conjunto finito o infinito numerable, el proceso será llamado proceso estocástico en tiempo discreto. Dada la definición de un proceso estocástico se procede a definir una serie de tiempo.

Serie de tiempo: Una serie de tiempo es una sucesión de observaciones generadas por un proceso estocástico $\{X_t, t \in T\}$ donde el conjunto de índices T

está en función del tiempo.

- **Series de tiempo discretas:** Una serie de tiempo se considera discreta cuando las observaciones se toman en un conjunto de tiempos discretos T_0 , generalmente, las observaciones se realizan en intervalos de tiempo fijos (x_1, x_2, \dots, x_n) .
- **Series de tiempo continuas:** Son aquellas que se obtienen cuando las observaciones se registran de manera continua en algún intervalo de tiempo (x_1, x_2, \dots, x_n) .

Modelo de serie de tiempo: Es la especificación de la distribución de probabilidad conjunta de una sucesión de variables aleatorias $\{X_t\}$, para las cuales cada conjunto de datos observados $\{x_t\}$ se toma una realización del proceso.

- **Modelos deterministas:** Se trata de métodos de extrapolación sencillos en los que no se hace referencia a las fuentes o naturaleza de la aleatoriedad subyacente en la serie.
- **Modelos estocásticos:** Se basan en la descripción simplificada del proceso aleatorio subyacente en la serie. Es decir, se asume que la serie observada (y_1, y_2, \dots, y_n) se extrae de una sucesión de variables aleatorias con cierta distribución conjunta no sencilla de determinar. Por lo que se construyen modelos aproximados que sean útiles para la generación de pronósticos. Por ejemplo los modelos Box-Jenkins.

2.1.1. Componentes de una Serie de Tiempo

A continuación se mencionan las componentes de una serie de tiempo.

- **Tendencia:** La tendencia de una serie de tiempo es el componente de largo plazo, representa el crecimiento o disminución en la serie sobre un periodo amplio.
- **Estacionalidad:** El componente estacional es un patrón de cambio que se repite a sí mismo año tras año.
- **Cíclos:** Se refiere a movimientos hacia arriba y hacia abajo alrededor del nivel de la tendencia. Generalmente demoran periodos largos (más de un año).
- **Fluctuaciones irregulares:** Son movimientos erráticos en una serie de tiempo que no siguen un patrón regular, ni reconocible. Tales movimientos representan “ lo que queda ” en una serie de tiempo después de que la tendencia, ciclos y variaciones estacionales han sido explicadas (Brockwell, 2002) y (Navidi, 2006)).

2.1.2. Modelos paramétricos para datos con tendencia

Considere el siguiente modelo para una serie de tiempo

$$y_t = T_t + e_t. \quad (2.1.1)$$

Donde

y_t es el valor de la serie de tiempo en el periodo t .

T_t es la tendencia de la serie de tiempo en el periodo t .

e_t es el error de la serie de tiempo en el periodo t .

Este modelo señala que la serie de tiempo y_t se puede representar mediante un nivel promedio que cambia con el tiempo, el término de error e_t representa las fluctuaciones aleatorias que ocasionan que los valores de la serie de tiempo y_t se desvíen del nivel promedio T_t .

Algunos modelos útiles para la tendencia son

1. **Sin tendencia:** Se modela como $T_t = \beta_0$, donde β_0 es una constante.
2. **Tendencia lineal:** Se modela como $T_t = \beta_0 + \beta_1 t$, significa que hay un crecimiento lineal durante un largo periodo ($\beta_1 > 0$ en el tiempo).
3. **Tendencia Cuadrática:** Se modela como $T_t = \beta_0 + \beta_1 t + \beta_2 t^2$, quiere decir que hay un cambio cuadrático o curvilíneo durante un largo periodo en el tiempo.
4. **Polinomial de p -ésimo orden:** Se modela como $T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_p t^p$, se utiliza para describir fenómenos no lineales tales como la progresión de epidemias de enfermedades.
5. **Exponencial:** Se modela como $T_t = \exp(\beta_0 + \beta_1 t)$.
6. **Logístico:** Se modela como $T_t = \frac{\beta_2}{1 + \beta_1 \exp(-\beta_0 t)}$.

Para la estimación de los parámetros $\beta = (\beta_0, \beta_1, \beta_2)$ en los modelos lineales y cuadráticos se utiliza el método de mínimos cuadrados ordinarios. En este método el vector de parámetros estimados $\hat{\beta}$ es el vector que produce el valor mínimo de la suma de los errores cuadrados. Se supone que el término de error e_t cumple con las suposiciones de normalidad, independencia y varianza constante. La suposición de varianza constante se puede verificar mediante una gráfica de residuos contra el tiempo. La suposición de normalidad se puede corroborar mediante la construcción de histogramas o gráficas normales de los residuos.

Al analizar datos de series de tiempo a menudo se transgrede la suposición de independencia, es muy común en los términos de error ordenados en el tiempo que éstos sean autocorrelacionados, es decir, la suposición de independencia o de errores no correlacionados generalmente no se satisface en series de tiempo. Por lo general los errores en los datos de series de tiempo tienen correlación

seriada, es decir, $E(\epsilon_i, \epsilon_{i+j}) \neq 0$. Se dice que estos términos de error están autocorrelacionados (Brockwell, 2002).

2.1.3. Autocorrelación en series de tiempo

La autocorrelación positiva existe cuando los términos de error positivos tienden a ser seguidos en el tiempo por términos de error positivos. Los términos de error negativos tienden a ser seguidos en el tiempo por términos de error negativos. Además existe autocorrelación negativa cuando los términos de error positivos tienden a ser seguidos en el tiempo por términos de error negativos y términos de error negativos tienden a ser seguidos en el tiempo por términos de error positivos.

La suposición de independencia plantea que los términos de error ordenados en el tiempo no manifiestan autocorrelación positiva o negativa, esto quiere decir que los términos de error ocurren en un patrón aleatorio en el tiempo.

Existe un tipo de autocorrelación positiva o negativa que se llama autocorrelación de primer orden, la cual plantea que el término de error en el periodo de tiempo t , se relaciona con el término de error en el periodo $t - 1$ mediante la ecuación

$$e_t = \phi_1 e_{t-1} + a_t. \quad (2.1.2)$$

Donde ϕ_1 es el coeficiente de correlación entre términos de error separados por un periodo y a_1, a_2, \dots son valores seleccionados al azar y en forma independiente de entre una distribución normal que tiene media cero y una varianza constante independiente del tiempo.

La prueba de Durbin-Watson (D-W), es una prueba formal para la autocorrelación de primer orden, positiva o negativa. La estadística para la prueba de Durbin-Watson, denominada estadística de D-W es

$$d = \frac{\sum_{t=2}^N (e_t - e_{t-1})^2}{\sum_{t=1}^N e_t^2} \quad (2.1.3)$$

donde e_1, e_2, \dots, e_N son los residuos ordenados en el tiempo (Brockwell, 2002).

Las pruebas de Durbin-Watson para autocorrelación de primer orden

Durbin y Watson demostraron que hay puntos ($d_{L,\alpha}$ y $d_{U,\alpha}$) tales que si α es la probabilidad de un error tipo I, entonces se cumplen las siguientes condiciones (Cuadro 2.1).

Para que se puedan llevar a cabo estas pruebas se han elaborado tablas que contienen los puntos $d_{L,\alpha}$ y $d_{U,\alpha}$ apropiados para varios valores de α , k (número de

H_0	H_a	Condición	
Los términos de error no están autocorrelacionados positivamente	Los términos de error están autocorrelacionados positivamente	$d < d_{L,\alpha}$	Se rechaza H_0 .
		$d > d_{U,\alpha}$	No se rechaza H_0 .
		$d_{L,\alpha} \leq d \leq d_{U,\alpha}$	La prueba no es concluyente
Los términos de error no están autocorrelacionados negativamente	Los términos de error están autocorrelacionados negativamente	$(4-d) < d_{L,\alpha}$	Se rechaza H_0 .
		$(4-d) > d_{U,\alpha}$	No se rechaza H_0 .
		$d_{L,\alpha} \leq (4-d) \leq d_{U,\alpha}$	La prueba no es concluyente
Los términos de error no están autocorrelacionados positiva ó negativamente	Los términos de error están autocorrelacionados positiva ó negativamente	$d < d_{L,\frac{\alpha}{2}}$	Se rechaza H_0 .
		$d > d_{U,\frac{\alpha}{2}}$	No se rechaza H_0 .
		$d_{L,\frac{\alpha}{2}} \leq d \leq d_{U,\frac{\alpha}{2}}$ ó	La prueba no es concluyente
		$d_{L,\frac{\alpha}{2}} \leq (4-d) \leq d_{U,\frac{\alpha}{2}}$	

Cuadro 2.1: Pruebas de Durbin–Watson para autocorrelaciones de primer orden

variables independientes) y n (número de observaciones) (Durbin et.al., 1950).

Observaciones

1. La validez de la prueba de Durbin-Watson depende del supuesto de que la población de todos los posibles residuos en cualquier tiempo t tiene una distribución normal.
2. La autocorrelación positiva se encuentra con más frecuencia en la práctica que la autocorrelación negativa.
3. No hay dependencia histórica entre las variables y_t (Bowerman et.al., 2007).

2.1.4. Series Estacionarias

Serie estacionaria.- Una serie de tiempo $\{X_t; t \in T\}$, es estacionaria en sentido estricto si y sólo si

$$F[x_{t_1}, x_{t_2}, \dots, x_{t_n}] = F[x_{t_1+k}, x_{t_2+k}, \dots, x_{t_n+k}] \quad \forall t_1, t_2, \dots, t_n \in T, k \in \mathbb{N}$$

es decir, si la función de distribución en cualquier conjunto finito de n variables aleatorias de la serie no se altera si se desplaza k periodos en el tiempo.

Serie estacionaria de segundo orden o débilmente estacionaria.- Una serie de tiempo $\{X_t; t \in T\}$, es estacionaria de orden 2 o débilmente si y sólo si

1. Es estacionaria en media, es decir, todas las variables aleatorias de la serie tienen la misma media y ésta es finita:

$$E(X_t) = \mu < \infty, \forall t \in T. \quad (2.1.4)$$

2. Todas las variables tienen la misma varianza y es finita, es decir, la dispersión en torno a la media constante a lo largo del tiempo es la misma para todas las variables de la serie de tiempo:

$$V(X_t) = E[X_t - \mu]^2 = \sigma_x^2 < \infty, \forall t \in T. \quad (2.1.5)$$

3. Las autocovarianzas sólo dependen del número de periodos de separación entre las variables y no del tiempo, es decir, la covarianza lineal entre dos variables aleatorias de la serie de tiempo que disten k periodos de tiempo es la misma que existe entre cualesquiera otras dos variables que estén también separadas k periodos, independientemente del momento concreto de tiempo al que están referidas

$$\begin{aligned} cov(X_{t+k}, X_t) &= E[X_{t+k} - \mu][X_t - \mu] \\ &= \gamma_{|(t+k)-t|} \\ &= \gamma_k < \infty, \forall k \in Z^+ \end{aligned} \quad (2.1.6)$$

(Brockwell, 2002).

De forma similar, la autocorrelación con lapso k es

$$\rho_x(k) = \frac{E(x_t - \mu)(x_{t+k} - \mu)}{\sqrt{E(x_t - \mu)^2 E(x_{t+k} - \mu)^2}} = \frac{E(x_t - \mu)(x_{t+k} - \mu)}{\sigma_x^2} = \frac{\gamma_x(k)}{\gamma_x(0)}. \quad (2.1.7)$$

Por lo tanto, se dice que un proceso estocástico estacionario en covarianza si y sólo si

$$E[X_t] = \mu < \infty \quad (2.1.8)$$

$$Cov[X_{t+k}, X_t] = \begin{cases} Var[X_t], & \sigma_X^2 < \infty \quad si \ k < 0 \\ \gamma_k < \infty, & si \ k \geq 1. \end{cases} \quad (2.1.9)$$

En resumen si una serie de tiempo es estacionaria, su media, su varianza y su autocovarianza (en diferentes rezagos) permanecen iguales sin importar el momento en el cual se midan; es decir, son invariantes respecto al tiempo.

Proceso estocástico de segundo orden o débilmente estacionario.- Un proceso estocástico se dice que es estacionario 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.

Sea X_t una serie de tiempo entonces con estas propiedades

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

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

Covarianza: $\gamma_k = E[(X_t - \mu)(X_{t+k} - \mu)]$,

donde γ_k , la covarianza (o autocovarianza) al rezago k , es la covarianza entre los valores que están separados k periodos (Villavicencio, 2018).

Proceso Estacionario de orden r .- Es un proceso cuando todos sus momentos hasta el de orden r existen y son estables.

2.2. Operadores y Polinomio de Retraso

Para eliminar la tendencia de las series, existen métodos que nos ayudan a lograr la estacionariedad. Estos métodos tienen como herramienta principal a los operadores de diferencia y de retraso.

1. Operador de retraso.- Es aquél denotado por B y que aplicado a una función dependiente del tiempo, proporciona esa misma función pero retardada un periodo

$$BX_t = X_{t-1}, \forall t \quad (2.2.1)$$

de forma general se expresa como

$$B^k X_t = X_{t-k}, \forall t. \quad (2.2.2)$$

2. Operador de adelanto.- Es la operación inversa al operador anterior, se define como $F = B^{-1}$, o bien

$$FX_t = X_{t+1}, \forall t \quad (2.2.3)$$

de forma general se expresa como

$$F^k X_t = X_{t+k}, \forall t. \quad (2.2.4)$$

3. Operador diferencia.-Es otro operador importante denotado con ∇ , definido por

$$\nabla X_t = X_t - X_{t-1}, \forall t. \quad (2.2.5)$$

Que puede ser escrito en términos de B como

$$\nabla X_t = X_t - X_{t-1} = (1 - B)X_t. \quad (2.2.6)$$

Ahora bien, en el análisis de series de tiempo se utilizan operadores de retraso en forma de polinomios, es decir el polinomio

$$X_t - g_1 X_{t-1} - g_2 X_{t-2} - \dots - g_k X_{t-k} = X_t - \sum_{j=1}^k g_j X_{t-j} = - \sum_{j=0}^k g_j X_{t-j} \quad (2.2.7)$$

es un polinomio de retraso que puede expresarse como $G(B)X_t$, en donde

$$G(B) = 1 - g_1 B - g_2 B^2 - \dots - g_k B^k = 1 - \sum_{j=1}^k g_j B^j = - \sum_{j=0}^k g_j B^j \quad (2.2.8)$$

con $g_0 = -1$ y los coeficientes g_1, g_2, \dots, g_k son constantes que ponderan la importancia de los retrasos con los cuales están asociados, además k puede ser $0, 1, 2, \dots$

El uso de polinomios de retraso es de particular importancia en el estudio de series de tiempo porque permiten expresar, de una manera concisa y simple, algunos de los modelos que han probado ser de mayor utilidad en la práctica para representar fenómenos reales. A reserva de presentar con cierto detenimiento las características más sobresalientes de dichos modelos, cabe mencionar que dentro de estos modelos se encuentran los de promedios móviles, que se representan mediante la expresión

$$X_t - \mu = \theta(B)a_t \quad (2.2.9)$$

en donde, $\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$, μ denota a la media de la serie (o, dicho de otra manera, el nivel del proceso).

$X_t - \mu$ representa a la desviación de X_t respecto a la media.

$\{a_t\}$ es una sucesión de variables aleatorias con ciertas características que facilitan su manejo y

$\theta_1, \dots, \theta_q$ son parámetros que sirven para relacionar a las sucesiones $\{a_t\}$ y $\{X_t\}$. Asimismo, los modelos *autorregresivos*, que se definen como

$$\phi(B)(X_t - \mu) = a_t \quad (2.2.10)$$

en donde, $\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$ y $\phi_1, \phi_2, \dots, \phi_p$ son parámetros.

A las combinaciones de los modelos anteriores, se les conoce con el nombre de modelos autorregresivos de promedios móviles también se les representa con la expresión

$$\phi(B)(X_t - \mu) = \theta(B)a_t. \quad (2.2.11)$$

Por último, con el uso de polinomios de retraso y del operador diferencia se obtienen las representaciones

$$\phi(B)\nabla^d X_t = \theta(B)a_t \quad (2.2.12)$$

que constituyen los modelos autorregresivos integrados y de promedios móviles.

Proceso Estacionario.- Es un proceso estocástico cuya distribución de probabilidad en un instante de tiempo fijo o una posición fija, es la misma para todos los instantes de tiempo o posiciones. En consecuencia, parámetros como la media μ y la varianza σ^2 , si existen, no varían a lo largo del tiempo o la posición.

Capítulo 3

Metodología Box-Jenkins

Los modelos clásicos de Box-Jenkins describen series temporales estacionarias. Si no es así, entonces tenemos que transformar la serie temporal (aparte de contener una tendencia y/o ciclos de larga duración, muestre fluctuaciones que se repiten anualmente, quizá con cambios graduales a través de los años) en una serie de valores, para tener una serie temporal estacionaria, en forma intuitiva, una serie temporal es estacionaria si, las propiedades estadísticas de la serie temporal son esencialmente constantes a través del tiempo. Si graficamos n valores y_1, y_2, \dots, y_n de una serie temporal, los cuales fluctúan con variación constante respecto a la media constante μ , entonces es razonable pensar que la serie temporal es estacionaria (Guerrero, 2009). En una forma más formal, primero definiremos.

3.1. Funciones de autocovarianza y autocorrelación

El comportamiento de una variable aleatoria X puede caracterizarse a través de su función de densidad $f(x)$. Similarmente, dos variables aleatorias están descritas por su función de densidad conjunta. En general en las series de tiempo existe cierta correlación entre las observaciones. La covarianza entre X_t y otra observación como X_{t+k} , es llamada autocovarianza de intervalo k , la cual nos indica el grado de variación de dos observaciones de la serie de tiempo, aportando la tendencia entre las variables, es decir, si una variable tiene valores grandes y la otra también, entonces, la autocovarianza será positiva mientras que si pasa lo contrario será negativa. Se tiene que si el proceso es estacionario, la

autocovarianza entre dos observaciones, X_t y X_{t+k} es

$$\begin{aligned} Cov(X_t, X_{t+k}) &= Cov(X_{t-t}, X_{t+k-t}) \\ &= Cov(X_0, X_k) \\ &= Cov(X_{t-t}, X_{|k|}) \\ &= \gamma(k) \end{aligned}$$

Es así como se puede decir que la autocovarianza sólo dependerá de su periodo k . En resumen, la autocovarianza entre X_t y X_{t+k} con periodo k está definida por

$$\gamma(k) = Cov(X_t, X_{t+k}) = E(X_t - \mu)(X_{t+k} - \mu). \quad (3.1.1)$$

La colección de los valores de γ_k , $k = 1, 2, \dots$ es llamada la función de autocovarianza. De esta forma obtenemos el coeficiente de autocorrelación del periodo o lapso k

$$\rho_k = \frac{E(X_t - \mu)(X_{t+k} - \mu)}{\sqrt{E(X_t - \mu)^2(X_{t+k} - \mu)^2}} = \frac{Cov(X_t, X_{t+k})}{Var(X_t)} = \frac{\gamma_k}{\gamma_0} \quad (3.1.2)$$

Notar que $Var(X_t) = Var(X_{t+k}) = \gamma_0$.

La colección de los valores ρ_k , $k = 1, 2, \dots$ es llamado la función de autocorrelación (FAC). Con lo anterior, las funciones de autocovarianza γ_k y autocorrelación tienen las siguientes propiedades (Robert et.al, 2010) y (William, 2006).

1. $\gamma_0 = Var(X_t)$; $\rho_0 = 1$.
2. $|\gamma_k| \leq \gamma_0$; $\rho_k \geq 0$.
3. $\gamma_k = \gamma_k$ y $\rho_k = \rho_{-k}$ para toda k , lo que quiere decir que son funciones simétricas a partir del lapso $k=0$. La diferencia entre X_t y X_{t-k} es la misma.
4. Las funciones γ_k y la función de autocorrelación ρ_k son semidefinidas positivas en el sentido que

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \gamma_{|t_i - t_j|} \geq 0 \quad y \quad \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \rho_{|t_i - t_j|} \geq 0 \quad (3.1.3)$$

para cualquier conjunto de puntos t_1, t_2, \dots, t_n y cualquier número $\alpha_1, \alpha_2, \dots, \alpha_n$.

