

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

COORDINACION DE ESTUDIOS DE POSGRADO

INSTITUTO DE INVESTIGACIONES EN MATEMATICAS APLICADAS Y EN SISTEMAS

ANÁLISIS DE UN ÍNDICE BURSÁTIL CON MODELOS NO ESTACIONARIOS

TESINA PARA OBTENER EL TÍTULO DE ESPECIALISTA EN ESTADÍSTICA APLICADA

PRESENTA

GUSTAVO EDUARDO PESCADOR JIMÉNEZ

12 DE AGOSTO DE 2013



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A mi madre por su continuo apoyo, enseñanzas y cariño.

Agradecimientos

Le agradezco al Dr. Fernando Baltazar por su tiempo y apoyo para terminar a tiempo este trabajo.

Al Dr. Alberto Contreras y Act. Miguel Angel Chong por sus observaciones y aportación de su conocimiento en series de tiempo.

Al IIMAS y la UNAM por darme las bases necesarias en la práctica y teoría.

¡Muchas gracias a todos!

Índice general

Introducción	III
1. Series de Tiempo en Finanzas y sus Características	1
1.1. Rendimiento de los Activos	1
1.2. Otros procesos en consideración	3
2. Modelos Lineales y Aplicación	7
2.1. Propiedades y Definiciones	7
2.1.1. Procesos Estocásticos	8
2.1.2. Estacionariedad	8
2.1.3. Estimación y eliminación de tendencia y componen- te estacional	10
2.1.4. Procesos lineales	13
2.2. Modelo Autorregresivo (AR)	14
2.2.1. Propiedades	15
2.2.2. Identificar un AR	18
2.3. Modelo de Promedios Móviles (MA)	23
2.3.1. Propiedades	23
2.3.2. Identificar un MA	24
2.4. Modelos ARMA	27
2.4.1. Propiedades de un ARMA(1,1)	27
3. Modelos no Lineales y Aplicación	31
3.1. ARCH	33
3.1.1. Propiedades	34
3.1.2. Desventajas	35
3.1.3. Efecto ARCH	35
3.2. GARCH	36

3.2.1. Propiedades	36
3.3. Contrastes de Estructura y Bondad de Ajuste	42
4. Aplicación	45
Conclusiones	63
Bibliografía	63

Introducción

Datos relacionados con el mercado financiero o el área de economía tienen en común que se cuentan con registros periódicos, por ejemplo, las tasas de interés semanales, mensuales y anuales, el precio diario de acciones de apertura o cierre, los precios mensuales de índices bursátiles, etc. Esta característica no es propiamente de estas áreas, también podemos encontrarlos en el área de meteorología, agricultura, geología, etc. Este tipo de cifras se pueden modelar con series de tiempo. El principal propósito de las series de tiempo es explicar, describir y modelar el comportamiento que tienen estos datos así como predecir valores futuros a partir de la historia de la serie, es decir, datos del presente y pasado.

Usualmente en finanzas la varianza condicional de los rendimientos es considerada como un medida de riesgo para el activo en estudio. Ésta es una parte importante para la valuación de un activo financiero y el cálculo del valor en riesgo (VaR por sus siglas en inglés).

Problemas recientes en el área de finanzas han motivado al estudio de la variabilidad, comúnmente llamada volatilidad. Un problema muy común en el área económica financiera es el de predecir la volatilidad. Además de los modelos como los autorregresivos de media móvil (ARMA) que suponen una varianza constante, existen otros modelos que se desarrollaron para el manejo de este problema debido a que estos representan a la volatilidad de mejor manera. Otro tipo de modelos menos conocido pero de gran eficiencia son los modelos como los autoregresivos con heteroscedasticidad condicional (ARCH), presentados por Engle (1982)¹, la generalización de estos (GARCH, modelo que será de interés para este trabajo), por Bollerslev (1986)² y otras generalizaciones como

¹Engle, R.F. *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kindom inflations*. *Econometrica* 50:987-1007.

²Bollerslev, T. *Generalized autoregressive conditional heteroskedasticity*. *Journal of Econo-*

IGARCH, EGARCH, TARCH, entre otros. Recientemente modelos de volatilidad estocástica han sido de mucho interés por ser modelos de espacio-estado, estos modelos son similares a los ARCH pero con la diferencia de que tienen una variable de ruido estocástico, véase [10].

Usaremos modelos de heteroscedasticidad condicional y no los llamados modelos de volatilidad estocástica (véase [10] capítulo 6) pues una de las razones es que los modelos ARCH tienen un proceso generador de datos diferente que los hace más adecuados para ciertos objetivos, como el agrupamiento de la volatilidad, estacionariedad, etc. Otra razón es que, si bien estos modelos pueden aplicarse a diversos problemas estadísticos de series temporales, estos tienen una mejor aplicación a series financieras, por ejemplo a la relación que se tiene entre riesgo y rendimiento. Además, algunas características que tienen los rendimientos, como la presencia de valores atípicos, es que son impredecibles, y una característica muy peculiar es que tanto los periodos de más agitación como lo más tranquilos están agrupados en el tiempo. Estas mismas características son las que se modelan con modelos ARCH. Cuando la volatilidad es alta (baja), lo más seguro es que permanezca elevada (baja), además estos periodos son limitados y, bajo ciertas características, la predicción de la volatilidad tenderá a la varianza de los rendimientos, es decir, será menos extrema, véase [5]. Un proceso ARCH produce un mayor número de valores extremos de lo que se esperaría de una distribución normal estándar, ya que los valores extremos durante el periodo de alta volatilidad son mayores de los que se habrían podido anticipar con un proceso de volatilidad constante.