Para facilitar la identificación de modelos a partir de valores muestrales, en la Figura 3.1 se muestran los comportamientos típicos de la función de autocorrelación.

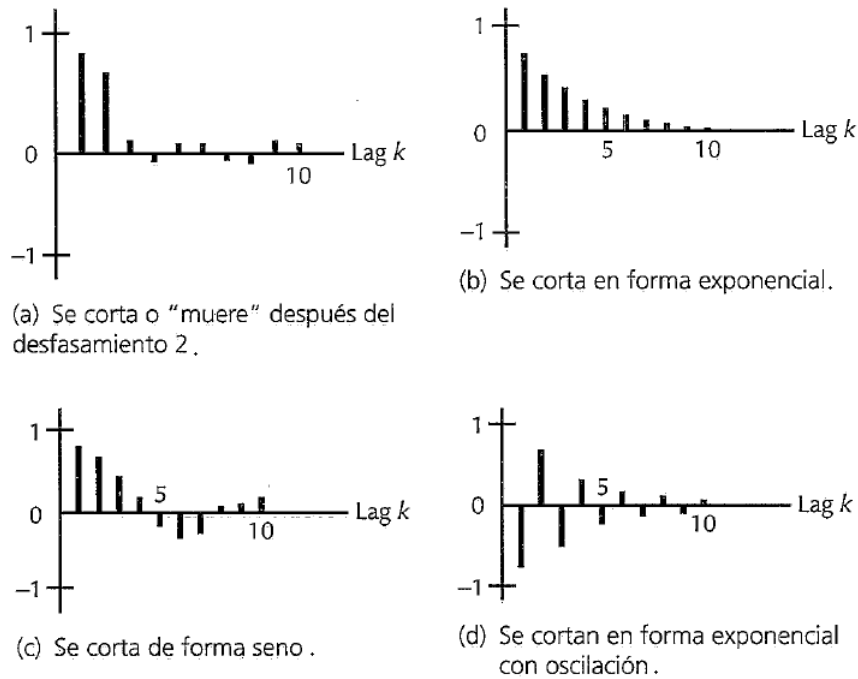


Figura 3.1: Comportamiento de las Funciones de Autocorrelación (Guerrero, 2009).

Función de Autocorrelación Parcial

La Función de Autocorrelación Parcial (FACP) de un proceso X_t se define como

$$\phi_{1,1} = Corr(X_{t+1}, X_t) = \rho_1 \tag{3.1.4}$$

$$\phi_{k,k} = Corr(X_t, X_{t+k} | X_{t+1}, X_{t+2}, \dots, X_{t+k-1}); k = 0, 1, 2, 3... \tag{3.1.5}$$

Es la correlación entre X_t y X_{t+k} después de su dependencia lineal mutua en las que intervienen las variables $X_{t+1}, X_{t+2}, \dots, X_{t+k-1}$ que se han eliminado.

La correlación parcial entre dos variables aleatorias es la correlación que queda si el impacto posible de todas las otras variables aleatorias ha sido eliminado. Para esto haremos uso de la correlación condicional.

$$Corr(X_t, X_{t+k} | X_{t+1}, \dots, X_{t+k-1}) \tag{3.1.6}$$

que es mejor conocida como la autocorrelación parcial en el análisis de series de tiempo.

La autocorrelación parcial puede ser obtenida considerando un modelo de regresión, donde la variable dependiente X_{t+k} de un proceso estacionario con media

cero es retrasado en k variables $X_{t+k-1}, X_{t+k-2}, \dots, X_t$, es decir, La FACP se puede obtener mediante un modelo autorregresivo de orden k de la siguiente forma:

$$X_{t+k} = \phi_z(k_1)X_{t+k-1} + \phi_z(k_2)X_{t+k-2} + \dots + \phi_z(k_k)X_t + e_{t+k} \quad (3.1.7)$$

donde, $\phi_z(k_i)$: el i -ésimo parámetro de regresión .
 e_{t+k} : es el término de error con media cero que no está correlacionado con X_{t+k-j} para $j=1,2,\dots,k$. Multiplicando ambos lados de la ecuación anterior por X_{t+k-j} , obtenemos

$$\gamma_z(j) = \phi_z(k_1)\gamma_z(j-1) + \phi_z(k_2)\gamma_z(j-2) + \dots + \phi_z(k_k)\gamma_z(j-k) \quad (3.1.8)$$

y de ahí

$$\rho_z(j) = \phi_z(k_1)\rho_z(j-1) + \phi_z(k_2)\rho_z(j-2) + \dots + \phi_z(k_k)\rho_z(j-k). \quad (3.1.9)$$

Para $j=1,2,\dots,k$ se tiene el sistema de ecuaciones

$$\rho_z(1) = \phi_z(k_1)\rho_z(0) + \phi_z(k_2)\rho_z(1) + \dots + \phi_z(k_k)\rho_z(k-1) \quad (3.1.10)$$

$$\rho_z(2) = \phi_z(k_1)\rho_z(1) + \phi_z(k_2)\rho_z(2) + \dots + \phi_z(k_k)\rho_z(k-2) \quad (3.1.11)$$

$$\dots \quad (3.1.12)$$

$$\rho_z(k) = \phi_z(k_1)\rho_z(k-1) + \phi_z(k_2)\rho_z(k-2) + \dots + \phi_z(k_k)\rho_z(o). \quad (3.1.13)$$

Usando la regla de Cramer sucesivamente para $k=1,2,\dots$ tenemos

$$\phi_z(11) = \rho_z(1) \quad (3.1.14)$$

$$\phi_z(22) = \frac{\begin{vmatrix} 1 & \rho_x(1) \\ \rho_x(1) & \rho_x(2) \end{vmatrix}}{\begin{vmatrix} 1 & \rho_x(1) \\ \rho_x(1) & 1 \end{vmatrix}} \quad (3.1.15)$$

•
•
•

$$\phi_z(kk) = \frac{\begin{vmatrix} 1 & \rho_x(1) & \rho_x(2) & \dots & \rho_x(k-2) & \rho_x(1) \\ \rho_x(1) & 1 & \rho_x(1) & \dots & \rho_x(k-3) & \rho_x(2) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_x(k-1) & \rho_x(k-2) & \rho_x(k-3) & \dots & \rho_x(1) & \rho_x(k) \end{vmatrix}}{\begin{vmatrix} 1 & \rho_x(1) & \rho_x(2) & \dots & \rho_x(k-2) & \rho_x(k-1) \\ \rho_x(1) & 1 & \rho_x(1) & \dots & \rho_x(k-3) & \rho_x(k-2) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_x(k-1) & \rho_x(k-2) & \rho_x(k-3) & \dots & \rho_x(1) & 1 \end{vmatrix}}. \quad (3.1.16)$$

El gráfico de la Función de Autocorrelación y de la Función de Autocorrelación Parcial brindan información importante sobre el modelo que se ajusta al comportamiento de la serie de tiempo. Puede presentar comportamiento de extinción exponencial, extinción lenta, movimientos sinoidales o truncamiento.

3.1.1. Matriz de covarianza de un proceso estacionario

Una matriz de covarianza Γ_n es simétrica y con elementos constantes en la diagonal principal, llamada matriz de autocovarianza definida de la forma siguiente

$$\Gamma_n = \begin{bmatrix} \gamma_z(0) & \gamma_z(1) & \gamma_z(2) & \dots & \gamma_z(n-1) \\ \gamma_z(1) & \gamma_z(0) & \gamma_z(1) & \dots & \gamma_z(n-2) \\ \gamma_z(2) & \gamma_z(1) & \gamma_z(0) & \dots & \gamma_z(n-3) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \gamma_z(n-1) & \gamma_z(n-2) & \gamma_z(n-3) & \dots & \gamma_z(0) \end{bmatrix}. \quad (3.1.17)$$

$$= \sigma_z^2 \begin{bmatrix} 1 & \rho_z(1) & \rho_z(2) & \dots & \rho_z(n-1) \\ \rho_z(1) & 1 & \rho_z(1) & \dots & \rho_z(n-2) \\ \rho_z(2) & \rho_z(1) & 1 & \dots & \rho_z(n-3) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_z(n-1) & \rho_z(n-2) & \rho_z(n-3) & \dots & \rho_z(0) \end{bmatrix}$$

$$= \sigma_z^2 \mathbb{P}_n.$$

La matriz de correlación correspondiente \mathbb{P}_n es llamada **matriz de autocorrelación**.

Las matrices de autocovarianza y autocorrelación son definidas positivas para cualquier proceso estacionario.

Existe una condición que la autocorrelación de un proceso estacionario satisfice, pues la definición positiva de la matriz de autocorrelación implica que su determinante es mayor que cero. Por ejemplo, para $n = 2$, tenemos

$$\begin{vmatrix} 1 & \rho_z(1) \\ \rho_z(1) & 1 \end{vmatrix} > 0 \Rightarrow 1 - \rho_z^2(1) > 0 \Rightarrow -1 < \rho_z(1) < 1.$$

3.1.2. Funciones Muestrales

Cuando tratamos con fenómenos reales de cualquier índole, se tienen datos observados pero se desconocen los parámetros de las variables aleatorias del proceso. Para elegir un modelo adecuado de los datos, usaremos la función de autocorrelación muestral (FACM). Si los datos provienen de una serie de tiempo estacionaria, entonces la FACM nos dará un estimador de la función de autocorrelación de $\{X_t\}$.

Sea $\{z_1, z_2, \dots, z_n\}$ una serie de tiempo, entonces la media muestral se denota como

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i. \quad (3.1.18)$$

La varianza muestral es

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2. \quad (3.1.19)$$

La autocovarianza muestral para $-n < k < n$ es

$$\hat{\gamma}^2(k) := \frac{1}{n} \sum_{i=1}^{n-|k|} (X_{i+|k|} - \bar{X})(X_i - \bar{X}). \quad (3.1.20)$$

Sin embargo, en la práctica sólo se puede calcular la función de autocorrelación muestral, $\hat{\rho}_k$, la cual está en función de $\hat{\gamma}_k$, la covarianza muestral en el lapso k y la varianza muestral $\hat{\gamma}_0$.

De este modo

$$\hat{\gamma}_k = \frac{\sum (X_t - \bar{X})(X_{t+k} - \bar{X})}{n - k} \quad (3.1.21)$$

y

$$\hat{\gamma}_0 = \frac{\sum (X_t - \bar{X})^2}{n - 1} \quad (3.1.22)$$

Y así, la autocorrelación muestral para $-n < k < n$ es

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad (3.1.23)$$

donde, n es el tamaño de la muestra y \bar{X} la media muestral.

La Función de Autocorrelación Muestral puede gráficamente darnos información (también llamado correlograma muestral) de la serie según sea el caso

1. Si se corta rápido o tiende a cero, entonces la serie de tiempo es estacionaria.
2. Si tiende a cero lentamente, entonces la serie no será considerada estacionaria.

Cuando se tiene una serie de tiempo finita de tamaño T , se pueden estimar a lo más $T - 1$ coeficientes de autocorrelación.

Ejemplo (Ruido blanco).- Si $\{X_t; t \in T\}$ es una sucesión de variables aleatorias no correlacionadas, cada una con media cero y varianza σ^2 , entonces se dice que la sucesión es un ruido blanco. Para indicar que la sucesión $\{X_t; t \in T\}$ es un ruido blanco se usará la siguiente notación:

$$\{X_t\} \sim WN(0, \sigma^2). \quad (3.1.24)$$

La función de autocovarianza de un proceso de ruido blanco es

$$\gamma_k = \begin{cases} \sigma^2, & \text{si } k = 0; \\ 0, & \text{si } k \neq 0; \end{cases} \quad (3.1.25)$$

Función de autocorrelación

$$\rho_k = \begin{cases} 1, & \text{si } k = 0; \\ 0, & \text{si } k \neq 0; \end{cases} \quad (3.1.26)$$

Y función de autocorrelación parcial

$$\phi_{k,k} = \begin{cases} 1, & \text{si } k = 0; \\ 0, & \text{si } k \neq 0; \end{cases} \quad (3.1.27)$$

no depende de t , entonces el ruido blanco es estacionario.

Cuando se hable de autocorrelación y autocorrelación parcial se hace referencia a ρ_k y $\phi_{k,k}$ para $k \neq 0$, en otro caso $\rho_0 = \phi_{0,0} = 1$.

Estas dos funciones son iguales a cero en un proceso de ruido blanco. Un proceso de ruido blanco será Gaussiano si su distribución conjunta es normal.

3.2. Modelos Lineales Estacionarios

3.2.1. Modelo lineal general

Cuando el objetivo es predecir el valor que tomará en el momento t una variable que presenta dependencia temporal, la metodología indica, a grandes rasgos, recolectar información pasada de la variable, observar y detectar patrones, para finalmente usarlos en el pronóstico de la misma. La estructura de la dependencia

temporal de un proceso estocástico se almacena en las funciones de autocovarianza y autocorrelación. Se utiliza la información que brindan ambas, para generar un patrón sistemático y posteriormente, un modelo que represente el comportamiento de la serie, mismo que servirá para generar pronósticos.

En un modelo de series de tiempo, la serie se descompone en dos términos; el de error o innovación y la parte sistemática. La innovación es la parte aleatoria que indica el término de error entre la serie observada y la modelada. La parte sistemática del modelo se estima usando el conjunto de información recolectada en el pasado.

El problema de la modelación de series de tiempo, reside en el ajuste de la parte sistemática de manera que la innovación tenga la distribución de un Ruido Blanco.

Consideraremos dentro de los procesos estocásticos estacionarios solo a los que se pueden representar como combinación lineal de variables aleatorias.

Bajo condiciones generales, para procesos estacionarios con distribución normal y media cero, el proceso se puede representar como una combinación lineal de los valores pasados más un ruido blanco, de la forma

$$\tilde{X}_t = \pi_1 \tilde{X}_{t-1} + \pi_2 \tilde{X}_{t-2} + \dots + a_t = \sum_{j=1}^{\infty} \pi_j \tilde{X}_{t-j} + a_t \quad \forall t, t = 1, 2, \dots \quad (3.2.1)$$

Cumpliendo con las condiciones

- * El proceso no deber ser anticipante, es decir, el futuro no puede determinar el presente.
- * El proceso debe ser invertible, es decir, el presente depende de forma convergente de su propio pasado. En otras palabras, los parámetros del modelo deben cumplir que

$$\sum_{i=1}^{\infty} \pi_i^2 < \infty. \quad (3.2.2)$$

Existen al menos, tres formas de representar un modelo lineal

1. Forma autorregresiva, en donde el valor presente de la variable se expresa en función de su propio pasado más una innovación simultánea a la variable.
2. Medias móviles, en donde el valor presente de la variable se expresa en función de todas las innovaciones presentes y pasadas.
3. Forma finita, en donde el valor de X_t depende del pasado de X hasta el momento $t - p$ (parte autorregresiva), de la innovación actual y de su

pasado hasta el momento $t - q$ (parte de medias móviles). Este modelo recibe el nombre de Autorregresivo de Medias Móviles de orden (p, q) ; $ARMA(p, q)$.

Cuando el modelo es conocido, se puede utilizar cualquier representación dependiendo del objetivo que se persiga. Si el modelo no es conocido y debe plantearse y estimarse, se utiliza la formulación finita.

Cabe recalcar que cuando se construye un modelo, se pretende replicar una realidad compleja, y el objetivo es lograr un modelo que se ajuste con suficiente precisión a las características que refleje la función de autocorrelación de la serie bajo estudio.

3.2.2. Procesos Autorregresivos (AR)

El procesos autorregresivos de orden p se denota como $AR(p)$. El proceso $AR(p)$ expresa a X_t en función de su pasado hasta el retardo $t - p$ más una innovación contemporánea, y se denota como

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + a_t \quad (3.2.3)$$

o de forma equivalente, como

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = a_t \Rightarrow \phi_p(B) X_t = a_t.$$

Donde $\phi_p(B)$ es el polinomio autorregresivo y $(\phi_1, \phi_2, \dots, \phi_p)$ es el vector de parámetros autorregresivos.

Es necesario comprobar si el proceso $AR(p)$ es estacionario, por lo que haremos uso del siguiente teorema

Teorema 1.1 *Un proceso autorregresivo finito $AR(p)$ es estacionario si las raíces del polinomio autorregresivo $\phi_p(B)$ se encuentran fuera del círculo unitario (Box, 2008).*

Todo modelo autorregresivo finito cumple con las condiciones del modelo lineal general; no anticipante e invertible, para cualquier valor de los parámetros. Los procesos $AR(p)$ son no anticipantes porque su formulación hace depender al valor de X_t de su pasado y no de su futuro, y es invertible porque su fórmula finita hace que se cumpla la condición

$$\sum_{i=1}^{\infty} \pi_i^2 < \infty. \quad (3.2.4)$$

Un proceso estacionario $AR(p)$ tiene como características

1. Media

$$E[X_t] = E[\phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + a_t] \quad (3.2.5)$$

$$= \phi_1 E[X_{t-1}] + \phi_2 E[X_{t-2}] + \dots + \phi_p E[X_{t-p}] + E[a_t].$$

Como el proceso es estacionario, la media es constante

$$(1 - \phi_1 - \phi_2 - \dots - \phi_p)E[X_t] = 0 \Rightarrow E[X_t] = 0. \quad (3.2.6)$$

Este modelo se puede generalizar para representar series con media distinta de cero. El modelo AR(p)

$$X_t = \delta + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + a_t \quad (3.2.7)$$

tiene como media

$$(1 - \phi_1 - \phi_2 - \dots - \phi_p)E[X_t] = \delta \Rightarrow E[X_t] = \frac{\delta}{1 - \phi_1 - \phi_2 - \dots - \phi_p}. \quad (3.2.8)$$

2. La función de autocovarianza de un proceso AR(p), queda expresada de la forma

$$\gamma_k = Cov(X_t, X_{t+k}) \quad (3.2.9)$$

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p}, \quad (3.2.10)$$

3. Ahora para obtener la función de autocorrelación ρ_k de un proceso AR(p) se divide entre γ_0 dado por $\gamma_0 = \sigma^2$ y así se tiene

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \quad k \geq 1 \quad (3.2.11)$$

tiene la misma estructura que la de un proceso AR(1) es decir $\rho_k = \phi_1^k$; decrece exponencialmente hacia cero sin truncarse.

4. La función de autocorrelación parcial para el modelo AR(p) es el siguiente

$$\gamma_k = \begin{cases} \phi_{k,k} = \phi_p, & \text{si } k = p; \\ \phi_{k,k} = 0, & \text{si } k > p. \end{cases} \quad (3.2.12)$$

Para $k < p$ los valores de $\phi_{k,k}$ pueden ser fácilmente calculados del último componente de

$$\phi_{k,k} = \Gamma_h^{-1} \gamma_h \quad (3.2.13)$$

donde $\Gamma = [\gamma_{i-j}]_{i,j=1}^k$ y $\gamma_k = [\gamma_1, \gamma_2, \dots, \gamma_k]$. Por lo tanto $\gamma_{k,k}$ desaparece después del lapso p . Es por eso, que sabemos que la **FACP** desaparece después del retraso p .

Notar que cuando $k > p$ la última columna del determinante en el numerador de (3.1.16) puede ser escrita como una combinación lineal de las columnas previas del determinante, de esta manera el determinante es igual a cero y la FACP se corta después del lapso p .

3.2.3. Proceso de Medias Móviles (MA)

Otro tipo de modelo de gran importancia práctica en la representación de series de tiempo es el proceso finito de medias móviles. Estos procesos representan series de tiempo de memoria corta, describen fenómenos que producen eventos con un efecto inmediato y con periodos cortos de tiempo. Un proceso de medias móviles tiene la forma

$$X_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (3.2.14)$$

donde $a_j \sim WN(0, \sigma^2)$, $j = t - q, \dots, t$.

Modelo MA(q)

Se puede expresar en términos del operador de retraso de la forma

$$X_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \Rightarrow X_t = \theta_q(B) a_t, \quad (3.2.15)$$

donde $\theta_q(B)$ es el polinomio de medias móviles dado por $(1 - B - B^2 - \dots - B^q)$ y $(\theta_1, \theta_2, \dots, \theta_q)$ es el vector de parámetros de medias móviles.

Al ser un modelo con más retardos, la memoria de éste aumenta y la estructura dinámica representada por el modelo es más abundante, pues la perturbación a_t en un modelo $\mathbf{MA}(q)$ permanece q periodos en el sistema. Esta memoria se reflejará en la estructura de las funciones de autocovarianza y autocorrelación. Como el modelo de medias móviles es un modelo lineal truncado en el retardo q , entonces éste será estacionario bajo las mismas condiciones, es decir, que la sucesión de los parámetros del modelo sea convergente

$$\sum_{i=1}^q \theta_i^2 < \infty. \quad (3.2.16)$$

Esta condición siempre se cumple ya que el número de parámetros de un modelo $\mathbf{MA}(q)$ siempre es finito.

Con respecto a la función de autocovarianza de un modelo $\mathbf{MA}(q)$, se tiene que la autocovarianza del lapso k está dada por

$$\gamma_k = Cov(X_t, X_{t+k}) \quad (3.2.17)$$

$$\gamma_k = \begin{cases} (1 + \theta_1^2 + \dots + \theta_q^2) \sigma^2, & \text{si } k = 0; \\ (-\theta_k - \theta_1 \theta_{k+1} - \dots - \theta_{q-k} \theta_q) \sigma^2, & \text{si } k = 1, \dots, q; \\ 0, & \text{si } k > q. \end{cases} \quad (3.2.18)$$

Por ende, la función de autocorrelación se denota tomando en cuenta que

$$\rho_q(k) = \begin{cases} 1, & \text{si } k = 0; \\ \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2}, & \text{si } k = 1, \dots, q; \\ 0, & \text{si } |k| > q. \end{cases} \quad (3.2.19)$$

La función de autocorrelación de un proceso $MA(q)$ se corta después del retraso q , lo que nos permite identificar si un serie de tiempo dada es generada por un proceso de medias móviles. Un modelo $MA(q)$ es no anticipante, porque el futuro no influye en el pasado y será invertible si su representación autorregresiva es tal que la influencia de X_{t-k} es menor conforme se aleja del pasado. El siguiente teorema proporciona las condiciones necesarias y suficientes para que el modelo de medias móviles sea invertible.

Teorema 1.2. *Un proceso de medias móviles finito es invertible si las raíces del polinomio de medias móviles $\theta_q(B)$ se encuentran fuera del círculo unitario (Box, 2008).*

3.2.4. Procesos autorregresivos de medias móviles o **AR-MA**(p,q)

Un proceso estacionario e invertible puede ser representado de la forma autorregresiva o de medias móviles. En la práctica, pueden sucitarse problemas en su representación al contener demasiados parámetros, aunque el modelo sea de orden finito, por lo que se procede a la unión de ambos modelos en uno solo que recibe el nombre de proceso autorregresivo de medias móviles **ARMA**, pues un modelo con demasiados parámetros reduce la eficiencia de la estimación. Un proceso ARMA determina a X_t en función de su pasado hasta el retardo p , de la innovación contemporánea y el pasado de la innovación hasta el retardo q

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad (3.2.20)$$

donde $\{X_t\}$ es estacionaria y $\{a_t\} \sim WN(0, \sigma^2)$, de forma equivalente se expresa como

$$(1 - \psi_1 B - \dots - \psi_p B^p) X_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t \quad (3.2.21)$$

o bien,

$$\Phi_p(B) X_t = \Theta_q(B) a_t \quad (3.2.22)$$

donde $\Phi_p(B)$ es el polinomio autorregresivo y $\Theta_q(B)$ es el polinomio de medias móviles.

Las condiciones de estacionariedad del modelo $ARMA(p,q)$ vienen impuestas por la parte autorregresiva, dado que la parte de medias móviles finita siempre es estacionaria, mientras que las condiciones de invertibilidad del modelo se comprueban a partir de la parte de medias móviles, pues la parte autorregresiva finita siempre lo será.

En otras palabras para que este proceso sea invertible, se requiere que las raíces de la ecuación $\theta_q(B) = 0$ y para que sea estacionario, se requiere que las raíces de la ecuación $\phi_p(B) = 0$ estén fuera del círculo unitario.

Se asumirá que ambas ecuaciones no comparten raíces comunes.

La función de autocovarianza de un proceso **ARMA**(p, q) es

$$\begin{aligned}\gamma_k &= E[x_t, x_{t-k}] \\ &= E[(\phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q})x_{t-k}] \\ &= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p}, \quad k \geq (q+1)\end{aligned}$$

y de esta podemos obtener la función de autocorrelación

$$\begin{aligned}\rho_k &= \frac{\phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p}}{\gamma_0} \\ &= \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \quad k \geq (q+1).\end{aligned}$$

Obsérvese que, la FAC simple del proceso ARMA es análoga a la del proceso AR(p). Entonces la FAC simple del proceso ARMA(p, q), se corta después del lapso q así como ocurre en un proceso AR(p).

Para la FACP, dado que el proceso ARMA(p, q) contiene al proceso MA(q) como un caso particular, la FACP tendrá una mezcla de decaimientos exponenciales u ondas sinusoidales amortiguadas dependiendo de las raíces de $\Phi(B)$ y $\Gamma(B)$.

Por último, en el Apéndice F, se da una descripción de la FAC y FACP para cada uno de los procesos vistos anteriormente.

3.3. Modelos no estacionarios

Nótese que hasta ahora se ha centrado la exposición en modelos que son estacionarios, es decir, se basan en el supuesto de estacionariedad en covarianza, en los que la media y la varianza son constantes finitas y las autocovarianzas no dependen del tiempo sino del número de periodos de separación entre las variables. En la práctica lo más común es que las series que se analizan sean no estacionarias, pues tienen media o varianza no constante.

Cuando una serie no es estacionaria en varianza, se utilizarán las transformaciones *Box-Cox*

$$X_t^{(\lambda)} = \begin{cases} \frac{X_t^\lambda - 1}{\lambda}, & \text{si } \lambda \neq 0; \\ \log(X_t), & \text{si } \lambda = 0 \end{cases} \quad (3.3.1)$$

donde λ es el parámetro de transformación. Usualmente, las transformaciones Box-Cox, además de estabilizar la varianza, mejoran la aproximación a la distribución normal del proceso $\{X_t\}$.

En las series económicas, la tendencia es una característica dominante, que puede ser creciente, decreciente, lineal o exponencial. Las series que presentan este

tipo de comportamientos, no son estacionarias. La no estacionariedad en media se puede modificar modelando tendencias, mediante modelos globales que especifiquen a la tendencia como una función del tiempo: $X_t = f(t) + \mu_t$, donde $f(t)$ es una función determinista del tiempo, ya sea lineal, cuadrática, exponencial, etc., y μ_t es un proceso estocástico estacionario con media cero. Estos modelos reciben el nombre de modelos de tendencia determinista ya que suponen que la serie evoluciona de forma perfectamente predecible.

Cuando no hay estacionariedad en media, es posible modelar a través de los modelos **ARMA**(p, q), pues éste no será estacionario si alguna raíz de su polinomio AR no está fuera del círculo unitario, por que hay dos situaciones.

Si alguna raíz está dentro del círculo unitario, la serie decrecerá o crecerá rápidamente al infinito. Esta evolución no es comúnmente observada en series económicas.

Si alguna raíz es igual a la unidad, la serie seguirá un modelo que describe realizaciones con comportamiento similar a lo largo del tiempo, pero que varían de nivel. Esta conducta es comúnmente observada en series económicas, por lo que procederemos a modelar series no estacionarias a través de modelos ARMA(p, q) no estacionarios.

La no estacionariedad también puede estar influenciada por algún factor de tipo semideterminista como puede ser la **estacionalidad**, si en cambio, el problema es que se aprecia una **tendencia** en el comportamiento de la serie, es bastante posible que dicha tendencia pueda eliminarse mediante aplicación del operador diferencia, lo cual da origen a los modelos ARIMA

3.3.1. Modelos ARIMA(p, d, q)

Los modelos ARIMA permiten describir un valor como una función lineal de datos anteriores y errores aleatorios, además de que incluyen un componente cíclico o estacional. Supongamos un modelo ARMA(p, q) de la forma

$$\Phi_p(B)X_t = \Theta_q(B)a_t \quad (3.3.2)$$

donde el polinomio AR se puede factorizar en función de sus p raíces B_1, B_2, \dots, B_p ,

$$\Phi_p(B) = (1 - B_1^{-1}B)(1 - B_2^{-1}B)\dots(1 - B_p^{-1}B). \quad (3.3.3)$$

Supongamos ahora que $(p-1)$ raíces son estacionarias (con módulo no unitario) y la restante es una raíz unitaria. Entonces, el polinomio AR se puede escribir de la forma

$$\Phi_p(B) = (1 - B_1^{-1}B)(1 - B_2^{-1}B)\dots(1 - B_p^{-1}B) \quad (3.3.4)$$

$$\begin{aligned} &= \phi_{p-1}(B)(1 - (1)^{-1}B) \\ \Rightarrow \Phi_p(B) &= \phi_{p-1}(B)(1 - B) \end{aligned}$$

donde el polinomio $\Phi_{p-1}(B)$ es el producto de los $(p-1)$ polinomios de orden uno asociados a las raíces fuera del círculo unidad. Sustituyendo este polinomio en el modelo ARMA(p, q) obtenemos

$$\Phi_{p-1}(B)(1-B)X_t = \Theta_q(B)a_t \Rightarrow \Phi_{p-1}(B)\nabla X_t = \nabla(B)a_t \quad (3.3.5)$$

donde el polinomio $\Phi_{p-1}(B)$ es estacionario porque todas sus raíces están fuera del círculo unitario y polinomio $\nabla = (1-B)$ contiene a la raíz unitaria.

El modelo $\Phi_{p-1}(B)\nabla X_t = \nabla_q(B)a_t$ representa el comportamiento de un proceso X_t que no es estacionario porque contiene una raíz unitaria y recibe el nombre de proceso integrado de orden 1.

En general, el polinomio AR del modelo puede contener más de una raíz unitaria, y se descompone de la forma

$$\Phi_p(B) = \phi_{p-d}(B)(1-B)^d \quad (3.3.6)$$

que sustituyendo en el modelo ARMA(p, q) nos da como resultado

$$\Phi_{p-d}(B)\nabla^d X_t = \Theta_q(B)a_t \quad (3.3.7)$$

donde el polinomio $\Phi_{p-d}(B)$ es estacionario y el polinomio de orden d , $\nabla^d = (1-L)^d$, contiene las d raíces unitarias no estacionarias.

Este proceso recibe el nombre de proceso integrado de orden d y se denota $X_t \sim I(d)$. Mencionaremos una definición de relevancia

Definición: Un proceso X_t es integrado de orden d , $X_t \sim I(d)$, si X_t no es estacionario, pero su diferencia de orden d , $\nabla^d X_t$, sigue un proceso ARMA(p, d, q) estacionario e invertible.

En este proceso, el orden de integración es el número de diferencias que hay que tomar para lograr la estacionariedad en media. En la práctica, d casi siempre toma los valores 0, 1 y más 2.

En general, si una serie $\{Z_t\}$ es integrada de orden d , se representa con el modelo

$$\Phi_p(B)\nabla^d X_t = \delta + \Theta_q(B)a_t \quad (3.3.8)$$

donde el polinomio autorregresivo estacionario $\Phi_p(B)$ y el polinomio invertible de medias móviles $\Theta_q(B)$ no tienen raíces comunes.

Este modelo recibe el nombre de Modelo Autorregresivo Integrado de Medias Móviles de orden (p, d, q) o de forma sistetizada ARIMA(p, d, q), donde p es el orden del polinomio autorregresivo estacionario, d es el orden de integración de la serie, es decir, el número de diferencias que hay que tomar a la serie para que sea estacionaria, y q es el orden del polinomio de medias móviles invertibles.

3.4. Modelación ARIMA

En esta sección nos referiremos en particular al análisis univariado ya que realizaremos una modelación ARIMA la cual, supone que la observación secuenciada por tiempo en una serie de tiempo podría ser estadísticamente dependiente. Luego, intentamos encontrar una buena manera de la naturaleza de esa relación estadística. Para la modelación ARIMA, haremos uso de la metodología Box-Jenkins. En este caso, se conocen los valores de la serie temporal $\{X_t\}$ y se busca determinar el modelo $ARIMA(p, d, q)$ que la pudo haber generado. Box y Jenkins proponen un procedimiento iterativo de cuatro pasos.

Paso 1. Identificación tentativa: Se utilizan datos antiguos para identificar en forma tentativa un modelo apropiado de Box-Jenkins.

Paso 2. Estimación: Se utilizan datos antiguos para estimar los parámetros del modelo identificado en forma tentativa.

Paso 3. Comprobación del diagnóstico: Se utilizan varios diagnósticos para comprobar si es adecuado el modelo identificado en forma tentativa, y, si es necesario, para recomendar un modelo mejorado, el cual considera entonces como un nuevo modelo identificado en forma tentativa.

Paso 4. Pronóstico: una vez que se obtuvo el modelo final, se usa para pronosticar valores futuros de series temporales.

3.4.1. Identificación

Se selecciona el modelo $ARIMA(p, d, q)$ que reproduzca las características de la serie. Hay dos fases para identificar el modelo

- a) Análisis de estacionariedad, en donde se determina el número de transformaciones necesarias para obtener una serie estacionaria
 1. Estacionariedad en varianza. Una serie será estacionaria cuando se mantenga el supuesto de que la variabilidad de la serie en torno a su media se mantiene constante a lo largo del tiempo. Cuando la serie no es estacionaria en varianza, se utilizan las transformaciones de Box-Cox, que incluyen una familia infinita de funciones. Como las series económicas suelen ser positivas y sin valores cero, la transformación más utilizada es la logarítmica.
Para analizar la estacionariedad en varianza de una serie se utilizan el gráfico de la serie original y el gráfico de las transformaciones correspondientes.
 2. Estacionariedad en media. Ahora, debemos identificar si la serie oscila en torno a un nivel constante o no, es decir, si es estacionaria en

media. Para llegar a esto, se analizarán las características que diferencian a las series estacionarias de las no estacionarias.

Características de las series estacionarias

- a) Fluctúan alrededor de una media constante.
- b) La función de autocorrelación decae exponencialmente.

Características de las series no estacionarias

- a) Presentan intervalos con medias diferentes.
- b) Los procesos con alguna raíz unitaria generalmente tienen una función de autocorrelación muestral con decrecimiento lento.
- c) Raíz autorregresiva cercana a la unidad.
- d) Sobrediferenciación al elegir un orden de integración d cuando la serie $\nabla^{d-1}X_t$ ya es estacionaria. Cabe mencionar que el objetivo en esta fase de la modelización es determinar el menor número de diferencias d que conviertan a la serie en estacionaria.

- b) Elección de los órdenes p y q . Una vez que se tiene una serie estacionaria, se determinará el proceso que lo generó.

Para esto, haremos uso de dos herramientas

1. Correlogramas muestrales y gráficos de la serie original.
2. Correlogramas muestrales y gráficos de determinadas transformaciones de la serie: logaritmos, diferencias, etc. (en caso de ser necesario).
3. Contrastes de la prueba de raíces unitarias, que permiten hacer inferencia sobre la existencia de raíces unitarias en una serie.
4. Contraste de la prueba de Dickey-Fuller aumentado.

Identificación de los órdenes p y q

Las funciones de autocorrelación contienen la información de las características dinámicas del proceso estacionario, por lo que será el instrumento básico para identificar los órdenes p y q del modelo ARMA adecuado para representar las características de la serie estacionaria $\{X_t\}$.

Para identificar los órdenes p y q , se compararán las funciones de autocorrelación muestral con las FAC teóricas de los modelos ARMA cuyas características ya conocemos (Cuadro F.1).

Los procesos estacionarios AR tienen

- a) FAC teóricas que se estabilizan hacia cero con algún tipo de decaimiento exponencial o un patrón de onda sinusoidal amortiguado.

- b) FAC teóricas que se cortan en cero después de retraso p (el orden del proceso AR)

Los procesos estacionarios MA tienen

- a) FAC teóricas que se cortan en cero después de retraso q (el orden del proceso MA) o bien se anulan para $j > q$.
- b) FACP teóricas que colapsan hacia cero con algún tipo de decaimiento exponencial o un patrón de onda sinusoidal amortiguado.

Los procesos estacionarios ARMA tienen

- a) FAC teóricas que se colapsan hacia cero después de los primeros rezagos $q-p$ y no se anula pero decrece rápidamente.
- b) FACP teóricas que colapsan hacia cero después del primer retardo $p-q$.

Proceso	FAC	FACP
AR	Decaimiento exponencial o variación cíclica.	Corta a cero después de <i>lag p</i> .
MA	Corta a cero después de <i>lag p</i> .	Decaimiento exponencial o variación cíclica.
ARMA	Se desvía a cero	Se desvía a cero.

Cuadro 3.1: Características distintivas de FAC y FACP para un proceso estacionario.

Inclusión del término independiente

La media de un proceso $ARMA(p, q)$ estacionario está directamente relacionada con la constante δ . Para saber si se incluye un término independiente no nulo en el modelo, se contrastarán las siguientes hipótesis

$$H_0 : E[X_t] = 0 \quad vs. \quad H_a : E[X_t] \neq 0. \quad (3.4.1)$$

El estadístico de contraste es

$$t = \frac{\bar{X}}{\hat{\sigma}} \sim t(T - d - 1) \quad (3.4.2)$$

donde $\hat{\sigma}_X^2$ es el estimador de la varianza de la media muestral \bar{X} que se define como

$$\hat{\sigma}_X^2 = \frac{C_0}{T - d} (1 + 2\hat{\rho}_1 + 2\hat{\rho}_2 + \dots + 2\hat{\rho}_n) \quad (3.4.3)$$

donde $C_0 = \sum \frac{(X_t - \bar{X})^2}{(T-1)}$ es la varianza muestral de la serie estacionaria y $(1 + 2\hat{\rho}_1 + 2\hat{\rho}_2 + \dots + 2\hat{\rho}_n)$ las n primeras autocorrelaciones muestrales significativas $\{X_t\}$.

Para calcular esta varianza se utiliza la aproximación

$$\hat{\sigma}_X^2 \sim \frac{C_0}{T - d} \quad (3.4.4)$$

se rechazará la hipótesis nula a un nivel de significancia α , dando como consecuencia la inclusión del parámetro δ en el modelo, si

$$t > t_{\frac{\alpha}{2}}(T - d - 1). \quad (3.4.5)$$

3.4.2. Estimación

Cuando ya se han identificado los modelos que pudieron haber generado a la serie $\{X_t\}$, se procede a estimar los parámetros desconocidos de los mismos

$$\beta = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q) \quad y \quad \sigma_a^2. \quad (3.4.6)$$

Estos parámetros se pueden estimar de forma consistente por Mínimos Cuadrados o Máxima Verosimilitud. Ambos métodos de estimación se basan en el cálculo de las innovaciones a partir de los valores de la serie estacionaria. El método de Mínimos Cuadrados minimiza la suma de cuadrados.

3.4.3. Validación

En esta fase se procede a determinar que tanto se ajustaron los modelos a los datos. Se debe tener en cuenta que

1. Si las estimaciones de los coeficientes del modelo son significativas y cumplen las condiciones de estacionariedad e invertibilidad que deben satisfacer los parámetros del modelo.
2. Si los residuos del modelo tienen un comportamiento similar a las innovaciones, es decir, si son Gaussianos.

En la etapa de verificación de diagnóstico determinamos si un modelo es estadísticamente adecuado. En particular, probamos si los choques aleatorios son independientes. Si no se cumple esta suposición, existe un patrón de autocorrelación en la serie original que no ha sido explicado por el modelo ARIMA.

Nuestro objetivo, sin embargo, es construir un modelo que explique completamente cualquier autocorrelación en la serie original.

En la práctica, no podemos observar los choques aleatorios en un proceso, pero tenemos estimaciones de ellos en forma de residuales en la etapa de estimación. Cuando los residuos están autocorrelacionados debemos considerar cómo el modelo ARIMA estimado podría ser reformulado. Algunas veces esto significa regresar para reexaminar las FAC y FACP estimados iniciales. Sin embargo, en esta etapa también se proporcionan pistas sobre cómo se podría mejorar el modelo.