El índice de volatilidad de un mercado financiero ha empezado a ser un instrumento financiero. El índice VIX de volatilidad se clasificó por el Consejo de Cambio de Opciones de Chicago (CBOE por sus siglas en inglés) que empezó el intercambio de futuros en Marzo del 2004.

Los modelos con heteroscedasticidad condicional pueden ser clasificados en dos categorías. Aquellos que son de la primera, usan una función exacta para la evolución de la volatilidad, mientras que los de la segunda usan una ecuación estocástica para describir el comportamiento de la volatilidad. Los modelos GARCH pertenecen a la primera categoría y los modelos de volatilidad estocástica pertenecen a la segunda.

En este trabajo nos enfocaremos en la modelación de la volatilidad de un índice bursátil, en este caso el Standard & Poor's 500(S&P 500). La im-

portancia de este índice es la representatividad de la situación real del mercado, incluso expresa mejor al mercado accionario y de bonos que muchos otros índices.

En el capítulo 1 daremos a conocer los rendimientos de una serie y propiamente los log-rendimientos de la misma. Se darán bases probabilísticas necesarias para el estudio y estimación de los datos presentados. Para poder presentar los modelos ARCH necesitamos dar a conocer otro tipo de modelos de series de tiempo y sus herramientas que serán de ayuda para el análisis deseado, esto lo veremos en el capítulo 2, para después introducir, en el siguiente capítulo, los modelos ARCH y su generalización, GARCH, así como dar una descripción más a fondo su aplicación. También se presentarán contrastes de estructura y bondad de ajuste, herramientas útiles y necesarias para la comparación y, finalmente en el capítulo 4, veremos una aplicación analizando una serie de datos reales, a los que les ajustaremos los diferentes modelos sugeridos por las herramientas vistas en los capítulos anteriores.

Capítulo 1

Series de Tiempo en Finanzas y sus Características

En este capítulo definiremos los rendimientos de un activo que son la base para el estudio y análisis financiero, así como su función de distribución de probabilidades ($F_r(\cdot)$) para poder calcular la función de verosimilitud, necesaria para la estimación de los parámetros de $F_r(\cdot)$ y mencionaremos otro tipo de procesos que también son de interés en ésta área.

La importancia de fijar el estudio de una serie en los rendimientos es que estos son más representativos pues modelan el cambio de precio que sufre un activo de un día a otro.

1.1. Rendimiento de los Activos

Por la gran representación que tienen los rendimientos, muchos estudios financieros los toman en cuenta en lugar de utilizar el precio de los activos. Campbell, Lo y MacKinlay (1997), véase [3] para mayor detalle, dan dos razones por las cuales utilizar rendimientos:

- Para el promedio de los inversionistas, los rendimientos poseen información completa y un resumen libre de la escala de oportunidades de inversión.
- Los rendimientos son más fáciles de manejar que los precios y sus propiedades estadísticas, como la estacionariedad, son más interesantes .

2CAPÍTULO 1. SERIES DE TIEMPO EN FINANZAS Y SUS CARACTERÍSTICAS

Para un mejor entendimiento en relación a los rendimientos de un activo supongamos que éstos no pagan dividendos y denotaremos como P_t al precio del mismo a tiempo t , donde $t \geq 0$.

Hay que hacer notar que las observaciones de P_t pueden ser tanto a tiempo continuo como discreto, esto depende de cómo se defina el conjunto en donde vive t , i.e., si $t \in T$ con T un conjunto numerable diremos que P_t es a tiempo discreto y en caso contrario diremos que es a tiempo continuo. Para este trabajo, tomaremos el caso discreto que es el que analizaremos de ahora en adelante.

Rendimiento Simple (un periodo)

Este tipo de rendimiento se refiere a tener el precio del activo del tiempo $t - 1$ a t con $t \in \mathbb{N}$. Entonces el rendimiento simple lo definiremos como:

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}, \quad (1.1)$$

el rendimiento simple lo podemos ver como uno mas el rendimiento neto ($1 + R_t$).