- a) **La FAC residual.** Utilizada para probar la hipótesis de que los choques aleatorios son independientes.
- b) **Prueba-t.** Habiendo calculado y trazado las autocorrelaciones residuales, es importante determinar si cada uno es significativamente distinto de cero. Usamos la fórmula Bartlett, para estimar los errores estándar de las autocorrelaciones residuales. Cuando se aplica a autocorrelaciones residuales, la fórmula es

$$s[r_k(\hat{a})] = (1 + 2 \sum_{j=1}^{k-1} r_j(\hat{a})^2)^{\frac{1}{2}} n^{-\frac{1}{2}}. \quad (3.4.7)$$

Después de haber encontrado los errores estándar estimados de $r_k(\hat{a})$ de (3.4.7) podemos probar la hipótesis nula de que H_0 es $\rho_k(a) = 0$ para cada coeficiente de autocorrelación residual. El símbolo ρ y el a en paréntesis indica que estamos probando una hipótesis acerca de los choques aleatorios en un proceso.

No tenemos $p_k(a)$ valores disponibles, pero tenemos estimaciones de ellos en forma de autocorrelaciones residuales $r_k(\hat{a})$.

Probamos la hipótesis nula calculando cuántos errores estándar (t) se alejan de cero, cada coeficiente de autocorrelación residual cae

$$t = \frac{r_k(\hat{a}) - 0}{s[r_k(\hat{a})]}. \quad (3.4.8)$$

En la práctica, si el valor absoluto de un FAC residual es menor que 1.25 en los *lags* (también llamados retrasos) 1, 2, y 3, y menor que 1,6 en los *lags* más grandes, concluimos que los choques aleatorios en esos “lags” son independientes. Podríamos estar equivocados en esta conclusión, pero siempre corremos el riesgo cuando hacemos decisiones basadas sobre muestras de información.

c) Prueba chi-cuadrada (Apéndice B).

Un buen modelo

Al final el modelo debe cumplir con las siguientes características de un buen modelo (Alan Pankratz, 1983).

1. Parsimonioso. Esta característica usa el más pequeño número de coeficientes necesarios para explicar los datos disponibles.
2. Un buen modelo AR es estacionario cuando tiene coeficientes AR los cuales satisfacen algunas desigualdades matemáticas (Cuadro 3.2)

Tipo de Modelo	Condiciones de Estacionariedad
ARMA(0, q)	Siempre estacionario
AR(1) o ARMA(1, q)	$ \phi_1 < 1$
	$ \phi_2 < 1$
AR(2) o ARMA(2, q)	$\phi_2 + \phi_1 < 1$
	$\phi_2 - \phi_1 < 1$

Cuadro 3.2: Resumen de condiciones estacionarias para coeficientes AR (Alan Pankratz, 1983)

3. Un buen modelo MA es invertible cuando tiene coeficientes MA los cuales satisfacen algunas desigualdades matemáticas (Cuadro 3.3).
4. Alta calidad en los coeficientes.

(a) Prueba t . Cada coeficiente estimado tiene un error estándar porque es un estadístico basado en información muestral. Una muestra diferente podría dar diferentes estimaciones de ϕ_1 y θ_1 , así que cada coeficiente

Tipo de Modelo	Condiciones de Invertibilidad
ARMA($p, 0$)	Siempre invertible
MA(1) o ARMA($p, 1$)	$ \theta_1 < 1$
MA(2) o ARMA($p, 2$)	$ \theta_2 < 1$ $\theta_2 + \theta_1 < 1$ $\theta_2 - \theta_1 < 1$

Cuadro 3.3: Resumen de condiciones de invertibilidad para coeficientes ARMA (Alan Pankratz, 1983).

estimado tiene una distribución muestral con cierto error estándar. Una aproximación para lograr dicha estimación de tal coeficiente es aplicando la prueba $-t$ (también conocida como $t - value$) para probar la hipótesis de que el coeficiente verdadero es cero. Es calculada de la siguiente manera

$$t = \frac{\text{coeficiente estimado} - \text{valor del coeficiente especulado}}{\text{error estándar estimado del coeficiente}} \quad (3.4.9)$$

Como regla práctica deberíamos considerar no excluir algún coeficiente con un valor $|t| \geq 2$, esto significa que el coeficiente es significativamente distinto de cero a un nivel de confianza del 5%.

(b) ϕ 's y θ 's no están altamente correlacionados.

5. Residuales estadísticamente independientes.
 - a) FAC residual: Los correlogramas de los residuos estimados son puramente aleatorios.
 - b) Prueba t . Cuando ya se han calculado y graficado las autocorrelaciones residuales, es importante determinar si cada residual es significativamente distinto de cero.
 - c) Prueba Chi-Cuadrada y Prueba Ljung-Box (Apéndice B y C respectivamente).
6. Ajuste a los datos (el pasado) es bueno para adecuarse al análisis en la etapa de estimación .
 - a) Se agrega algún otro coeficiente para ver si el modelo resultante es mejor.
 - b) Ajustando un subconjunto de datos.
 - c) Reformulación del modelo.
7. Cuenta con un error de pronóstico bastante pequeño.

3.4.4. Predicción

La última etapa de la modelación ARIMA, es pronosticar valores futuros de una serie de tiempo. En primera consideraremos como se derivan algebraicamente los pronósticos *punto* (valor numérico único) de un modelo ARIMA estimado.

El álgebra de un pronóstico ARIMA

Forma ecuación diferencia. La forma más conveniente para producir pronósticos punto de un modelo ARIMA es escribiendo el modelo en la forma ecuación diferencia. Aquí mencionaremos algunos ejemplos.

Sea t el periodo de tiempo actual. Cuando pronosticamos estamos interesados en valores futuros de una serie de tiempo variable, denotada z_{t+l} , donde $l \geq 1$. Periodo t , es llamado el pronóstico origen, y l es llamada pronóstico *lead time*. En el análisis ARIMA, el pronóstico depende de las observaciones disponibles sobre la variable z hasta el periodo t . Permitamos que la información contenida en el conjunto de observaciones disponibles (z_t, z_{t-1}, \dots) sea designada como I_t . Entonces el pronóstico de z_{t+l} , es designado por \hat{z}_{t+l} que es la esperanza matemática de z_{t+l} dado I_t

$$\hat{z}_t(l) = E(z_{t+l}|I_t) \quad (3.4.10)$$

donde la línea vertical significa "dado".

Como ejemplo, consideremos al modelo ARIMA(1,0,1), que desarrollado en forma algebraica es

$$(1 - \phi_1 B)\tilde{z} = (1 - \theta_1 B)a_t \quad (3.4.11)$$

ó

$$z_t = \mu(1 - \phi_1) + \phi_1 z_{t-1} - \theta_1 a_{t-1} + a_t. \quad (3.4.12)$$

Ahora, sea $l = 1$. Usaremos (3.4.12) para cambiar del tiempo t para obtener la expresión z_{t+1} :

$$z_{t+1} = \mu(1 - \phi_1) + \phi_1 z_t - \theta_1 a_t + a_{t+1} \quad (3.4.13)$$

Aplicando (3.4.10) a (3.4.13), encontramos que el pronóstico de z_{t+1} es

$$\hat{z}_t(1) = E(z_{t+1}|I_t) = \mu(1 - \phi_1) + \phi_1 z_t - \theta_1 a_t. \quad (3.4.14)$$

Desde que a_{t+1} es desconocido el tiempo t , le asignamos su valor esperado de cero. En este ejemplo z_t y a_t juntos constituyen I_t . Eso es, z_t y a_t son toda la información relevante acerca de los z 's pasados necesarios para pronosticar z_{t+1} . (Recuerda que los términos MA son parsimoniosos sustitutos algebraicos para los términos AR; así a_t representa un conjunto de z 's pasados.)

Si ahora, $l = 2$, uso (3.4.12) para escribir una expresión para z_{t+2} . Entonces el valor esperado condicional de esa expresión es el pronóstico $\hat{z}_t(2)$:

$$\hat{z}_t(2) = E(z_{t+2}|I_t) = \mu(1 - \phi_1) + \phi_1 z_{t+1} - \theta_1 a_{t+1}. \quad (3.4.15)$$

Desde que z_{t+1} es desconocido en el origen t este debe ser reemplazado por su esperanza condicional $\hat{z}_t(1)$ desde (3.4.14). Igualmente, a_{t+1} es desconocido en el origen t y es reemplazado por su valor esperado cero. Con estas dos sustituciones, (3.4.15) llega a ser

$$\hat{z}_t(2) = \mu(1 - \phi_1) + \phi_1 \hat{z}_t(1). \quad (3.4.16)$$

Si seguimos el procedimiento como arriba, encontramos que cada subsecuente pronóstico para este ARIMA(1,0,1) está basado sobre el valor pronóstico anterior de z . Eso es, $\hat{z}_t(3)$ depende de $\hat{z}_t(2)$, $\hat{z}_t(4)$ depende de $\hat{z}_t(3)$ y así:

$$\hat{z}_t(3) = \mu(1 - \phi_1) + \phi_1 \hat{z}_t(2) \quad (3.4.17)$$

$$\hat{z}_t(4) = \mu(1 - \phi_1) + \phi_1 \hat{z}_t(3) \quad (3.4.18)$$

$$\hat{z}_t(5) = \mu(1 - \phi_1) + \phi_1 \hat{z}_t(4) \quad (3.4.19)$$

$$\begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{array}$$

En el ejemplo anterior, los pronósticos para $l > 1$ son llamados pronóstico *bootstrap* porque ellos son basados sobre pronósticos z 's más que z 's observados.

Los pronósticos de otros modelos ARIMA son encontrados esencialmente de la misma forma. En la práctica, μ es desconocido y es reemplazado por su media estimada $\hat{\mu}$. Igualmente, los coeficientes ϕ y θ son reemplazados por sus estimadores, $\hat{\phi}$ y $\hat{\theta}$.

Como se demostró en el ejemplo, las observaciones pasadas z son empleadas cuando están disponibles. Ellas están disponibles hasta el momento t , el pronóstico origen; por tanto, ellas deben ser reemplazadas por su correspondiente estimador, la estimación residual \hat{a}_t , cuando estos residuales están disponibles. Pero cuando el subíndice de tiempo sobre un choque aleatorio excede el pronóstico origen t , ese choque es reemplazado por su valor esperado cero. Esto es que sucedió cuando nos mudamos de (3.4.15) a (3.4.16) en el caso de ARIMA(1,0,1): aquí no hay estimación residual \hat{a}_{t+1} disponible cuando pronosticamos desde el origen t , así que nosotros sustituimos cero.

Forma choque-aleatorio. En un modelo ARIMA es su forma MA. Eso es, invirtiendo y expandiendo al operador AR, reemplazamos algún término AR con una serie infinita de término MA.

Los coeficientes en la forma choque-aleatorio son denotados por el símbolo ψ_t , con i correspondiente al tiempo *lag* de el choque-aleatorio asociado al pasado

$$z_t = \mu + \psi_0 a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \cdot \cdot \cdot \quad (3.4.20)$$

donde $\psi_0 = 1$.

La dispersión del pronóstico ARIMA

Usando la forma ecuación diferencia de un modelo ARIMA, podemos producir una serie de puntos pronóstico, donde "punto" significa el pronóstico de un solo valor más que un rango. Usando la forma de choque-aleatorio, podemos encontrar la varianza de los errores pronóstico. Esto nos permite construir intervalos de confianza aproximados alrededor de nuestros pronósticos, así daremos alguna información sobre como realmente el pronóstico podría ser.

Varianza pronóstico-error y desviación estándar

Primero, define un error de pronóstico para el origen t y al tiempo *lead* l , designamos a $e_t(l)$ como la z observada para el periodo $t+l$ menos el pronóstico z para ese periodo

$$e_t(l) = z_{t+l} - \hat{z}_t(l). \quad (3.4.21)$$

Usaremos (3.4.20) para escribir z_{t+l} en forma de choque-aleatorio como

$$z_t = \mu + \psi_0 a_{t+l} + \psi_1 a_{t-1+l} + \psi_2 a_{t-2+l} + \dots \quad (3.4.22)$$

El valor del pronóstico $\hat{z}_t(l)$, el cual es la esperanza matemática condicional $E(z_{t+l}|I_t)$, es encontrado desde (3.4.22) para ser

$$\hat{z}_t(l) = E(z_{t+l}|I_t) = \mu + \psi_0 a_{t+l} + \psi_1 a_{t-1+l} + \psi_2 a_{t-2+l} + \dots \quad (3.4.23)$$

Pronóstico desde datos en forma logarítmica.

Usualmente estamos interesados en pronosticar los datos originales más que los valores logarítmicos. Esto podría tentarnos a calcular el antilogaritmo de el pronóstico log para conseguir el pronóstico real, es decir, no debemos olvidar que si la varianza de una realización es convertida a estacionaria por una transformación de los datos originales dentro de valores log naturales (z 's), quizá no pronosticamos la serie original (z_t) solo buscando su antilogaritmo, si no que en su lugar, debemos tomar en cuenta que la varianza pronóstico-error de z'_{t+1} de el pronóstico log en esta forma es

$$\hat{z}_t(l) = \exp\{z'_t(l) + \frac{1}{2}\sigma^2[e'_t(l)]\}. \quad (3.4.24)$$

Así, no solo simplemente encontramos el antilog de $\hat{z}'_t(l)$. En su lugar, debemos tomar en cuenta la varianza de el pronóstico logarítmico como en (3.4.24). Sin embargo, los límites de confianza inferior y superior alrededor de $\hat{z}_t(l)$ son encontrados tomando los antilogaritmos de los límites alrededor de $\hat{z}'_t(l)$. Eso es, si U y L son los límites superior e inferior de un intervalo de confianza α percentil alrededor de $\hat{z}'_t(l)$, entonces $\exp(U)$ y $\exp(L)$ son los α - percentil límites superior e inferior alrededor de $\hat{z}_t(l)$. Esto es porque el intervalo alrededor de $\hat{z}_t(l)$ no es simétrico desde que el intervalo alrededor de $\hat{z}'_t(l)$ es simétrico.

Finalmente, nota que ese pronóstico de $z'_t(l)$ quizá podría ser interpretado en términos de z_t sin encontrar antilogaritmos porque el cambio de un valor log es el cambio porcentual del correspondiente valor antilog (Alan Pankratz, 1983).

Modelo de paseo aleatorio con deriva.

El paseo aleatorio con deriva resulta de añadir una constante al modelo.

$$X_t = X_{t-1} + a_t + C \quad (3.4.25)$$

$$\nabla X_t = a_t + C$$

En este caso, la inclusión de una constante en el modelo implica la inclusión de una tendencia determinista con pendiente C , junto con la tendencia estocástica (Gujarati, 2009)

Capítulo 4

Pronóstico del precio de las acciones de Grupo Financiero Inbursa

4.1. Aspectos Generales

Los datos con los que se trabajó en el presente estudio son del periodo Enero del 2006 a Diciembre 2016, proporcionados por el sitio web Yahoo Finance bajo el nombre GFINBURO.MX., considerandose un total de 132 observaciones. La variable bajo estudio serán las observaciones lineales de los precios de cierre de las acciones de Inbursa, de ahora en adelante se manejará con el nombre $GFIN_t$.

4.2. Análisis estadístico

Analizar los datos mediante la serie de tiempo de los precios de las acciones de Grupo Financiero Inbursa.

4.3. Análisis de los datos con metodología Box–Jenkins

4.3.1. Identificación del modelo

Estabilización de la varianza

En la Figura 4.1 se observa que la varianza de la serie no es constante a través del tiempo, y el nivel o tendencia es creciente, por tanto la serie no es estacionaria. Iniciaremos estabilizando la varianza con el fin de volver estacionaria la serie, para ello seleccionaremos una transformación estabilizadora de varianza, esta

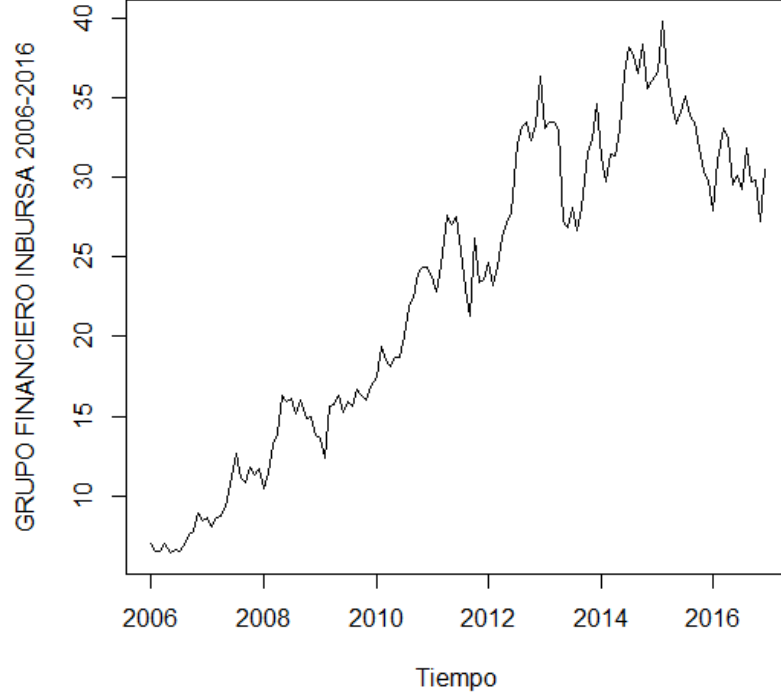


Figura 4.1: Precios de cierre mensual de las acciones GFINBURO.MX del 2006-2016.

transformación sugiere elegir la potencia λ , de tal manera que se cumpla la relación siguiente.

$$\text{constante} = \frac{\sigma_t}{\mu_t^{1-\lambda}}, t = 1, 2, \dots, N, \quad (4.3.1)$$

en donde σ_t y μ_t denotan a la desviación estándar y a la media del proceso Z_t y N es el número de observaciones. Guerrero (1993) sugiere un método para estabilizar la varianza de la serie, dicho método se trabaja de la siguiente manera, primero divídase a las N observaciones de la serie en H grupos que contengan $R = (N - n)/H$ observaciones contiguas cada uno, dejando fuera de los cálculos a un total de n observaciones ($0 \leq n < R$) ya sea del principio o del final de la serie; si Z_r es la r -ésima observación del grupo H , entonces

$$\bar{Z}_h = \sum_{r=1}^R \frac{Z_{h,r}}{R}, S_h = \sqrt{\sum_{r=1}^R \frac{(Z_{h,r} - \bar{Z}_h)^2}{(R-1)}}. \quad (4.3.2)$$

Se debe cuidar que exista homogeneidad entre los grupos y que éstos contengan el mismo número de observaciones, de esta forma se tendrán H parejas de valores comparables $\{S_h, \bar{Z}_h\}$, $h = 1, \dots, H$ con los cuales se construye el Cuadro 4.1. El coeficiente de variación se calcula de la siguiente manera

Grupo	Potencia (λ)				
	-1	-0.5	0	0.5	1
1	$\frac{S_1}{\bar{Z}_1^2}$	$\frac{S_1}{\bar{Z}_1^{1.5}}$	$\frac{S_1}{\bar{Z}_1}$	$\frac{S_1}{\bar{Z}_1^{0.5}}$	S_1
2	$\frac{S_2}{\bar{Z}_2^2}$	$\frac{S_2}{\bar{Z}_2^{1.5}}$	$\frac{S_2}{\bar{Z}_2}$	$\frac{S_2}{\bar{Z}_2^{0.5}}$	S_2
...		
h	$\frac{S_h}{\bar{Z}_h^2}$	$\frac{S_h}{\bar{Z}_h^{1.5}}$	$\frac{S_h}{\bar{Z}_h}$	$\frac{S_h}{\bar{Z}_h^{0.5}}$	S_h
...		
H	$\frac{S_H}{\bar{Z}_H^2}$	$\frac{S_H}{\bar{Z}_H^{1.5}}$	$\frac{S_H}{\bar{Z}_H}$	$\frac{S_H}{\bar{Z}_H^{0.5}}$	S_H
Coeficiente de variación	CV(-1)	CV(-0.5)	CV(0)	CV(0.5)	CV(1)

Cuadro 4.1: Cálculos para seleccionar la potencia λ de la transformación estabilizadora de varianza.

$$CV(\lambda) = \frac{de(\lambda)}{M(\lambda)} \tag{4.3.3}$$

donde $M(\lambda)$ se calcula de la siguiente forma:

$$M(\lambda) = \sum_{h=1}^H \frac{\left(\frac{S_h}{\bar{Z}_h^{1-\lambda}}\right)}{H} \tag{4.3.4}$$

y $de(\lambda)$ es la desviación estándar de λ y se calcula como sigue

$$de(\lambda) = \sqrt{\sum_{h=1}^H \frac{\left[\left(\frac{S_h}{\bar{Z}_h^{1-\lambda}}\right) - M(\lambda)\right]^2}{(H-1)}}. \tag{4.3.5}$$

Puesto que se desea satisfacer la relación (4.3.3) con los valores muestrales, deberá elegirse la potencia que proporcione el mínimo coeficiente de variación, debido a que ese valor satisficará de manera aproximada la relación

$$\frac{S_h}{\bar{Z}_h^{1-\lambda}} = \text{constante}, \quad \text{para } h = 1, \dots, H \tag{4.3.6}$$

y así la serie transformada resultante será

$$T(Z_t) = \begin{cases} Z_t^\lambda & , \quad \text{si } \lambda \neq 0 \\ \log(Z_t) & , \quad \text{si } \lambda = 0 \end{cases} \tag{4.3.7}$$

Nótese el espaciamiento entre los valores consecutivos de λ en el (Cuadro 4.2), ésta es una sugerencia y podría ensayarse un espaciamiento diferente al mostrado para encontrar con la exactitud deseada aquel valor de λ que vuelva mínimo el coeficiente de variación. Debe tenerse en cuenta, que el método está basado en aproximaciones. Por otro lado, en ocasiones el conocimiento del fenómeno en estudio puede indicar el tipo de transformación por ser empleada, como sucede con la logarítmica, cuyo uso resulta adecuado cuando existe interés en analizar las variaciones porcentuales de la serie.

Potencia (λ)						
Grupo	-0.5	0	0.5	-0.5	0	0.5
2006 (1)	$\frac{S_1}{Z_1^{1.5}}$	$\frac{S_1}{Z_1}$	$\frac{S_1}{Z_1^{0.5}}$	0.042	0.113	0.303
2007 (2)	$\frac{S_2}{Z_2^{1.5}}$	$\frac{S_2}{Z_2}$	$\frac{S_2}{Z_2^{0.5}}$	0.046	0.149	0.470
2008 (3)	$\frac{S_3}{Z_3^{1.5}}$	$\frac{S_3}{Z_3}$	$\frac{S_3}{Z_3^{0.5}}$	0.030	0.130	0.495
2009 (4)	$\frac{S_4}{Z_4^{1.5}}$	$\frac{S_4}{Z_4}$	$\frac{S_4}{Z_4^{0.5}}$	0.020	0.082	0.325
2010 (5)	$\frac{S_5}{Z_5^{1.5}}$	$\frac{S_5}{Z_5}$	$\frac{S_5}{Z_5^{0.5}}$	0.027	0.125	0.570
2011 (6)	$\frac{S_6}{Z_6^{1.5}}$	$\frac{S_6}{Z_6}$	$\frac{S_6}{Z_6^{0.5}}$	0.016	0.083	0.413
2012 (7)	$\frac{S_7}{Z_7^{1.5}}$	$\frac{S_7}{Z_7}$	$\frac{S_7}{Z_7^{0.5}}$	0.027	0.149	0.810
2013 (8)	$\frac{S_8}{Z_8^{1.5}}$	$\frac{S_8}{Z_8}$	$\frac{S_8}{Z_8^{0.5}}$	0.017	0.099	0.548
2014 (9)	$\frac{S_9}{Z_9^{1.5}}$	$\frac{S_9}{Z_9}$	$\frac{S_9}{Z_9^{0.5}}$	0.015	0.088	0.520
2015 (10)	$\frac{S_{10}}{Z_{10}^{1.5}}$	$\frac{S_{10}}{Z_{10}}$	$\frac{S_{10}}{Z_{10}^{0.5}}$	0.013	0.081	0.473
2016 (11)	$\frac{S_{11}}{Z_{11}^{1.5}}$	$\frac{S_{11}}{Z_{11}}$	$\frac{S_{11}}{Z_{11}^{0.5}}$	0.03	0.056	0.310
Coefficiente de variación	CV(-0.5)	CV(0)	CV(0.5)	0.481	0.289	0.30

Cuadro 4.2: Cálculos para seleccionar la potencia λ de la transformación estabilizadora de varianza.

Ahora, consideremos la serie en estudio para realizar la aplicación del procedimiento anterior. Dicha serie cubre el periodo que va de enero 2006 a finales de diciembre de 2016, por lo cual $N = 132$. De manera un tanto arbitraria se dividió la serie en subseries anuales, en nuestro caso ningún dato se ha quedado fuera así $n = 0$, $N - n = 132 - 0$, $H = 10$ y $R = 12$.

El Cuadro 4.2 muestra los cálculos necesarios para determinar la transformación potencia estabilizadora de varianza para esta serie.

En el cuadro 4.3 se observa que el mínimo coeficiente de variación (dentro del grupo de valores λ elegido) se logra con $\lambda = 0$, por lo cual una posible transformación para estabilizar aproximadamente la varianza de $\{GFINT_t\}$ (GFINT por sus siglas de Grupo Financiero Inbursa) es

$$T(GFINT) = \log(GFINT_t), \text{ para } \lambda = 0. \quad (4.3.8)$$

	Potencia (λ)				
Grupo	-1	-0.5	0	0.5	1
Coefficiente de variación	0.766	0.481	<u>0.289</u>	0.30	0.443

Cuadro 4.3: Determinando el mínimo coeficiente de variación.

Estabilización del nivel y empleo de la FAC

Una vez determinada la transformación apropiada para estabilizar la varianza de una serie para obtener $T(GFIN_t)$, se procede a estabilizar el nivel de la serie mediante la aplicación del operador diferencia un número apropiado de veces. La principal herramienta para determinar el grado de diferenciación apropiado es la FAC muestral, ya que un decaimiento rápido de las autocorrelaciones a cero es indicativo de que la serie es estacionaria, en cuanto a nivel se refiere. En la Figura 4.2 podemos observar a la FAC muestral de los datos originales, es decir, sin ninguna transformación ni diferenciación, después vamos a observar el comportamiento de la serie transformada y a continuación con uno y dos grados de diferencia, pues por lo regular no es necesario recurrir a más niveles de diferencia, además se debe evitar la sobrediferenciación de la serie porque esto podría causar problemas al tratar de identificar un posible proceso generador de la serie observada.

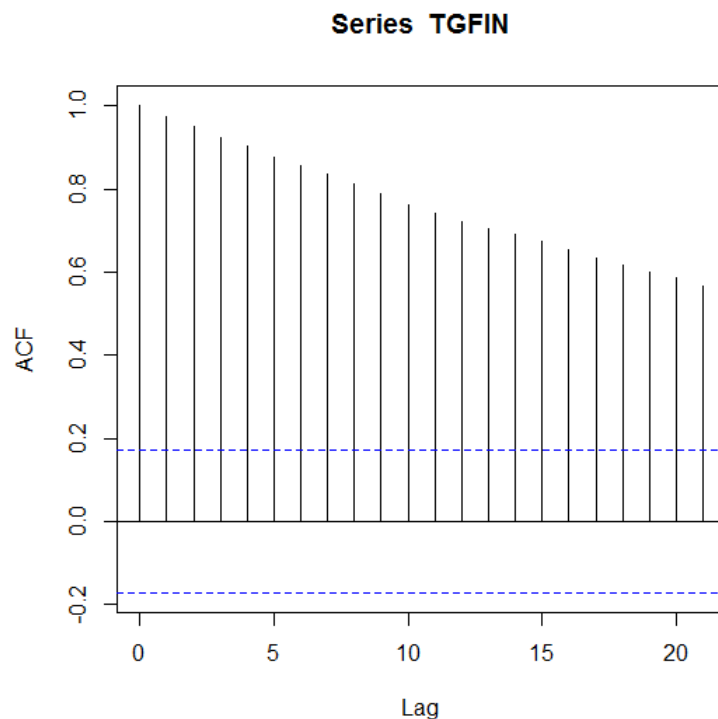


Figura 4.2: FAC estimada de GFINBURO y $T(GFIN_t)$.

El grado de diferenciación requerido para volver estacionaria la serie, podría determinarse mediante el cálculo de la desviación estándar muestral de las series $\{T(GFIN_t)\}$, $\{\nabla T(GFIN_t)\}$, $\{\nabla^2 T(GFIN_t)\}$, que se denotará por $S(0)$, $S(1)$

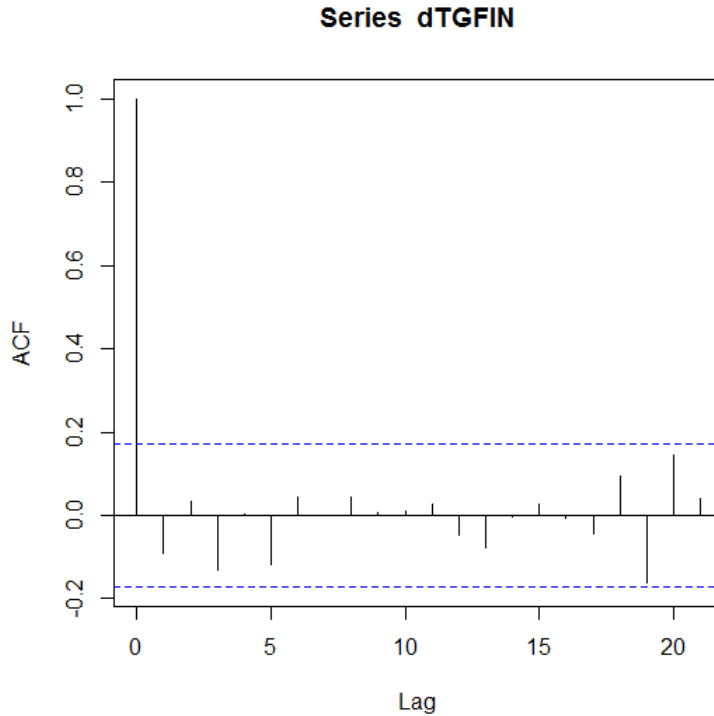


Figura 4.3: FAC estimada de $\nabla T(GFIN_t)$.

y $S(2)$, si d es el grado de diferenciación requerido, entonces

$$S(d) = \min\{S(j), j = 0, 1, 2\}. \quad (4.3.9)$$

Los resultados obtenidos con la ayuda de R-program son los siguientes

$$S(d) = \min\{S(0) = 0.51, S(1) = 0.07, S(2) = 0.10\} = 0.0717. \quad (4.3.10)$$

Así que el valor más pequeño como lo vemos en la ecuación 4.3.10 es $S(1)$ o bien la desviación estándar de $\nabla T(GFIN_t)$. Esta herramienta es más utilizada como complemento de las FACs muestrales, en ellas se puede apreciar el comportamiento de la FAC $\nabla T(GFIN_t)$, donde hasta aquí hemos observado que ninguna rebasó los límites de las bandas.

Ahora bien, para decidir si las autocorrelaciones son cero ó no a partir de un cierto lapso q , deben compararse los valores r_k con sus correspondientes desviaciones estándar, es decir, se utilizarán las autocorrelaciones muestrales del cuadro en la Figura 4.4, correspondientes a la serie $\nabla T(GFIN_t)$; bajo el supuesto de que sólo una diferencia es requerida para volver estacionaria la serie $T(GFIN_t)$. Si se parte del supuesto de que $\rho_k = 0$ para $k > 0$ (lo cual significaría que la serie es ruido blanco), la desviación estándar estimada de r_k , para

Autocorrelations of series 'dTGFIN', by lag											
k	0	1	2	3	4	5	6	7	8	9	10
r _k	1.000	-0.091	0.034	-0.132	0.003	-0.120	0.042	0.000	0.043	0.008	0.009
k	11	12	13	14	15	16	17	18	19	20	21
r _k	0.027	-0.048	-0.078	-0.005	0.027	-0.007	-0.045	0.095	-0.161	0.145	0.039

Figura 4.4: Valores función de autocorrelación muestral de $\nabla T(GFIN_t)$.

$k=1,2,\dots$ viene dada por

$$\sqrt{\widehat{Var}(r_k)} = \sqrt{\frac{1}{132-1}[1+2(0)]} = 0.087. \quad (4.3.11)$$

Es decir, para lograr estabilizar el nivel exitosamente se requiere realizar solo una diferenciación, $d = 1$. Al comparar el valor de r_1 con esta desviación estándar se tiene que $|r_1|$ es mayor que su desviación estándar, lo cual implica que casi es razonable suponer que $\rho_k = 0$. La razón para tal afirmación se basa en que, cuando $N \geq 50$, la distribución de la autocorrelación r_k puede aproximarse por una distribución normal y se sabe que para una distribución normal con media cero, los valores $\pm 1,96 * (\text{desviación estándar})$ constituyen los límites para determinar observaciones significativamente distintas de cero en el nivel de significación de 5% es decir

$$\begin{aligned} [\bar{x} - 1.96\sigma, \bar{x} + 1.96\sigma] &= [0 - 1.96(0.087), 0 + 1.96(0.087)] \\ &= [-0.17052, 0.17052] \end{aligned} \quad (4.3.12)$$

Una forma más sencilla de conocer si una autocorrelación r_k es significativamente distinta de cero es si

$$|r_k| > 2 \sqrt{\frac{1}{N-d} \left(1 + 2 \sum_{j=1}^q r_j^2\right)} \quad \text{para } k > q. \quad (4.3.13)$$

Supóngase ahora que $\rho_1 \neq 0$, pero que $\rho_k = 0$, para $k > 1$; en estas condiciones, la desviación estándar estimada de r_k , para $k = 2, 3, \dots$ es

$$\sqrt{\widehat{Var}(r_k)} = \sqrt{\frac{1}{132-1}[1+2(-0.091)^2]} = 0.0880 \quad (4.3.14)$$

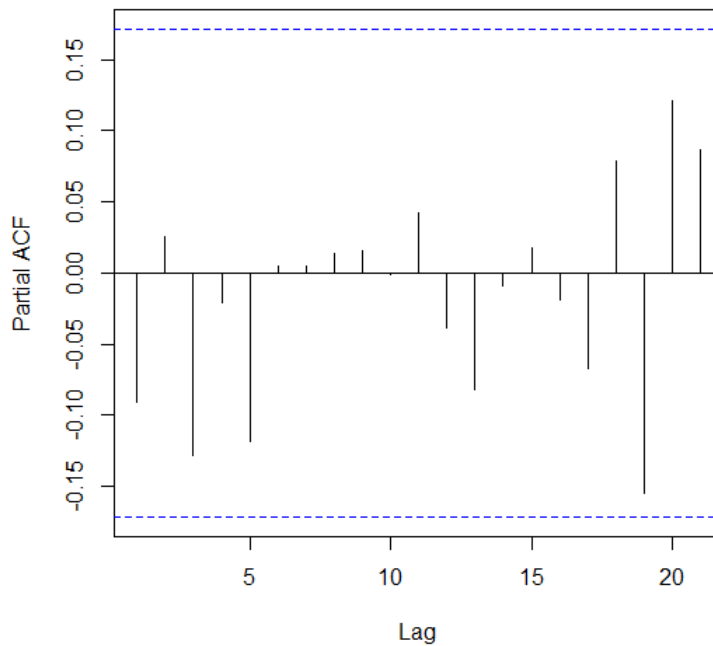
$$|r_k| > 2 \sqrt{\widehat{Var}(r_k)} = 2 * \sqrt{\frac{1}{132-1}[1+2(-0.091)^2]} = 2 * 0.0880 = 0.176 \quad (4.3.15)$$

Puesto que ninguna autocorrelación muestral con retraso $k \geq 2$ satisface la relación (4.3.15) para $q = 1$, resulta razonable suponer que ninguna autocorrelación es distinta de cero, lo cual conduce a postular como un posible modelo para la serie a un modelo ARIMA(0,1,0), con lo cual se tendría tentativamente identificado un modelo para la serie $T(GFIN_t)$. De aquí se pasa entonces a la siguiente etapa del proceso Box-Jenkins, la etapa de estimación de parámetros donde haremos uso de la FACP.

4.3.2. Estimación de parámetros

Empleo de la función de autocorrelación parcial

Se ha logrado observar que para la identificación de un proceso de promedios móviles puro puede llevarse a cabo con el empleo de la FAC muestral y con la aproximación de la varianza en las autocorrelaciones respectivas, pero en el caso de un proceso autorregresivo, la identificación del modelo a partir de la FAC muestral no es tan simple, de hecho en general, el orden de un proceso $AR(p)$ no es posible detectarlo con el solo uso de la FAC muestral.



Partial autocorrelations of series 'dTGFIN', by lag											
k	1	2	3	4	5	6	7	8	9	10	
r _k	-0.091	0.026	-0.128	-0.021	-0.118	0.005	0.005	0.014	0.016	-0.001	
k	11	12	13	14	15	16	17	18	19	20	21
r _k	0.042	-0.038	-0.082	-0.009	0.017	-0.019	-0.067	0.079	-0.155	0.121	0.087

Series dTGFIN

Figura 4.5: Función de Autocorrelación Parcial muestral de $\nabla T(GFIN_t)$ y valores de la función parcial muestral.

Por lo anterior, se requiere de la FACP porque nos permite encontrar el orden de procesos AR, la FACP adquiere determinadas características que dependen del orden del proceso y del tipo de parámetros involucrados. El cálculo de las autocorrelaciones parciales se logro con el apoyo de R-program (ver Figura 4.5),

para así proceder a una aproximación sugerida por Quenouille (1949), quien indica que si el proceso es $AR(p)$, las autocorrelaciones parciales muestrales se distribuyen de manera independiente, con media

$$E(\widehat{\phi}_{ii}) = \phi_{ii}. \quad (4.3.16)$$

y con

$$Var(\widehat{\phi}_{ii}) = \frac{1}{N-d} = \frac{1}{131} = 0.007 \quad \text{para } i > p. \quad (4.3.17)$$

A partir de esta expresión, se establece que ϕ_{ii} es distinto de cero (a un nivel de significancia de aproximadamente 5%) si el valor calculado de $\widehat{\phi}_{ii}$ se encuentra fuera del intervalo definido por

$$|2\sqrt{Var(\widehat{\phi}_{ii})}| = |2 * \sqrt{0.007}| = 0.174 \quad \text{para } i > p. \quad (4.3.18)$$

Sobra decir, que ninguno de los valores ϕ_{ii} se encuentra fuera del intervalo estimado, eso significa que ningún valor es significativamente distinto de cero a un nivel de significancia del 5% en un proceso autorregresivo para este caso.

Por consiguiente, de la etapa de identificación de modelos para la serie $T(GFIN_t)$, surge como posibilidad el modelo $ARIMA(0, 1, 0): \nabla T(GFIN_t) = a_t$ o $T(GFIN_t) = T(GFIN_{t-1}) + a_t$.

4.3.3. Verificación

Análisis de los residuos

Una de las formas para detectar las violaciones a los supuestos de los modelos es a través del análisis de residuos, donde los residuos miden la discrepancia entre los valores observados y los valores estimados del modelo. Además, cuando el tamaño de la muestra es grande, los errores aleatorios y los residuos (que también son variables aleatorias) son esencialmente iguales; por esta razón, al analizar los residuos observados $\{\widehat{a}_t\}$ se analiza básicamente lo que debería ser una realización del proceso de ruido blanco $\{a_t\}$.

Por lo anterior, los supuestos acerca del proceso $\{a_t\}$ pueden verificarse y posiblemente corregirse de la siguiente manera.

SUPUESTO 1. $\{a_t\}$ tiene media cero

Calcúlese la media aritmética de los residuos

$$m(\widehat{a}) = \sum_t^N \frac{\widehat{a}_t}{N-d-p} = 0.0119 \quad \text{con } d=1 \quad \text{y } p=0. \quad (4.3.19)$$

SUPUESTO 2. $\{a_t\}$ tiene varianza constante.

La gráfica de los residuos contra el tiempo nos ayuda visualmente para conocer si la varianza parecer o no ser constante y como vemos en la Figura 4.7, el comportamiento de la varianza en gran mayoría parecer ser constante.