Rendimiento Simple (multiperiodo)

Por un rendimiento múltiple de periodo k nos referimos a tener un activo por k -periodos entre $t - k$ y t lo denotamos como $1 + R_t(k)$ que es simplemente el producto de k rendimientos simples de $t - k + 1$ a t , i.e.

$$\begin{aligned} 1 + R_t(k) &= \frac{P_t}{P_{t-k}} \\ &= \frac{P_t}{P_{t-1}} \frac{P_{t-1}}{P_{t-2}} \cdots \frac{P_{t-k+1}}{P_{t-k}} \\ &= (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1}) \\ &= \prod_{j=1}^k (1 + R_{t-j}). \end{aligned} \quad (1.2)$$

A estos rendimientos con multiperiodo se les llama rendimiento compuesto.

La dificultad de tratar con la media geométrica para los rendimientos anualizados motiva otro tipo de rendimientos compuestos como los rendimientos compuestos continuos, o mayormente conocidos como log-rendimientos (r_t).

Al logaritmo natural del rendimiento neto simple de un activo normalmente se le dice log-rendimiento y está dado por

$$r_t = \ln(1 + R_t) = \ln \frac{P_t}{P_{t-1}} = p_t - p_{t-1}, \quad (1.3)$$

donde $p_t = \ln(P_t)$.

Los rendimientos compuestos continuos r_t tienen algunas ventajas sobre los rendimientos netos simples R_t . Primero

$$\begin{aligned} r_t(k) &= \ln(1 + R_t(k)) \\ &= \ln[(1 + R_t)(1 + R_{t-1})\dots(1 + R_{t-k+1})] \\ &= \ln(1 + R_t) + \ln(1 + R_{t-1}) + \dots + \ln(1 + R_{t-k+1}) \\ &= r_t + r_{t-1} + \dots + r_{t-k+1}. \end{aligned} \quad (1.4)$$

Esto nos dice que el rendimiento compuesto continuo es la suma de los rendimientos de un periodo.

Y segundo, las propiedades estadísticas de los log-rendimientos son más manejables puesto que en el área de series de tiempo es más cómodo trabajar con un proceso aditivo que con un proceso multiplicativo.

1.2. Otros procesos en consideración

Además de los rendimientos, también se toman en cuenta otro tipo de procesos como la volatilidad y el comportamiento de rendimientos extremos. La volatilidad está relacionada con la dispersión de los rendimientos, ésta puede ser medida usando la desviación estándar o la varianza entre los rendimientos del mismo activo, comúnmente entre más volatilidad mayor el riesgo. Ésta se utiliza y juega un papel muy importante en la valuación de opciones y administración de riesgos.

Por valores extremos nos referimos a los rendimientos muy positivos o negativos. Por este tipo de valores hacemos alusión a la diferencia que se tiene entre los rendimientos del activo en estudio contra un activo de

4CAPÍTULO 1. SERIES DE TIEMPO EN FINANZAS Y SUS CARACTERÍSTICAS

referencia, usualmente este activo de referencia no tiene tanto riesgo y por lo general es un activo de corto tiempo, véase ([8]). Normalmente se toman como activos de referencia a los bonos gubernamentales. Los valores extremos de rendimientos negativos son importantes en la administración de riesgos mientras que los positivos son de gran riesgo si se tiene una posición corta, véase ([8]). En diferentes análisis financieros se utilizan las propiedades y aplicaciones de los valores extremos para el estudio del impacto de variables económicas, la frecuencia de ocurrencia, entre otras. Otro tipo de series financieras consideradas son las tasas de interés, el tipo de cambio y el rendimiento de los bonos (gubernamentales o corporativos).

En la figura (1.2) ponemos ejemplos de este tipo de procesos antes mencionados. El rendimiento de un índice bursátil (Standard and Poor's, S&P 500), rendimiento de una acción (America Móvil, AMX) y tipo de cambio (euro-dólar) respectivamente. Algo que es muy notorio y hay que resaltar en las gráficas de los rendimientos son los grandes saltos que se tienen, además de que los rendimientos, por lo general, están alrededor del cero. Estos saltos se dan por los grandes cambios en los precios debido a diferentes razones económicas, en el caso del S&P 500, tiene una gran caída en el 2008 debido a la última crisis económica mundial.

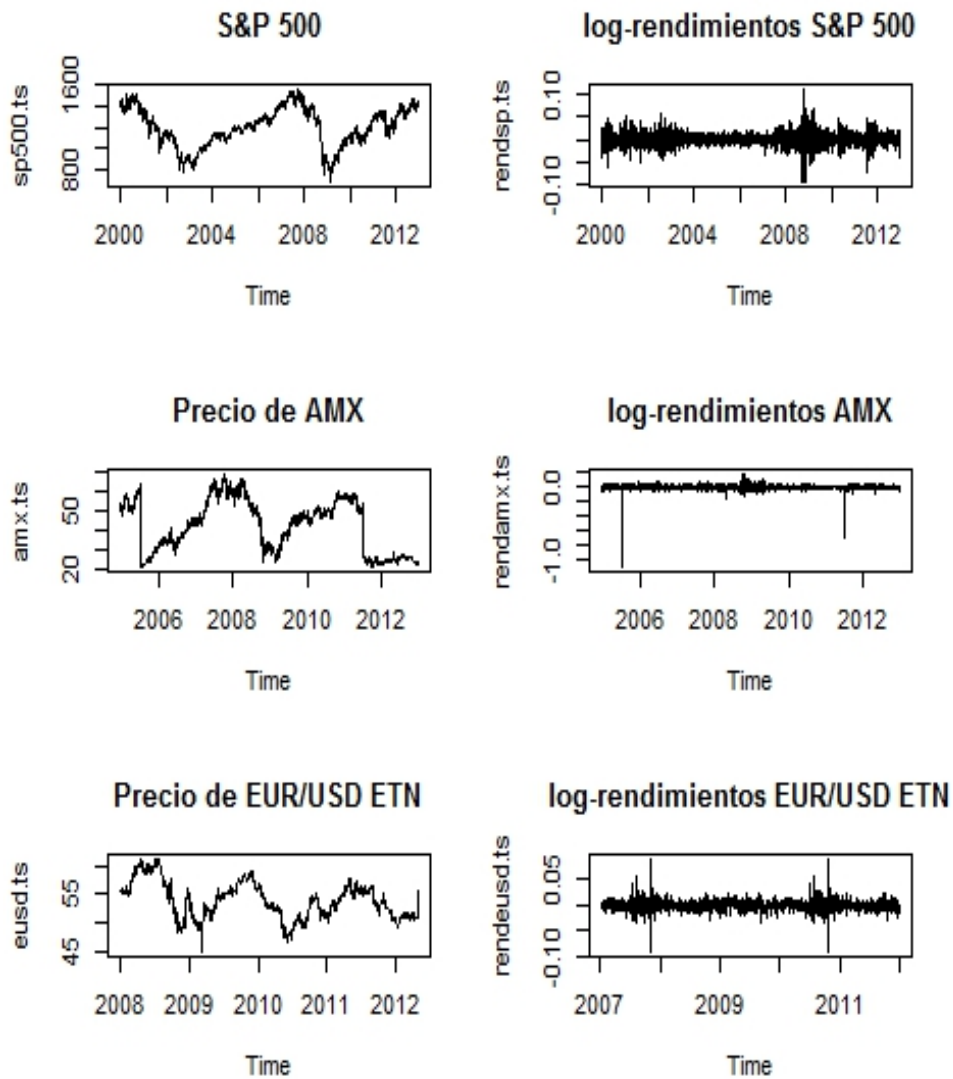


Figura 1.1: La gráfica del lado izquierdo representa el precio y la gráfica del lado derecho los ILog-rendimientos del índice S&P 500, la acción AMX, y el tipo de cambio euro-dólar respectivamente.

6CAPÍTULO 1. SERIES DE TIEMPO EN FINANZAS Y SUS CARACTERÍSTICAS

Capítulo 2

Modelos Lineales y Aplicación

En esta sección nos enfocaremos principalmente a la parte teórica en donde se darán definiciones y conceptos necesarios para poder realizar el análisis deseado a series financieras con series de tiempo. Se introducirán conceptos como el de proceso estocástico, estacionariedad de una serie entre otros. Utilizaremos procesos autorregresivos (AR) y de medias móviles (MA) y daremos a conocer su versión más sencilla, AR(1) y MA(1), para después generalizarlos. En base a estos dos modelos definiremos a los procesos ARMA, autorregresivos de medias móviles; éste tipo de procesos nos serán de ayuda para analizar otros modelos que veremos en el capítulo 3.

2.1. Propiedades y Definiciones

Algunas ramas de estudio en donde podemos encontrar series de tiempo son la Economía, Física, Ingeniería, Sociología y Telecomunicación, por ejemplo:

- Precios de divisas.
- Índice de precios.
- Energía solar.
- Censos de Población.
- Análisis y procesamiento de señales.

La idea general en el estudio de series de tiempo es que, una vez que se tienen los datos, lo que se desea hacer es inferencia sobre los parámetros deseados. Para lograr esto se debe tener o dar uno o varios modelos hipotéticos para poder representar los datos de la forma más adecuada. Una vez dado el/los modelo(s) se hace estimación sobre los parámetros y se realiza una prueba de bondad de ajuste para verificar qué modelo sirve, posteriormente se seleccionan los modelos. Una vez escogido el mejor modelo, éste se utiliza para mejorar nuestro entendimiento acerca de las variables que se tienen. Dado que ya se tomó un modelo apropiado se puede utilizar para diferentes propósitos, esto dependiendo el área de aplicación, la predicción de valores futuros y el control de los mismos. A esta metodología se le conoce como la de Box-Jenkins.

2.1.1. Procesos Estocásticos

Un proceso estocástico es una colección de variables aleatorias, esto es, para cada $t \in T$, X_t es una variable aleatoria. Generalmente tomamos a t como tiempo y X_t es el proceso al tiempo t . Si T es un índice de conjuntos numerable, entonces llamaremos al proceso X_t como un proceso estocástico a tiempo discreto y, si T es continuo, lo llamaremos proceso estocástico a tiempo continuo.