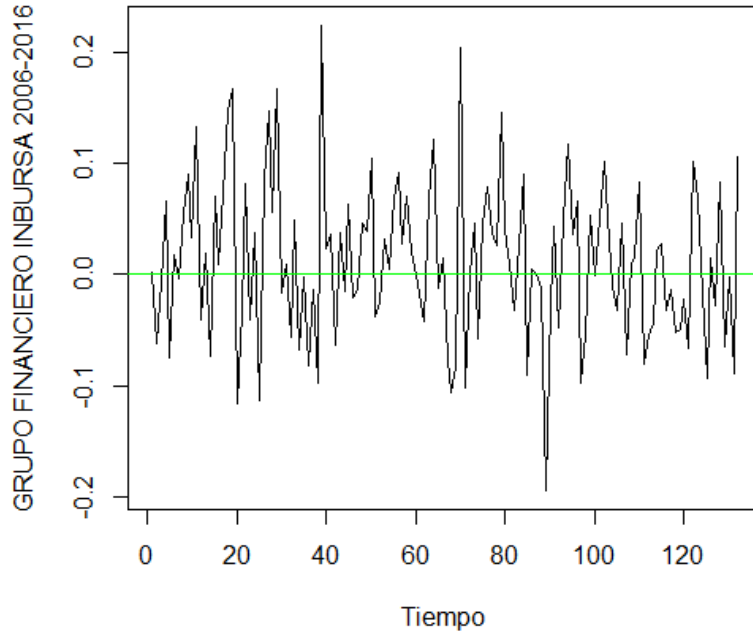


Figura 4.6: Media residual de GFIN.

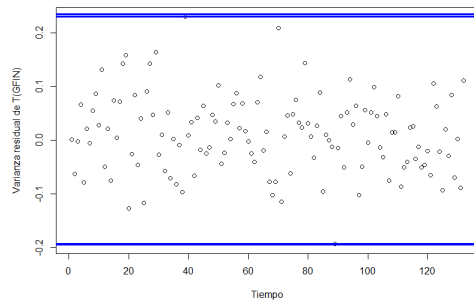


Figura 4.7: Varianza residual de GFIN.

SUPUESTO 3. Las variables aleatorias $\{a_t\}$ son mutuamente independientes.

i) FAC muestral de los residuos.

Debido a que la independencia significa no autocorrelación, se pueden calcular los valores de la FAC muestral de los residuos $\{r_k(\hat{a}_t)\}$, con media 0 (4.3.19) y

desviación estándar

$$\sqrt{V\widehat{ar}[r_k(\widehat{a}_t)]} = \frac{1}{\sqrt{N-d-p}} = 0.07121138 \quad (4.3.20)$$

Para determinar la significación estadística individual de las autocorrelaciones de los residuos se dira que la autocorrelación k -ésima es significativamente distinta de cero si

$$|r_k(\widehat{a}_t)| \geq \frac{2}{\sqrt{N-d-p}} = \frac{2}{\sqrt{131}} = 0.1747 \quad (4.3.21)$$

Autocorrelations of series 'residual', by lag											
k	0	1	2	3	4	5	6	7	8	9	10
r _k	1.000	-0.029	0.025	-0.133	-0.010	-0.119	0.036	0.004	0.044	0.011	0.009
k	11	12	13	14	15	16	17	18	19	20	21
r _k	0.025	-0.051	-0.080	-0.009	0.026	-0.009	-0.042	0.082	-0.146	0.140	0.051

Figura 4.8: k -ésima autocorrelación r_k .

Se puede observar mejor en la Figura 4.8 y 4.9, que en ninguno de los casos los valores de $r_k(\widehat{a}_t)$ logra rebasar la regla de significancia establecida, con esto podemos decir que las variables aleatorias son mutuamente independientes.

ii) Prueba Ljung-Box (ver Apéndice C).

Se debe probar si existe independencia entre los rezagos de la serie de tiempo, utilizando la prueba Ljung-Box. En nuestra prueba de hipótesis

$$H_0 : a_t \approx WN(0, \sigma^2) \quad (4.3.22)$$

$$H_1 : a_t \text{ no son } WN(0, \sigma^2) \quad (4.3.23)$$

La prueba fue realizada en R-program, la cual nos proporciona un valor $-p = 0.7331$, indicador de que si $p > 0.05$ entonces no se rechaza H_0 . Es decir, los residuales de nuestro modelo son ruido blanco o bien, mutuamente independientes.

SUPUESTO 4. a_t tiene una distribución normal, para toda t .

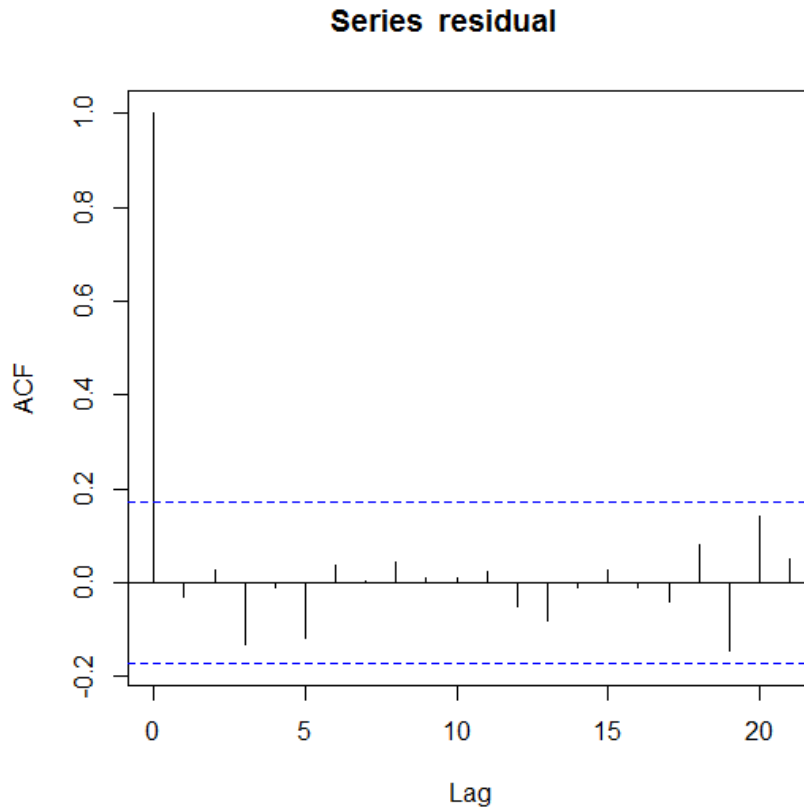


Figura 4.9: Gráfico de FAC de las autocorrelaciones de r_k .

Se sabe que para una distribución normal, aproximadamente el 95 % de las observaciones deben estar dentro de un intervalo que se extienda dos desviaciones estándar por abajo y por arriba de la media; entonces, si se cumple que la media de los residuos sea cero, se esperaría que a lo más un total de $\frac{N-d-p}{20} = 6.55 \approx 7$ observaciones estarán fuera del intervalo $(-2\hat{\sigma}_a, 2\hat{\sigma}_a) = (-2 * 0.07119309, 0.1423862)$, lo podemos ver en el gráfico de los residuos (Figura 4.11) pues solamente 6 son las observaciones que se encuentran fuera del rango permitido y en el histograma (Figura 4.12) apreciamos la forma de nuestros datos.

SUPUESTO 5. Implícitamente se ha supuesto que no existen observaciones posiblemente ajenas a la serie de estudio.

La gráfica de residuos contra el tiempo permitirá visualizar si existe este tipo de observaciones anómalas. Por ejemplo, un residuo que se encuentre fuera de $(-3\hat{\sigma}_a, 3\hat{\sigma}_a)$ implicará que, o bien sucedió un evento cuya probabilidad de

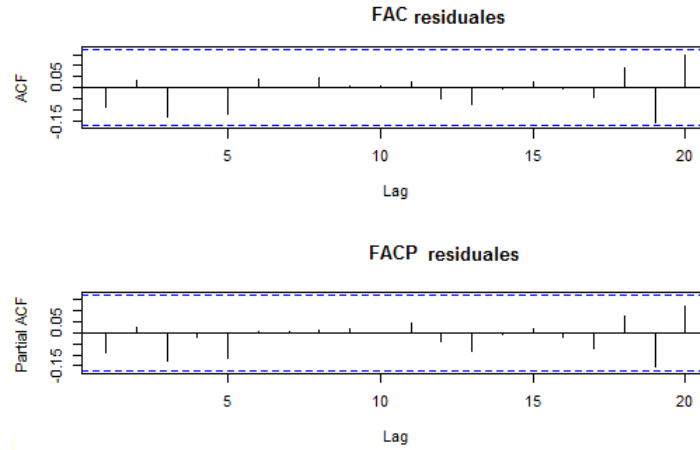


Figura 4.10: FAC y FACP de los residuales del modelo ARIMA(0,1,0).

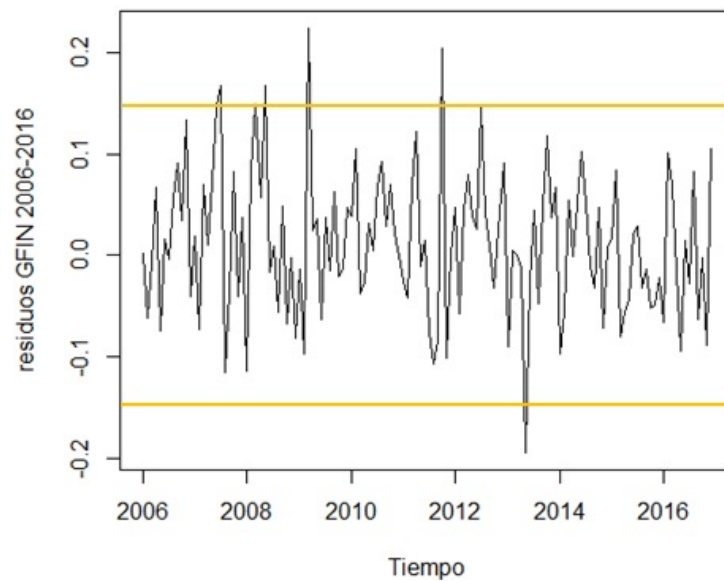


Figura 4.11: Gráfico de los límites de los residuales para una distribución normal.

ocurrencia era de aproximadamente 0.2% (lo cual sería muy extraño) o el residuo en cuestión corresponde a una observación que no fue generada por el mismo proceso generador del resto de la serie. De esta forma, como una regla empírica de trabajo, podría considerarse “sospechosas” las observaciones cuyos residuos estén fuera del intervalo $(-3\hat{\sigma}_a, 3\hat{\sigma}_a)$ es decir, el tamaño de nuestro

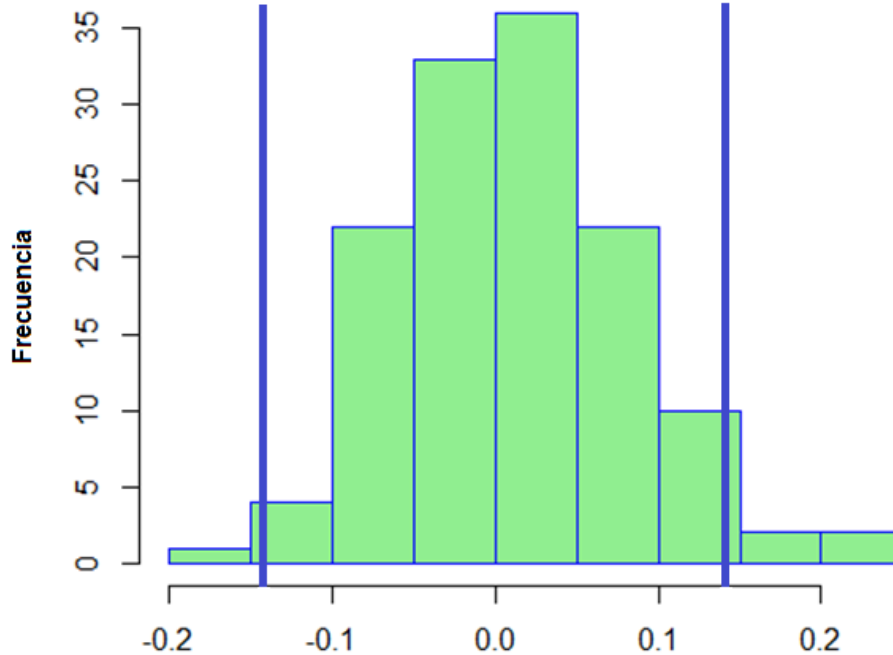


Figura 4.12: Histograma con los límites de los residuales $(-2\hat{\sigma}_a, 2\hat{\sigma}_a)$.

bandas límites es el siguiente

$$(-3\hat{\sigma}_a, 3\hat{\sigma}_a) = (-3 * 0.07119309, 0.213579). \quad (4.3.24)$$

En esta parte de los supuestos hemos detectado tal observación "sospechosa" señalada en la Figura 4.13 correspondiente a la fecha del mes de febrero año 2009 no se puede omitir ya que generaría una distorsión de nuestro modelo, es por esto que no hay que descartar la posibilidad de que realmente haya ocurrido un evento altamente improbable. Por lo mismo, se realizó una investigación para conocer la causa de tal cambio en el comportamiento del fenómeno, se descubrió que fue causado por una "intervención" exógena a la serie en estudio, tal intervención fue que durante el inicio del periodo presidencial del entonces presidente Felipe Calderón Hinojosa (2006–2012), el precio del barril de petróleo se elevó demasiado, ocasionando una reacción en cadena, dando origen al aumento de la inflación en la economía mexicana, las personas al verse carentes de dinero con los altos costos del mercado, deciden recurrir a diferentes instituciones bancarias para tramitar créditos (buscando una salida para obtener los recursos deseados), desafortunadamente la mayoría de los bancos les negó la autorización de tales trámites, a excepción de Banco Inbursa, quien decide dar el voto de confianza

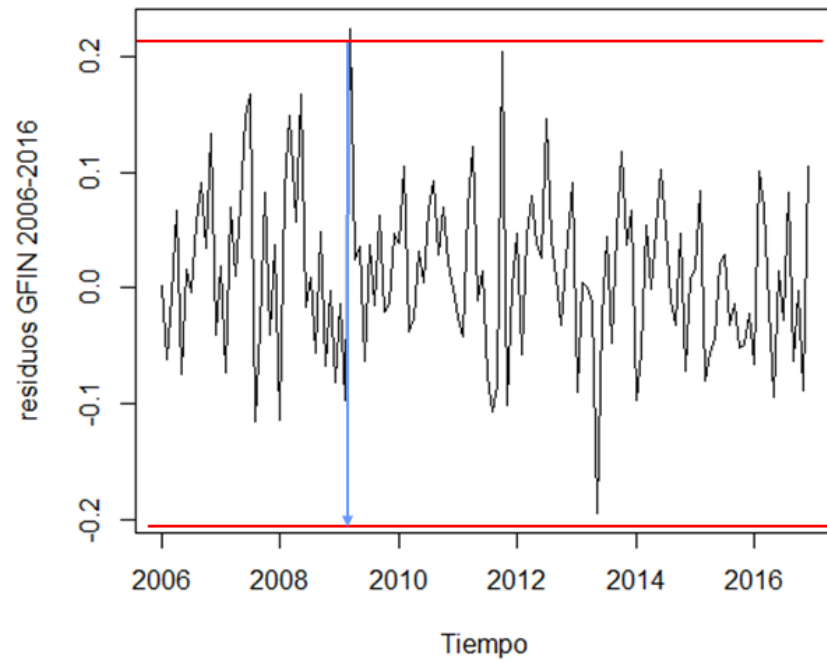


Figura 4.13: Límites para conocer observaciones poco probables.

(José Blanco, 2009). El resultado de tal decisión se ve reflejado al momento en que la economía logra estabilizarse, pagando los créditos otorgados más el pago de los intereses, esto eleva demasiado los precios de las acciones de una forma poco peculiar y común.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Modelo	Parámetro estimado	Intervalos del 95% de confianza	Correlación >0.5 o <-0.5	$m(\hat{a})$ y $\hat{\sigma}_a$	Ljung-Box valor p	$r_k(\hat{a}) \neq 0$ $\frac{z}{\sqrt{N-d}} = 0.183$	Residuales grandes
$(1 - \phi B)\nabla T(Z_t)$ $= (1 - \theta) a_t$	$\phi = -0.9496$ $\theta = 0.9012$	(0.72, 1.16) (0.60, 1.20)	0.96 Alta	0.011373 y 0.07065	0.7627	-	$\hat{a}_{39} = 3.12\hat{\sigma}_a$
$\nabla T(Z_t) = (1 - \theta)a_t$	$\theta = -0.036$	(-0.2196, 0.1468)	-	0.011836 y 0.07121	0.7331	-	$\hat{a}_{39} = 3.12\hat{\sigma}_a$
$\nabla T(Z_t) = a_t$	-	-	-	0.011169 y 0.07146	0.2957	-	$\hat{a}_{39} = 3.18\hat{\sigma}_a$

Figura 4.14: Resumen de resultados de la estimación de modelos.

La selección del modelo ARIMA por medio del Criterio de Información de Akaike(AIC) y Criterio de Información Bayesiano (BIC) (Cuadro 4.4) se establece el mejor modelo ARIMA. El modelo con menor AIC y BIC es el modelo ARIMA seleccionado.

Modelo	Prueba	Prueba
	AIC	BIC
ARIMA(0,1,0)	-315.56	-309.80
ARIMA(1,1,1)	-313.02	-307.82
ARIMA(0,1,1)	-312.87	-307.97

Cuadro 4.4: Comparación con prueba Akaike.

Para fines prácticos, resulta conveniente utilizar un formato único donde presentar los resultados numéricos necesarios para llevar a cabo la verificación de modelos, de tal manera que, con la ayuda de este formato (Figura 4.17) y el análisis de los residuales se logró llegar a la elección del modelo ARIMA(0,1,0), las razones son las siguientes

1. La FAC estimada de la serie diferenciada de INBURSA demostrada en la Figura (4.3), no decae rápidamente hacia cero, sugiriendo que la media de la primer diferencia sea estacionaria. Aquí, la primer diferencia parece ser una serie ruido blanco, sugiriendo el siguiente modelo ARIMA para los datos diferenciados

$$w_t = a_t \quad (4.3.25)$$

2. Al volver a realizar la estimación, se obtiene un modelo ARIMA(0,1,0) también conocido como caminata aleatoria presentado en la figura 4.17 como

$$(1 - B)\tilde{z} = a_t. \quad (4.3.26)$$

en su notación retroceso, donde $(1 - B)^d$ es el operador diferencia con

$d = 1$ ó bien $(1 - B)$, donde d indica el número de veces en que los datos han sido diferenciados (Alan Pankratz, 1983).

4.3.4. Predicción

El último paso de la modelación UBJ-ARIMA es pronosticar valores futuros de una serie de tiempo. Anteriormente se mencionó que la serie fue diferenciada una vez y aparentemente es compuesta de elementos estadísticamente independientes. El modelo que se obtuvo para ajustar los datos fue el modelo (4.3.26) el cual, ofrece una explicación adecuada estadísticamente. El FAC estimado en primer diferencia (Figura 4.3) no indica la necesidad de más términos MA ó AR.

Escribiendo el modelo (4.3.26) en ecuación diferencia obtenemos

$$z_t = z_{t-1} + a_t. \quad (4.3.27)$$

Desde (4.3.27) podemos observar que nuestro pronóstico (\hat{z}) ARIMA es simplemente el último dato observado: $\hat{z} = z_{t-1}$. Como no conocemos a_t o tenemos alguna estimación de la misma, ajustamos a_t a su valor esperado cero. Sin embargo, desconocemos el último valor observado $z(z_{t-1})$.

El modelo (4.3.27) no contiene un término constante, esto es porque z_t no es estacionaria y este no varía alrededor de una media fija pero, la serie en primer diferencia $w_t = z_t - z_{t-1}$ parece tener una media fija alrededor de cero, así que no existe un elemento con tendencia determinística (constante) en z_t . Si w_t no tiene media cero, entonces (4.3.26) y (4.3.27) tienen un término constante (C) igual a la media de w_t (Alan Pankratz, 1983). Entonces el modelo para w_t sería

$$(w_t - u_w) = a_t \quad (4.3.28)$$

Por tanto, realizaremos una prueba de hipótesis que nos muestra como la caminata aleatoria cuenta con una media distinta de cero la cual, nos indicará la tendencia de los precios de las acciones de INBURSA

$$H_0 : \mu_w = 0 \text{ vs } H_1 : \mu_w \neq 0 \quad (4.3.29)$$

para realizar esta prueba usaremos lo siguiente

$$Z_c = \frac{\bar{x} - \mu_w}{\frac{\sigma}{\sqrt{n}}} = \frac{0.0112399 - 0}{\frac{0.0717}{\sqrt{132}}} = 1.8 \quad (4.3.30)$$

donde, Media muestral de $\nabla T(Z_t) : \bar{x} = 0.0112399$

Media poblacional de $\nabla T(Z_t) := \mu_w$

Desviación estándar poblacional: $\sigma = 0.0717$

N: Número de datos=132.