Una serie de tiempo de n datos la pensaremos como una realización de un vector de n variables aleatorias $\{X_{t_1}, \dots, X_{t_n}\}$ que forman un proceso estocástico.

Definición 2.1.1 (Proceso estocástico).

Un proceso estocástico es una colección de variables aleatorias $\{X_t\}_{t \in T}$ indexadas a un conjunto de índices T definida sobre el mismo espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y toma valores en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, donde T es el conjunto de índices y $\mathcal{B}(\mathbb{R})$ es la sigma álgebra de Borel.

El conjunto de índices T normalmente es de la forma $\{1, 2, 3, \dots\}$, $[0, \infty)$, $(-\infty, \infty)$. En nuestro caso tomaremos a T como $\mathbb{N} \cup \{0\}$

2.1.2. Estacionariedad

Definición 2.1.2 (Función de autocovarianza).

Si el proceso $\{X_t\}_{t \in T}$ tiene $\text{Var}(X_t) < \infty$ entonces la función de autocovarianza

$\gamma_x(\cdot, \cdot)$ del proceso está dada por

$$\gamma_{t_1, t_2} = \text{COV}(X_{t_1}, X_{t_2}) = \mathbb{E}[(X_{t_1} - \mu_{t_1})(X_{t_2} - \mu_{t_2})], \text{ para cualesquiera } t_1, t_2 \in T$$

La correlación entre dos variables X y Y definida como

$$\text{COV}(X, Y) / \sqrt{\text{VAR}(X)\text{VAR}(Y)},$$

donde $\text{VAR}(\cdot)$ es la varianza, juega un papel muy importante en el estudio de los modelos que se definirán más adelante, en particular, correlaciones entre la variable de interés y sus valores pasados. A estas correlaciones se les conoce como autocorrelaciones.

Definición 2.1.3 (Proceso estacionario).

El proceso $\{X_t\}$ es estacionario si

- $\mathbb{E}|X_t|^2 < \infty$ para toda t ,
- $\mathbb{E}(X_t) = m$ para toda t ,
- $\gamma_X(r, s) = \gamma_X(r + t, s + t)$ para toda $r, s \in \mathbb{Z}$.

El último punto se refiere a que la función de autocovarianza es independiente para cada t .

Definición 2.1.4 (Función de autocorrelación (ACF)).

La función de autocorrelación, $\rho_X(\cdot)$, del proceso $\{X_t\}_{t \in T}$ con ventana de tiempo h (lag h) está dada por

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \text{Corr}(X_{t+h}, X_t).$$

En la literatura, a la definición dada se le conoce como “estacionariedad débil”. Si un proceso $\{X_t\}$ es estacionario entonces $\gamma_X(r, s) = \gamma_X(r - s, 0)$ para toda $r, s \in \mathbb{Z}$. Por esto es conveniente redefinir la función de autocovarianza de un proceso estacionario como la función de una sola variable.

$$\gamma_X(h) = \gamma_X(h, 0) = \text{COV}(X_{t+h}, X_t) \text{ para toda } h \in \mathbb{Z}.$$

Entonces la función $\gamma_X(\cdot)$ está referida a la función de autocovarianza de $\{X_t\}$ y $\gamma_X(h)$ es su valor con lag h . La ACF de $\{X_t\}$ la podemos ver como la función con lag h como

$$\rho_X(h) = \gamma_X(h) / \gamma_X(0) = \text{CORR}(X_{t+h}, X_t) \text{ para toda } h \in \mathbb{Z}.$$

Definición 2.1.5 (Proceso estrictamente estacionario).

El proceso $\{X_t\}_{t \in T}$ es estrictamente estacionario si para cualquier $\tau > 0$ y $t_i, t_{i+\tau} \in T, i \in \{1, \dots, n\}$, entonces $(X_{t_1}, \dots, X_{t_n})$ y $(X_{t_1+\tau}, \dots, X_{t_n+\tau})$ tiene la misma distribución conjunta, es decir:

$$\begin{aligned} F_{X_{t_1}, X_{t_2}, \dots, X_{t_n}}(a_1, a_2, \dots, a_n) &= \mathbb{P}(X_{t_1} \leq a_1, X_{t_2} \leq a_2, \dots, X_{t_n} \leq a_n) \\ &= \mathbb{P}(X_{t_1+\tau} \leq a_1, X_{t_2+\tau} \leq a_2, \dots, X_{t_n+\tau} \leq a_n) \\ &= F_{X_{t_1+\tau}, X_{t_2+\tau}, \dots, X_{t_n+\tau}}(a_1, a_2, \dots, a_n) \end{aligned}$$

Al inicio del capítulo se mencionó la idea general para el estudio de series de tiempo, para lograr esto, el primer paso en el análisis consiste en graficar la serie. La inspección de la gráfica puede sugerir la posibilidad de representar los datos como una realización de un proceso que puede tener todos o alguno de los siguientes componentes:

$$X_t = f(m_t, s_t, Y_t) = \begin{cases} m_t + s_t + Y_t, & \text{modelo aditivo} \\ m_t * s_t * Y_t, & \text{modelo multiplicativo,} \end{cases}$$

donde m_t es el componente de tendencia, s_t es el componente estacional de periodo d , donde $s_t = s_{t+d} = s_{t+2d} = \dots$ y Y_t es el componente aleatorio. Obsérvese que al modelo multiplicativo lo podemos llevar al aditivo con solo aplicarle logaritmo.

2.1.3. Estimación y eliminación de tendencia y componente estacional

A continuación veremos algunos métodos que se han propuesto para identificar y describir los componentes de tendencia y la parte estacional o cíclica de la serie de tiempo.

Método1

Si tenemos una serie que tiene solo la parte de tendencia $X_t = m_t + Y_t$, una manera de estimar m_t sería vía mínimos cuadrados. Es decir, procederemos a estimar la tendencia de una familia de funciones de la forma

$$m_t = a_0 + a_1 t + a_2 t^2,$$

y escogeremos los a_0, a_1, a_2 que minimizen $\sum (X_t - m_t)^2$. Así se obtendrá

$$\hat{m}_t = \hat{a}_0 + \hat{a}_1 t + \hat{a}_2 t^2.$$

Una vez estimada la tendencia podemos estimar la parte aleatoria de la siguiente manera:

$$\hat{Y}_t = X_t - \hat{m}_t, t \in \{1, \dots, n\}.$$

Método2

Suavizamiento de la media vía promedios móviles. Sea q un entero no negativo y consideremos el promedio móvil de dos lados como

$$\hat{m}_t = \sum_{j=-q}^q a_j X_{t+j} \text{ para } t \in \{q+1, q+2, \dots, n-q\},$$

donde $\sum_{j=-q}^q a_j = 1$.

Un caso particular de este tipo de ajuste de la tendencia es si suponemos que $a_j = \frac{1}{2q+1}$, es decir

$$\hat{m}_t = \frac{1}{2q+1} \sum_{j=-q}^q X_{t+j}, t \in \{q+1, q+2, \dots, n-q\}.$$

Antes de describir el último método introduciremos los operadores de retraso y diferencia.

Definición 2.1.6 (Operador de retraso).

Denotamos por B al operador retraso que actúa sobre el tiempo de la siguiente manera:

$$\begin{aligned} BX_t &= X_{t-1} \\ B^2 X_t &= B(BX_t) = BX_{t-1} = X_{t-2} \\ &\vdots \\ B^j X_t &= X_{t-j}. \end{aligned}$$

(2.1)

Por convención tomaremos a $B^0 X_t = X_t$.

Definición 2.1.7 (Operador de diferencia).

A partir del operador de retraso definimos el operador diferencia como

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

El operador de diferencia se puede manejar como si fuera un polinomio, es decir

$$\nabla^2 X_t = (1 - B)(1 - B)X_t = X_t - 2X_{t-1} + X_{t-2}.$$

Entonces si se desea diferenciar un proceso j veces se tiene

$$\begin{aligned} \nabla^j X_t &= (1 - B)^j X_t \\ &= \sum_{k=0}^j \binom{j}{k} (-1)^k X_{t-k}. \end{aligned} \tag{2.2}$$

Método 3

Cuando la serie de tiempo $\{X_t\}$ tiene una tendencia lineal, $m_t = at + b$, al aplicar el operador diferencia ∇ obtenemos $\nabla m_t = a$. De la misma forma, si tenemos una serie de tiempo $X_t = m_t + Y_t$ donde la tendencia es polinomial de grado k , $m_t = \sum_{j=0}^k a_j t^j$, al aplicar el operador diferencia a la tendencia, tenemos que $\nabla^k m_t = k! a_k$ y por lo tanto tenemos que

$$\nabla^k X_t = k! a_k + \nabla^k Y_t.$$

Lo que se pretende ahora es quitar la tendencia y la parte estacional en un modelo general

$$X_t = m_t + s_t + Y_t,$$

donde $\mathbb{E}[Y_t] = 0$, $s_t = s_{t+d}$ y $\sum_{j=1}^d s_j = 0$.

Definición 2.1.8.

Un proceso de ruido blanco denotado por $\{\epsilon_t\}_{t \in T}$, es un proceso estocástico formado por variables aleatorias independientes e idénticamente distribuidas, no correlacionadas de media cero y varianza constante σ_ϵ^2

Por lo tanto

$$\begin{aligned}\gamma_0 &= \sigma_\epsilon^2, \\ \gamma_h &= 0, h \neq 0\end{aligned}$$

y

$$\rho_h = \frac{\gamma_h}{\gamma_0} \text{ para } h \neq 0.$$

Para el ruido blanco, todas las ACF son cero.