Nuestra regla de decisión es rechazar H_0 si $Z_c > Z_{\frac{\alpha}{2}}$, donde $Z_{\frac{\alpha}{2}} = 1.645$.

Efectivamente, cuando $\alpha = 10\%$ rechazamos la hipótesis nula y aceptamos la hipótesis alternativa, es decir, la media de nuestros datos resulta ser distinta de

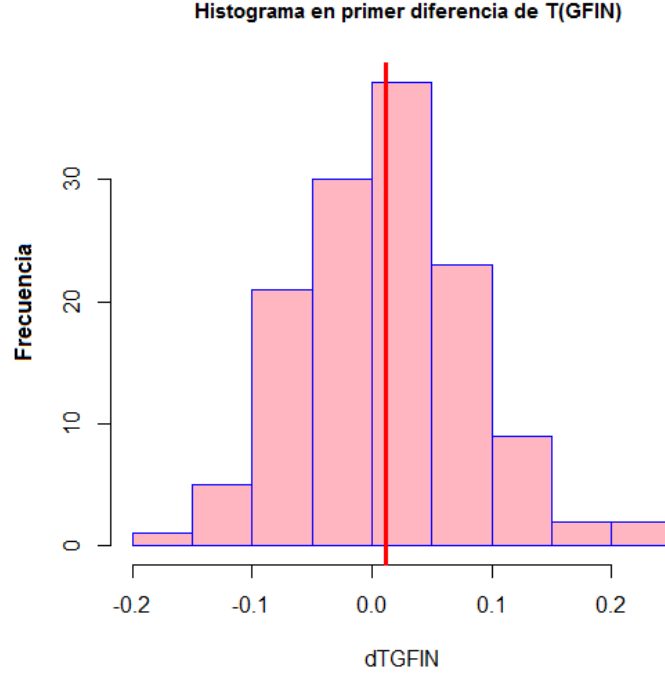


Figura 4.15: Histograma de $\nabla T(GFIN_t)$.

cero, con ello también podemos determinar nuestro intervalo de confianza para la media

$$P\left(\bar{x} - 1.645 \frac{\sigma}{\sqrt{n}} < \mu_w < \bar{x} + 1.645 \frac{\sigma}{\sqrt{n}}\right) \quad (4.3.31)$$

$$\Rightarrow \mu_w \in \left(0.0112 - 1.645 \frac{0.071}{\sqrt{132}}, 0.0112 + 1.645 \frac{0.071}{\sqrt{132}}\right) = (0.000969, 0.0215). \quad (4.3.32)$$

Como podemos observar en las frecuencias de el histograma (Figura 4.15) la media de nuestros datos es diferente de cero y con tendencia positiva (con un 90% de nivel de confianza). Así, el modelo para w_t debería ser

$$w_t = C + a_t. \quad (4.3.33)$$

donde $C = u_w$. Sustituyendo $z_t - z_{t-1}$ para w_t , (4.3.33) demostramos que z_t está cambiando por un monto constante ($C = 0.0112$) para cada periodo, y nuestro pronóstico debería ser

$$\hat{z}_t = 0.0112 + z_{t-1}. \quad (4.3.34)$$

En el Cuadro 4.5 se presenta una comparación entre los valores reales del precio de cierre de las acciones y los valores obtenidos mediante el pronóstico $ARIMA(0, 1, 0)$, como podemos notar, todos los valores se encuentran dentro

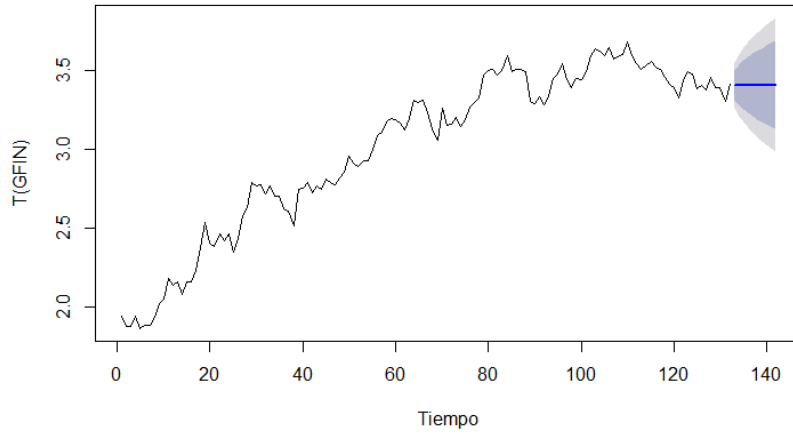


Figura 4.16: Pronóstico de los precios de las acciones de Grupo Financiero Inbursa.

Periodo	Precio Real	Pronósticos	Intervalo Inferior	Intervalo Superior
Enero/2017	30.91	30.41	26.13	34.69
Febrero/2017	28.68	30.66	24.90	36.42
Marzo/2017	31.03	30.91	23.95	37.87
Abril/ 2017	31.79	31.16	23.16	39.16
Mayo/ 2017	31.00	31.41	22.48	40.35
Junio/ 2017	32.17	31.66	21.87	41.46

Cuadro 4.5: Pronósticos del modelo de los precios de las acciones de Grupo Financiero Inbursa.

del intervalo inferior y superior, aun cuando el pronóstico en el mes de Febrero/2017 cuenta con una diferencia de 1.98, porque incluso aunque el valor real fue de \$28 comparado con el pronóstico de \$30, en el Cuadro 4.6 vemos como su error relativo junto con los otros errores de los otros pronósticos son aproximadamente cero.

Los resultados presentados en la Figura 4.16 son referentes a los valores pronosticados mediante el uso del modelo $ARIMA(0, 1, 0)$ junto con el intervalo inferior y superior. En la Figura 4.17, se presentan estos pronósticos en comparación con los datos reales y se aprecia que la estimación de los pronósticos son bastante similares a los valores reales.

Realizando una investigación para conocer el estado de nuestros resultados para

Periodo	Precio Real	Pronósticos	Error absoluto $e_a = P_r - P $	Error relativo $e_r = \frac{e_a}{P_r}$
Enero/2017	30.91	30.41	0.50	0.016
Febrero/2017	28.68	30.66	1.98	0.069
Marzo/2017	31.03	30.91	0.12	0.003
Abril/ 2017	31.79	31.16	0.63	0.019
Mayo/ 2017	31.00	31.41	0.41	0.013
Junio/ 2017	32.17	31.66	0.51	0.015

Cuadro 4.6: Error absoluto y relativo.



Figura 4.17: Precio mensual real de Grupo Financiero Inbursa vs Pronósticos.

compararlo con otras situaciones como lo hicimos con los pronósticos del precio promedio semanal del Crudo de Exportación Mexicano (Cuadro 4.7), podemos observar en los pronósticos de esta investigación que el tamaño de su valor fue bastante pequeño con respecto a los errores relativos de Precios Promedio Semanal. Esto comprueba que nuestros resultados son buenos.

N	Pronóstico PPS	Pronóstico GFIN	Error relativo PPS $e_r = \frac{e_a}{\bar{P}_r}$	Error relativo GFIN $e_r = \frac{e_a}{\bar{P}_r}$
1	37.981	30.41	0.074	0.016
2	37.970	30.66	0.135	0.069
3	37.988	30.91	0.103	0.003
4	38.013	31.16	0.151	0.019
5	38.039	31.41	0.305	0.013
6	38.066	31.66	0.382	0.015

Cuadro 4.7: Error relativo entre el precio promedio semanal de Crudo de Exportación Mexicano vs precio de cierre de una acción mensual del Grupo Financiero Inbursa (López, 2016).

Capítulo 5

Resultados

Inbursa es una empresa dispuesta a manejar dicha modelación para sus decisiones financieras, con el aporte de información proporcionado por el portal de Yahoo Finance se facilitó la obtención de los precios de las acciones mensuales para manejar los datos. Los datos mensuales de los que se dispuso fueron obtenidos de Yahoo Finance, lo que facilitó su uso, porque se encarga de corregir ciertos errores de captura, para así obtener una mejor interpretación.

Primero se realizó un análisis gráfico para observar el comportamiento general de los valores de las acciones, en el análisis se nota un incremento en la mayoría de los datos donde se observa una varianza no constante. De tal forma que se decide realizar un pronóstico mediante un modelo ARIMA, conduciendonos a realizar una estabilización de la varianza, aplicando una transformación logarítmica a la base de datos original.

La FAC estimada en la Figura 4.2 es bastante típica de una serie no estacionaria. Muchas series producen un FAC con una autocorrelación en un rezago cercano a 1 (0.9 ó más grande) seguido por un decaimiento lento. Pero una característica común de una serie con una media no estacionaria es la declinación lenta de las autocorrelaciones estimadas, no su nivel. La primer autocorrelación para una algunas series no estacionarias quizá sea minuciosamente menor que 1, pero las autocorrelaciones restantes se moverán hacia cero muy lentamente.

El FAC residual en la Figura 4.9 es casi similar a la FAC de la primer diferencia en la Figura 4.3. Esto se debe a que las FAC residuales en la Figura 4.9 representan el comportamiento de los datos originales después de que ellos han sido filtrados

El modelo obtenido es una caminata aleatoria, la cual es un buen modelo para muchas series de precios de las acciones. Este resultado empírico es consistente con la forma débil de la hipótesis del mercado eficiente.(Alan Pankratz, 1983). Antes de considerarlo se determina si el modelo ARIMA(0, 1, 0) cuenta con al-

guna tendencia determinística y quedó aprobada gracias a la prueba de hipótesis de una media distinta de cero de $\nabla T(GFIN)$ también conocida como *drift*. La prueba de hipótesis nos muestra que la caminata aleatoria cuenta con una media distinta de cero, la cual nos indica la tendencia de los precios de las acciones de INBURSA, precisando su valor con el histograma de frecuencia y la media de $\nabla T(Z_t)$.

No todos los modelos ARIMA con estimaciones significativas son sensatos. (Alan Pankratz, 1983). El modelo ARIMA(0, 1, 1) es un ejemplo que se ajusta a una realización corta, pero que no es aceptable por otros motivos. Nos esforzamos por obtener modelos parsimoniosos que se ajusten a los datos adecuadamente con coeficientes significativos, pero también debemos ajustar nuestros resultados estadísticos a la naturaleza de los datos estudiados.

Tomando en cuenta los resultados anteriores, con base al caso de estudio presentado por (Alan Pankratz, 1983), se concluye que la serie de tiempo de precios de INBURSA será modelada mediante el proceso ARIMA(0, 1, 0).

Por tanto, podemos comprender como una caminata aleatoria es el mejor modelo para muchas series de precios de las acciones, a veces el mejor modelo puede ser el más sencillo pero capaz de explicar el comportamiento de ciertas series de tiempo financieras. En este sentido se presentan los modelos ARIMA como una alternativa de solución para las empresas, porque tal modelación es efectiva en cuanto a calidad y tiempo, ya que las empresas requieren de modelos que cumplan ambas características por razones competitivas y financieras.

5.0.1. Conclusión

En el análisis de una serie de tiempo financiera como Inbursa, se vuelve muy útil para la toma de decisiones como accionista, debido a los movimientos frecuentes del mercado accionario. Los accionistas o inversionistas podrían considerar a Grupo Financiero Inbursa como una empresa de inversión a largo plazo, gracias a su lenta tendencia creciente.

5.0.2. Trabajos futuros

No quedando conformes con los resultados de un solo método, como trabajo futuro es recomendable aplicar los modelos de regresión ARCH y GARCH, dichos modelos son recomendados para este tipo de series de tiempo, en particular para las series de tiempo financieras por su alta volatilidad, ya que estos modelos consideran la información pasada de la variable y su volatilidad (o varianza) como factor altamente explicativo de su comportamiento presente.

Apéndice A

Código en R

Cargamos los datos de Inbursa del año 2006 a 2016.

```
data = read.csv("C:BUAPINBURSA2006.csv", header = TRUE)
```

```
data
```

```
attach(data)
```

```
tsdata <- ts(data, start = c(2006, 1), end = c(2016, 12), frequency = 12)
```

Realizamos la gráfica de los datos para conocer su comportamiento.

```
plot(tsdata, type = "l", xlab = "Tiempo", ylab = "GRUPOFINANCIEROINBURSA2006-2016")
```

En la etapa de identificación es donde realizar la estabilización de la varianza , así que procedemos con lo siguiente:

```
#E1:IDENTIFICACION
```

```
##Estabilizacion de la varianza
```

Paso1: Obtener la media de cada grupo (cargamos la serie pero agrupada)

```
data = read.csv("C : BUAPINBURSA2006lamda.csv", header = TRUE)
```

```
attach(data)
```

```
mean(2006) #7.204726, grupo 1
```

```
mean(X2007) #10.30205
```

```
mean(X2008) #14.34309
```

```
mean(X2009) #15.53116
```

```
mean(X2010) #20.65466
```

```
mean(X2011) # 24.65631
```

```
mean(X2012) #29.50051
```

```
mean(X2013) #30.68015
```

```
mean(X2014) #34.61281
```

```
mean(X2015) #34.10653
```

```
mean(X2016) #30.17741 , grupo 11
```

Paso2: Obtener la desviación estándar

```
sd(i..2006) #0.8143781, grupo 1
```

```

sd(X2007) #1.53612
sd(X2008) #1.876993
sd(X2009) #1.282613
sd(X2010) #2.601149
sd(X2011) #2.054872
sd(X2012) #4.403305
sd(X2013) #3.038291
sd(X2014) #3.061518
sd(X2015) #2.762921
sd(X2016) #1.704815 , grupo 11

```

Paso3: Relación desv entre media a la 1-lamda

Grupo(1) Año:2006

```

Ln11 <- -sd(2006)/(mean(2006))^2 #valor1(12) lamda=-1 #0.01568885
Ln51 <- -sd(2006)/(mean(2006))^1.5 #valor1(12) lamda=-0.5 #0.04211142
Ln01 <- -sd(2006)/(mean(2006)) #valor1(12) lamda=0 es #0.1130339
Lp51 <- -sd(2006)/(mean(2006))^0.5 #valor1(12) lamda=0.5 #0.3034012
Lp11 <- -sd(2006) #valor1(12) lamda=1 #0.8143781

```

Grupo(2) Año:2007

```

Ln12 <- -sd(X2007)/(mean(X2007))^2 #valor2(12) lamda=-1 #0.01447364
Ln52 <- -sd(X2007)/(mean(X2007))^1.5 #valor2(12) lamda=-0.5 #0.04645577
Ln02 <- -sd(X2007)/(mean(X2007)) #valor2(12) lamda=0 es #0.1491082
Lp52 <- -sd(X2007)/(mean(X2007))^0.5 #valor2(12) lamda=0.5 #0.4785898
Lp12 <- -sd(X2007) #valor2(12) lamda=1 # 1.53612

```

Grupo(3) Año:2008

```

Ln13 <- -sd(X2008)/(mean(X2008))^2 #valor3(12) lamda=-1 #0.009123828
Ln53 <- -sd(X2008)/(mean(X2008))^1.5 #valor3(12) lamda=-0.5 #0.03455401
Ln03 <- -sd(X2008)/(mean(X2008)) #valor3(12) lamda=0 es # 0.1308639
Lp53 <- -sd(X2008)/(mean(X2008))^0.5 #valor3(12) lamda=0.5 #0.4956114
Lp13 <- -sd(X2008) #valor3(12) lamda=1 # 1.876993

```

Grupo(4) Año:2009

```

Ln14 <- -sd(X2009)/(mean(X2009))^2 #valor4(12) lamda=-1 #0.005317264
Ln54 <- -sd(X2009)/(mean(X2009))^1.5 #valor4(12) lamda=-0.5 #0.02095512
Ln04 <- -sd(X2009)/(mean(X2009)) #valor4(12) lamda=0 es #0.08258326
Lp54 <- -sd(X2009)/(mean(X2009))^0.5 #valor4(12) lamda=0.5 #0.3254572
Lp14 <- -sd(X2009) #valor4(12) lamda=1 #1.282613

```

Grupo(5) Año:2010

```

Ln15 <- -sd(X2010)/(mean(X2010))^2 #valor5(12) lamda=-1 #0.00609718
Ln55 <- -sd(X2010)/(mean(X2010))^1.5 #valor5(12) lamda=-0.5 #0.0277101
Ln05 <- -sd(X2010)/(mean(X2010)) #valor5(12) lamda=0 es #0.1259352
Lp55 <- -sd(X2010)/(mean(X2010))^0.5 #valor5(12) lamda=0.5 #0.5723428