2.1.4. Procesos lineales

Un proceso $\{X_t\}$ es lineal si para todo $t \in \mathbb{N}$ se puede escribir:

$$X_t = \mu + \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \quad (2.3)$$

donde μ es la media de X_t , $\psi_0 = 1$ con $\sum_{j=0}^{\infty} |\psi_j| < \infty$ y $\{\epsilon_t\}$ es un proceso de ruido blanco. A las ψ se les conoce como los pesos de X_t . Como $\{X_t\}$ es estacionario, entonces podemos obtener fácilmente su media y varianza usando la independencia de $\{\epsilon_t\}$ como

$$\begin{aligned}\mathbb{E}(X_t) &= \mu \\ \text{Var}(X_t) &= \sigma_\epsilon^2 \sum_{i=0}^{\infty} \psi_i^2,\end{aligned} \quad (2.4)$$

donde σ_ϵ^2 es la varianza de ϵ .

La función de autocovarianza con lag l para el proceso $\{X_t\}$ es

$$\begin{aligned}\gamma_l = \text{Cov}(X_t, X_{t-l}) &= \mathbb{E} \left[\left(\sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \right) \left(\sum_{j=0}^{\infty} \psi_j \epsilon_{t-l-j} \right) \right] \\ &= \mathbb{E} \left(\sum_{i,j=0}^{\infty} \psi_i \psi_j \epsilon_{t-i} \epsilon_{t-l-j} \right) \\ &= \sum_{j=0}^{\infty} \psi_{j+l} \psi_j \mathbb{E}(\epsilon_{t-l-j}^2)^1 \\ &= \sigma_\epsilon^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+l}.\end{aligned} \quad (2.5)$$

Así mismo, la función de autocorrelación queda

$$\begin{aligned}\rho_l &= \frac{\gamma_l}{\gamma_0} \\ &= \frac{\sum_{i=0}^{\infty} \psi_i \psi_{i+l}}{1 + \sum_{i=1}^{\infty} \psi_i^2}, \quad l \geq 0.\end{aligned}\tag{2.6}$$

En este tipo de modelos se estudia la estacionariedad, la dependencia, la función de autocorrelación, la simulación y la tendencia. Los modelos que tienen este tipo de características y los que veremos en este trabajo son:

- Modelos Autorregresivos (AR).
- Modelos de Promedios Móviles (MA).
- ARMA.

2.2. Modelo Autorregresivo (AR)

Los modelos autorregresivos pueden describirse como aquellos en los que una variable o conjunto de variables se explican, al menos en parte, en función de los valores pasados de esa misma variable o conjunto de variables. Estos modelos han cobrado gran importancia en el campo de la econometría y la economía. Se ha demostrado que modelos sencillos de este tipo, con un pequeño número de variables y parámetros, compiten, incluso con ventaja, en su capacidad de predicción y simulación con grandes modelos macroeconómicos.

Definición 2.2.1.

El proceso $\{X_t\}$ es un autorregresivo de orden p , $AR(p)$, si para $p \geq 1$ un entero y ϕ_1, \dots, ϕ_p constantes reales tenemos que

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \epsilon_t,\tag{2.7}$$

donde ϵ_t es ruido blanco con media 0 y varianza σ_ϵ^2 .

Para el proceso $\{r_t\}$, mencionado en el primer capítulo, un $AR(1)$ tendrá la forma

$$r_t = \phi_0 + \phi_1 r_{t-1} + \epsilon_t.\tag{2.8}$$

Una vez definido un $AR(1)$ para el proceso $\{r_t\}$ continuaremos con propiedades de estos modelos en base a los log-rendimientos.

2.2.1. Propiedades

Suponiendo que el proceso $\{r_t\}$ es débilmente estacionario tenemos que $\mathbb{E}(r_t) = \mu$, $Var(r_t) = \gamma_0$ y $Cov(r_t, r_{t-j}) = \gamma_j$, donde μ, γ_0 son constantes y γ_j es una función de j . El cálculo de la esperanza y varianza del proceso es sumamente fácil y de esto resulta

$$\mathbb{E}(r_t) = \phi_0 + \phi_1 \mathbb{E}(r_{t-1}),$$

y bajo la condición de estacionariedad, $\mathbb{E}(r_t) = \mathbb{E}(r_{t-1})$, la esperanza queda

$$\mathbb{E}(r_t) = \mu = \frac{\phi_0}{1 - \phi_1}.$$

Es importante notar que la $\mathbb{E}(r_t)$ existe si $\phi_1 \neq 1$ y que ésta es cero si y sólo si $\phi_0 = 0$. Si despejamos ϕ_0 de la esperanza y lo sustituimos en (2.8) el modelo AR(1) queda

$$r_t - \mu = \phi_1(r_{t-1} - \mu) + \epsilon_t. \quad (2.9)$$

Si volvemos a hacer el mismo método repetidas veces, entonces la ecuación anterior queda

$$\begin{aligned} r_t - \mu &= \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_1^2 \epsilon_{t-2} + \dots \\ &= \sum_{i=0}^{\infty} \phi_1^i \epsilon_{t-i}. \end{aligned} \quad (2.10)$$

Si al resultado anterior lo elevamos al cuadrado y le sacamos esperanza (bajo el supuesto de estacionariedad), entonces obtenemos la varianza del proceso

$$Var(r_t) = \frac{\sigma_\epsilon^2}{1 - \phi_1^2}. \quad (2.11)$$

Una condición necesaria y suficiente para que un AR(1) sea débilmente estacionario es $|\phi_1| < 1$.