```

$Lp15 < -sd(X2010)$ #valor5(12) lamda=1 #2.601149

Grupo(6) Año:2011

$Ln16 < -sd(X2011)/(mean(X2011))^2$ #valor6(12) lamda=-1 #0.003380092

$Ln56 < -sd(X2011)/(mean(X2011))^{1.5}$ #valor6(12) lamda=-0.5 #0.01678389

$Ln06 < -sd(X2011)/(mean(X2011))$ #valor6(12) lamda=0 es #0.08334061

$Lp56 < -sd(X2011)/(mean(X2011))^{0.5}$ #valor6(12) lamda=0.5 #0.4138288

$Lp16 < -sd(X2011)$ #valor6(12) lamda=1 #2.054872

Grupo(7) Año:2012

$Ln17 < -sd(X2012)/(mean(X2012))^2$ #valor7(12) lamda=-1 #0.005059643

$Ln57 < -sd(X2012)/(mean(X2012))^{1.5}$ #valor7(12) lamda=-0.5 #0.02748113

$Ln07 < -sd(X2012)/(mean(X2012))$ #valor7(12) lamda=0 es #0.149262

$Lp57 < -sd(X2012)/(mean(X2012))^{0.5}$ #valor7(12) lamda=0.5 #0.81070738

$Lp17 < -sd(X2012)$ #valor7(12) lamda=1 #4.403305

Grupo(8) Año:2013

$Ln18 < -sd(X2013)/(mean(X2013))^2$ #valor8(12) lamda=-1 #0.003227857

$Ln58 < -sd(X2013)/(mean(X2013))^{1.5}$ #valor8(12) lamda=-0.5 #0.01787899

$Ln08 < -sd(X2013)/(mean(X2013))$ #valor8(12) lamda=0 es #0.09903116

$Lp58 < -sd(X2013)/(mean(X2013))^{0.5}$ #valor8(12) lamda=0.5 #0.54853038

$Lp18 < -sd(X2013)$ #valor8(12) lamda=1 #3.038291

Grupo(9) Año:2014

$Ln19 < -sd(X2014)/(mean(X2014))^2$ #valor9(12) lamda=-1 #0.002555425

$Ln59 < -sd(X2014)/(mean(X2014))^{1.5}$ #valor9(12) lamda=-0.5 #0.01503424

$Ln09 < -sd(X2014)/(mean(X2014))$ #valor9(12) lamda=0 es #0.08845043

$Lp59 < -sd(X2014)/(mean(X2014))^{0.5}$ #valor9(12) lamda=0.5 #0.52037738

$Lp19 < -sd(X2014)$ #valor9(12) lamda=1 #3.061518

Grupo(10) Año:2015

$Ln110 < -sd(X2015)/(mean(X2015))^2$ #valor10(12) lamda=-1

$Ln510 < -sd(X2015)/(mean(X2015))^{1.5}$ #valor10(12) lamda=-0.5

$Ln010 < -sd(X2015)/(mean(X2015))$ #valor10(12) lamda=0 es

$Lp510 < -sd(X2015)/(mean(X2015))^{0.5}$ #valor10(12) lamda=0.5

$Lp110 < -sd(X2015)$ #valor10(12) lamda=1

Grupo(11) Año:2016

$Ln111 < -sd(X2016)/(mean(X2016))^2$ #valor11(12) lamda=-1

$Ln511 < -sd(X2016)/(mean(X2016))^{1.5}$ #valor11(12) lamda=-0.5

$Ln011 < -sd(X2016)/(mean(X2016))$ #valor11(12) lamda=0 es

$Lp511 < -sd(X2016)/(mean(X2016))^{0.5}$ #valor11(12) lamda=0.5

$Lp111 < -sd(X2016)$ #valor11(12) lamda=1

Paso4: Obtener el coef de variancion

```

M(lamda)
mn1<-sum(c(Ln11,Ln12,Ln13,Ln14,Ln15,Ln16,Ln17,Ln18,Ln19,Ln110,Ln111))/11
mn5<-sum(c(Ln51,Ln52,Ln53,Ln54,Ln55,Ln56,Ln57,Ln58,Ln59,Ln510,Ln511))/11
mn0<-sum(c(Ln01,Ln02,Ln03,Ln04,Ln05,Ln06,Ln07,Ln08,Ln09,Ln010,Ln011))/11
mp5<-sum(c(Lp51,Lp52,Lp53,Lp54,Lp55,Lp56,Lp57,Lp58,Lp59,Lp510,Lp511))/11
mp1<-sum(c(Lp11,Lp12,Lp13,Lp14,Lp15,Lp16,Lp17,Lp18,Lp19,Lp110,Lp111))/11

de(lamda)
Cn1<-c((Ln11-mn1)^2,(Ln12-mn1)^2,(Ln13-mn1)^2,(Ln14-mn1)^2,(Ln15-mn1)^2,(Ln16-
mn1)^2,(Ln17-mn1)^2,(Ln18-mn1)^2,(Ln19-mn1)^2,(Ln110-mn1)^2,(Ln111-mn1)^2)
den1<-sqrt(sum(Cn1)/10)
Cn5<-c((Ln51-mn5)^2,(Ln52-mn5)^2,(Ln53-mn5)^2,(Ln54-mn5)^2,(Ln55-mn5)^2,(Ln56-
mn5)^2,(Ln57-mn5)^2,(Ln58-mn5)^2,(Ln59-mn5)^2,(Ln510-mn5)^2,(Ln511-mn5)^2)
den5<-sqrt(sum(Cn5)/10)
Cn0<-c((Ln01-mn0)^2,(Ln02-mn0)^2,(Ln03-mn0)^2,(Ln04-mn0)^2,(Ln05-mn0)^2,(Ln06-
mn0)^2,(Ln07-mn0)^2,(Ln08-mn0)^2,(Ln09-mn0)^2,(Ln010-mn0)^2,(Ln011-mn0)^2)
den0<-sqrt(sum(Cn0)/10)
Cp5<-c((Lp51-mp5)^2,(Lp52-mp5)^2,(Lp53-mp5)^2,(Lp54-mp5)^2,(Lp55-mp5)^2,(Lp56-
mp5)^2,(Lp57-mp5)^2,(Lp58-mp5)^2,(Lp59-mp5)^2,(Lp510-mp5)^2,(Lp511-mp5)^2)
dep5<-sqrt(sum(Cp5)/10)
Cp1<-c((Lp11-mp1)^2,(Lp12-mp1)^2,(Lp13-mp1)^2,(Lp14-mp1)^2,(Lp15-mp1)^2,(Lp16-
mp1)^2,(Lp17-mp1)^2,(Lp18-mp1)^2,(Lp19-mp1)^2,(Lp110-mp1)^2,(Lp111-mp1)^2)
dep1<-sqrt(sum(Cp1)/10)

C V ( lamda )
CVn1<-den1/mn1 # 0.7667536
CVn5<-den5/mn5 #0.4819464
CVn0<-den0/mn0 #0.2891669
CVp5<-dep5/mp5 #0.3054266
CVp1<-dep1/mp1 #0.4434787

Paso5: Obtener el mínimo coeficiente de variación : CV(lamda)

x<-c(CVn1,CVn5,CVn0,CVp5,CVp1)
min(x) #Min[CV(lamda=0)]=0.2891669 #T(GFINt)=log{GFINt}
A continuación también presentamos como se realizó la estabilización del nivel.

##Estabilización del nivel
Cargamos las librerías library(tseries)
library(forecast) #sirve para hacer proyecciones y auto.arima
cargamos nuestra base de datos pero ya con la transformación de varianza lla-
mada TGFIN.
#Fac muestral
TGFIN=log(i.INBURSA) #indica si serie es estacionaria
dTGFIN=diff(TGFIN) #en cuanto a nivel de la serie
ddTGFIN=diff(dTGFIN)

```

```
dddTGFIN=diff(ddTGFIN)
```

Conocemos cual será el comportamiento y valores de nuestras FAC y FACP con 0,1, o 2 veces diferenciada nuestra TGFIN.

```
acf(TGFIN)
acf(dTGFIN) #FAC muestral de {T(Xt)} acf(ddTGFIN) #FAC muestral de
{diffT^2(Xt)}
acf(dddTGFIN)
#Fac parcial
pacf(TGFIN) #para conocer el orden del proceso y parametros
pacf(dTGFIN) #FAC muestral de {T(Xt)}
pacf(ddTGFIN) #FAC muestral de {diffT^2(Xt)}
pacf(dddTGFIN)
sd(TGFIN) # 0.5173365
```

Aquí vamos a conocer cual es nuestro mínimo valor de nuestra desviación estándar dependiendo de las veces en que nuestra serie fue diferenciada.

```
sd(dTGFIN) # 0.07173448
sd(ddTGFIN) #0.1057995
sd(dddTGFIN) #0.1858733
y<-c(sd(TGFIN),sd(dTGFIN),sd(ddTGFIN),sd(dddTGFIN))
min(y) #sd(1° Diferencia TGFIN)= 0.07173448
```

#Empleo de la función de autocorrelacion muestral y autocorrelación muestral parcial.

Realizamos nuestra prueba de significancia.

```
r1k=acf(dTGFIN)
supuesto rhok=0 para k>0 i.e., serie es ruido blanco
var0k<-(1+2*0)/131 Nota:Se resta N-n° diferencias aplicadas
sdr0k<-sqrt((1+2*0)/131) #0.08737041 observamos que | r1k| >raiz(var0k)
```

```
supuesto rhok><0 para k>0
rho=0 para k>1 #desviación estándar estimada de rk, para k=2,3,...
```

```
var1k<-(1+2*(-0.091)^2)/131
sdr1k<-sqrt(var1k)
sdr1k #0.08809095 sign ><0 si |r1k|>2*raiz(var0k)
# | r1k | = 0.091<0.1761819
```

```
rho=0 para k>2
c<-c(0.091^2, 0.034^2)
var2k<-(1+2*sum(c))/131
sdr2k<-sqrt(var2k)
sdr2k #0.08819107 |r2k|=0.034<0.1763821
```

```

rho=0 para k>3
c<-c(0.091^2 ,0.034^2, .132)
var3k<-(1+2*sum(c))/131
sdr3k<-sqrt(var3k)
sdr3k      #0.09895924   | r3k |=0.132>0.1979185

```

```
p1k=pacf(dTGFIN)
```

```

Varphi=1/(132-1)  #para i>p=1
sdphi=sqrt(Varphi)
sdphi      #00.08737041 | phik|>2sdphi
#NO |-0.091|>0.1747408

```

Conocemos cuales serán los valores AIC y BIC de los resultados obtenidos anteriormente con ayuda de ciertos paquetes.

```

library(rlang)
forecast::auto.arima(TGFIN)
#R propone en autoarima ARIMA(0,1,0) con AIC=-315.56 y BIC=-309.8

```

Código para obtener el modelo propuesto junto con sus respectivos criterios y residuales. MODELO AUTOARIMA

```

x.auto<-arima(TGFIN,order=c(0,1,0))
x.res<-x.auto$residuals

```


Apéndice B

Prueba Chi-Cuadrada

Una forma de lidiar con los valores de FAC's residuales subestimados es probar las autocorrelaciones residuales como un conjunto en lugar de individualmente.

Un estadístico chi-cuadrado aproximada (el estadístico Ljung-Box) está disponible para esta prueba. Si este estadístico es significativo, deberíamos considerar la reformulación de modelo.

Probamos la siguiente hipótesis nula conjunta sobre las correlaciones entre los choque aleatorios.

$$H_0 : \rho_1(a) = \rho_2(a) = \dots = \rho_k(a) = 0 \quad (\text{B.0.1})$$

con este estadístico de prueba.

$$Q^* = n(n+2) \sum_{k=1}^K (n-k)^{-1} r_k^2(\hat{a}). \quad (\text{B.0.2})$$

donde n es el número de observaciones usadas para estimar el modelo. El estadístico Q^* aproximadamente sigue una distribución chi-cuadrada con $(K-m)$ grados de libertad, donde m es el número de parámetros estimados en el modelo ARIMA.

Si Q^* es grande (significativamente diferente de cero), dice que las autocorrelaciones residuales como un conjunto son significativamente diferentes de cero, y los choques aleatorios de modelo estimado probablemente estén autocorrelacionados. Entonces deberíamos reconsiderar la reformulación del modelo.

Apéndice C

Estadístico de Box-Pierce y Ljung-Box

Una buena manera de comprobar la suficiencia del modelo Box-Jenkins, es analizar los residuos obtenidos a partir del modelo, de la manera siguiente. Del mismo modo en que se calcularon las funciones de autocorrelación muestral, y autocorrelación parcial muestral de la serie de tiempo, se pueden determinar dichas funciones para los residuos.

Una forma de aprovechar los residuos para comprobar la suficiencia del modelo, es examinar una estadística o dato numérico que determine si las primeras K autocorrelaciones muestrales de los residuos, considerados juntos, indican suficiencia del modelo.

Box y Pierce desarrollaron un estadístico que, basado en los cuadrados de los primeros coeficientes e autocorrelación de los residuos, permite analizar si existe o no autocorrelación. Estos datos numéricos se resumen a continuación.

1. La estadística de Box-Pierce es

$$Q = n' \sum_{l=1}^K \hat{\rho}_l^2 \quad (\text{C.0.1})$$

siendo

$$\hat{\rho}_l = \frac{\sum_{t=l+1}^{n'} e_t e_{t-1}}{\sum_{t=1}^{n'} e_t^2} \quad (\text{C.0.2})$$

2. El estadístico de Ljung-Box es

$$Q^* = n'(n' + 2) \sum_{l=1}^K \frac{\hat{\rho}_l^2}{n - 1} \quad (\text{C.0.3})$$

Aquí, $n' = n - d$, donde n es el número de observaciones en la serie de tiempo original, y d es el grado de diferenciación no estacional usado para transformar los valores originales de la serie de tiempo en los valores estacionarios de la serie de tiempo. Además, $\hat{\rho}_l^2$ es el cuadrado de $\hat{\rho}_l$, la autocorrelación muestral de los residuos en el *lag* l , es decir, la autocorrelación muestral de los residuos separada por un desfase de unidades de tiempo l .

Apéndice D

Criterio de información de Akaike

Cuando se requiere que una serie $\{X_t\}$ sea ajustada por un modelo ARMA con media cero, se encuentra con el problema de encontrar el orden de p y q así como el encontrar su valor. Una aproximación alternativa para la selección de este modelo es llevar a cabo una serie de pruebas de hipótesis o usar un criterio de selección de modelos.

Existe un estadístico de ajuste que puede ser de ayuda para la selección de modelos. Un criterio muy utilizado es el de información de Akaike (AIC), está dado por

$$AIC = -2\ln(\text{máxima verosimilitud}) + 2r \quad (\text{D.0.1})$$

donde r denota el número de parámetros independientes que son ajustados por el modelo que está siendo evaluado. El AIC esencialmente escoge el modelo con el mejor ajuste, moderado por la función de verosimilitud, sometido a un término *penalty* para prevenir el sobreajuste, que incrementa con el número de parámetros en el modelo. Para un modelo $ARMA(p, d, q)$, note que $r = p + q + 1$ es la varianza residual que es incluida como un parámetro. Ignorando constantes arbitrarias, el primer término de verosimilitud es usualmente aproximado por $N\ln(\frac{S}{N})$, donde S denota la suma de cuadrados residuales, y N es el número de observaciones.

Apéndice E

Criterio de información Bayesiano

El criterio de información Bayesiana propuesto por Schwarz (1978). Es una medida de bondad de ajuste de un modelo estadístico, y es a menudo utilizado como un criterio para la selección de modelos entre un conjunto finito de modelos. Se basa en la función de probabilidad logarítmica y está estrechamente relacionado con el criterio de información de Akaike.

Al igual que en el AIC, BIC introduce un término de penalización para el número de parámetros en el modelo.

En la práctica los valores de

$$f_i(x_n) = \int_{\Theta_i} f_i(x_n|\theta)\pi_i(\theta)d\theta \quad (\text{E.0.1})$$

son difíciles de calcular, además de que para ello se requiere de la especificación de las densidades a priori $\pi_i(\theta)$. Una aproximación del logaritmo de (E.0.1) está dada por

$$\log[f_i(x_n)] \approx \log[f_i(x_n|\hat{\theta}_n^i)] - \frac{k_i}{2}\log(n) \quad (\text{E.0.2})$$

donde $\hat{\theta}_n^i$ es el EMV de θ^i y $\log[f_i(x_n|\hat{\theta}_n^i)]$ es la log-verosimilitud correspondiente al modelo M_i .

Así, ya que la función logaritmo es monótona creciente, seleccionar como mejor modelo el que maximice (E.0.1) equivale aproximadamente a seleccionar al que maximice (E.0.2), lo que a su vez equivale a seleccionar el modelo que minimice

$$BIC =: k_i \ln(n) - 2\ln(L) \quad (\text{E.0.3})$$

donde, k es el número de parámetros del modelo y $\ln(L) = \ln(\hat{\theta}_n^i)$ es la función de log-verosimilitud para el modelo estadístico.

Un BIC bajo implica un número menor de variables explicativas, mejor ajuste, o ambos.

Apéndice F

FAC y FACP para diferentes modelos

Modelo	FAC	FACP
Promedio móvil de primer orden $X_t = a_t + \theta a_{t-1}$.	Se corta después del lapso $k = 1$ $\frac{-\theta}{1+\theta^2}$, si $k = 1$ 0, si $k \geq 2$.	Se corta de manera que domina el decremento exponencial amortiguado. Para lag k es $\frac{-(-\theta)^k}{1+\theta^2+\dots+\theta^{2k}}$,
Promedio móvil de segundo orden $X_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}$.	Se corta después del lapso $k = 2$ $\frac{-\theta_1 + \theta_1 \theta_2}{1+\theta_1^2 + \theta_2^2}$, si $k = 1$ $\frac{-\theta_2}{1+\theta_1^2 + \theta_2^2}$, si $k = 2$ 0, si $k \geq 3$.	Se corta en forma de exponenciales amortiguadas, ondas en forma de seno amortiguada o ambos.
Promedio móvil de orden q $X_t = a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}$.	Se trunca después del lapso q 1, si $k = 0$ $\frac{-\theta_k + \theta_1 \theta_{k-1} + \dots + \theta_q \theta_{k-q}}{1+\theta_1^2 + \dots + \theta_q^2}$, si $k = 1, 2, \dots, q$ 0, si $k > q$.	Se extingue. Resulta una mezcla de decaimientos exponenciales y dependiendo de los signos $\Theta(B) = 0$.
Autorregresivo de primer orden $X_t = \phi_1 X_{t-1} + a_t$.	Se corta en forma de exponenciales amortiguadas. $\rho_k = \phi^k$, $k = 0, \pm 1, \dots$	Se corta después del lapso $k = 1$ $\phi_{1,1} = \rho_1$ $\phi_{k,k} = 0$, para $k > 1$.
Autorregresivo de segundo orden $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + a_t$.	Se corta en forma de exponenciales amortiguadas, ondas en forma de seno amortiguada o ambos. $\phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}$, $k = 1, 2, 3, \dots$	Se corta después del lapso $k = 2$. $\phi_{1,1} = \rho_1$ $\phi_{2,2} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$ $\phi_{3,3} = 0$, $k \geq 3$.
Autorregresivo de orden p $X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + a_t$.	Se extingue $\sum_{i=1}^p \phi_i \rho_{k-i}$, $k = 1, 2, \dots$	Se trunca después del lapso $k = p$ $\phi_{k,k} = \phi_k$, si $k = p$ $\phi_{k,k} = 0$, si $k > p$.
Autorregresivo de Promedios Móviles de orden $(1, 1)$ $X_t - \phi_1 X_{t-1} = \theta a_{t-1} + a_t$.	Tiene un decaimiento exponencial a cero a partir de p . $\frac{\phi^{k-1}(1-\phi\theta)(\phi-\theta)}{1-2\phi\theta+\theta^2}$, $k = 1, 2, \dots$	Se corta de manera que domina el decremento exponencial amortiguado $\phi_{1,1} = \rho_1$.
Autorregresivo de Promedios Móviles de orden (p, q) $X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} + a_t$.	Se extingue $\phi_1 \rho_{k-1} + \dots + \phi_p \rho_{k-p}$, para $k > q$ para $k < q$, ρ_k involucra a los parámetros $\theta_k, \theta_{k+1}, \dots, \theta_q$	Se extingue. Mezcla de decaimientos exponenciales u ondas sinusoidales amortiguadas dependiendo de las raíces de $\Theta(B)$ y $\Phi(B)$.

Cuadro F.1: Características distintivas de FAC y FACP para procesos estacionarios (Brockwell, 1997) y (Guerrero, 2009).

Bibliografía

- [1] Blanco, José: Recesión en México, recuperado el 5 de enero de 2019, de La Jornada.
- [2] Box, G. E.P., Jenkins, G., Reinsel, G. (2008). Time Series Analysis. Forecasting and Control. Quinta edición. Editorial Wiley.
- [3] Box, G. E. P. y G.M. Jenkins. (1976). Time series analysis: forecasting and control. Segunda edición. San Francisco: Holden-Day.
- [4] Bowerman, Bruce L, Richard T. O'Connell y Anne B. Koehler. (2007). Forecasting, time series and regression. Cuarta edición. Editorial Cengage Learning.
- [5] Brockwell, P. J. & Davis, R. A. (2002). Introduction to Time Series and Forecasting. Segunda edición. Editorial Springer-Verlag, p. 29-32, 34, 54.
- [6] Choi K, Thacker S (1981). An evaluation of influenza mortality surveillance 1962-1979: I Time Series Forecasts of expected pneumonia and influenza deaths. Am J Epidemiol, p. 113 y 215-26.
- [7] Douglas C. Montgomery(2002). Introducción al Análisis de regresión lineal.
- [8] Durbin, J., and G.S. Watson, (1950). Testing for serial correlation in least squares regression, I. Biometrika 37, p. 409-428.
- [9] Durbin, J., and G.S. Watson, (1950). Testing for serial correlation in least squares regression, II. Biometrika 37, p. 159-179.
- [10] Guerrero Guzmán, Victor, (2009). Análisis Estadístico y pronóstico de series de tiempo económicas. Tercera edición. Editorial Jit Press, p.13, 14, 107-120, 131-164,
- [11] Guerrero V., (1987). Modelos Estadísticos para series de tiempo

univariadas. V Coloquio del departamento de matemáticas. Centro de Investigación y Estudios Avanzados del IPN.

[12] Gujarati, D. N., y Porter, D.C. (2009). *Econometría* (Quinta edición ed.). México, D.F: McGRAW HILL, p.469.

[13] Mota López, Arantxa & Víctor Hugo Vázquez Guevara, (2016). *Pronóstico del precio del crudo de exportación mexicano con la metodología de Box–Jenkins para series de tiempo* (Tesis de Licenciatura). Benemérita Universidad Autónoma de Puebla, Puebla.

[14] Navidi, W.C.(2006). *Estadística para ingenieros y científicos*. Segunda edición. McGrall-Hill, p. 368 y 369.

[15] Pankratz,Alan. (1983). *Forecasting with Univariate Box-Jenkins Models. Concept and Cases*. Primera edición. Editorial John Wiley & Sons. Inc.DePaw University, en 81-83, 95-100, 103, 130, 149, 167, 193-197, 250, 256, 402-410 y 429.

[16] Parody, E., Charris, A & García, R. (2016). *Modelo Log-normal para la Predicción del Precio de las Acciones del Sector Bancario que Cotizan en el Índice General de la Bolsa de Valores de Colombia (IGBC)*. Dimensión Empresarial, p. 14, 137 -149.

Referencias bibliográficas en línea.

[1] Bancos mexicanos no se animan a la globalización. En Milenio. Sección: negocios. Recuperado el 1 de Enero del 2018 de www.milenio.com.

[2] Base de datos. En Yahoo Finance. Sección : Precio cerrado de las acciones GFINBURO.MX. Recuperado el 12 de Diciembre del 2017 de finance.yahoo.com

[3] El sexenio de Calderón y la economía. Revista Entrepreneur. Sección: Consultoría. Recuperado el 19 de Octubre de 2018 de www.entrepreneur.com

[4] Se duplica la deuda externa de México. Diario La razón Sección: Negocios. Recuperado el 19 de Octubre de 2018 de : www.razon.com.mx.