AR(1)

Multiplicando (2.9) por ϵ_t , usando la independencia de ϵ_t y r_{t-1} y tomando la esperanza obtenemos

$$\mathbb{E}[\epsilon_t(r_t - \mu)] = \sigma_\epsilon^2.$$

Multiplicando (2.9) por $(r_{t-l} - \mu)$, usando el resultado anterior y tomando la esperanza se tiene que la ACF de $\{r_t\}$ satisface

$$\gamma_l = \begin{cases} \phi_1 \gamma_1 + \sigma_\varepsilon^2, & \text{si } l = 0 \\ \phi_1 \gamma_{l-1}, & \text{si } l > 0; \end{cases}$$

Luego la ACF para $\{r_t\}$ es

$$\rho_l = \phi_1 \rho_{l-1}, \text{ para } l > 0.$$

Como $\rho_0 = 1$, tenemos que $\rho_l = \phi_1^l$. Esto indica que la ACF de un AR(1) decae exponencialmente.

AR(2)

Realizando los mismos cálculos que con AR(1), obtenemos

$$\mathbb{E}(r_t) = \mu = \frac{\phi_0}{1 - \phi_1 - \phi_2}.$$

Siguiendo con los mismos pasos hechos con el modelo AR(1), sustituimos ϕ_0 de la ecuación anterior, multiplicamos ésta por $(r_{t-l} - \mu)$ y tomando la esperanza, considerando que $\mathbb{E}[(r_{t-l} - \mu)\varepsilon_t]$, se obtiene

$$\gamma_l = \phi_1 \gamma_{l-1} + \phi_2 \gamma_{l-2}, \text{ para } l > 0.$$

Dividiendo lo último por γ_0 obtendremos la ACF para r_t :

$$\rho_l = \phi_1 \rho_{l-1} + \phi_2 \rho_{l-2}, \text{ para } l > 0, \quad (2.12)$$

Este resultado de la ACF satisface una ecuación en diferencias de segundo orden

$$(1 - \phi_1 B - \phi_2 B^2) = 0,$$

donde B es el operador de retraso.

El polinomio de segundo orden asociado a la ecuación en diferencias anterior es

$$1 - \phi_1 x - \phi_2 x^2 = 0, \quad (2.13)$$

las soluciones para esta ecuación en diferencias son

$$x = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}.$$

Si $\phi_1^2 + 4\phi_2 < 0$, entonces las soluciones para (2.13) serán complejas. En áreas de aplicación las raíces complejas son de suma importancia ya que dan pie a ciclos.

A continuación se muestra la simulación de un proceso con raíces complejas, donde se observá claramente la formación de ciclos. Para poder verlo de forma más clara, las líneas punteadas marcan estos ciclos.

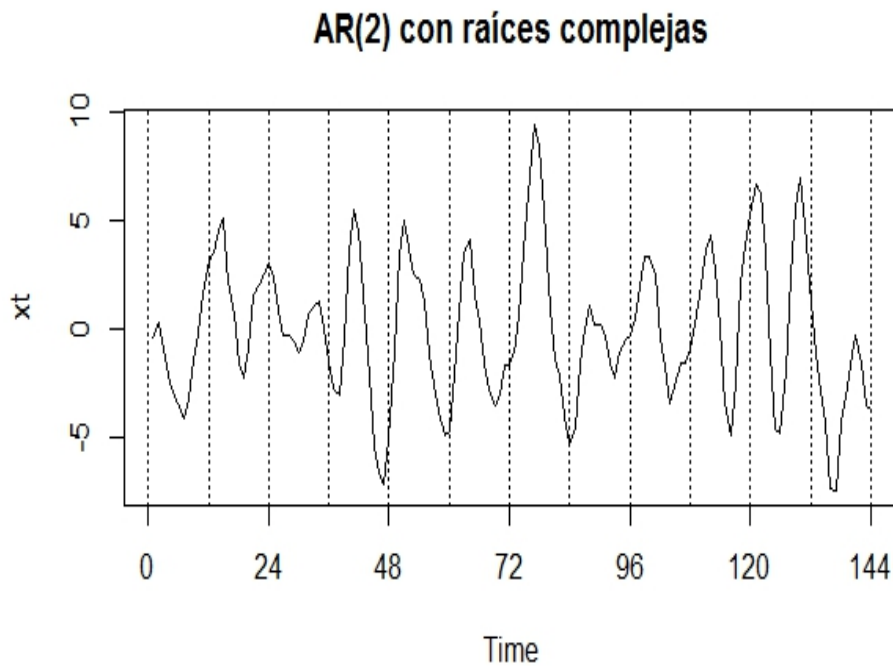


Figura 2.1: Simulación de un AR(2) con $\phi_1 = 1.5$ y $\phi_2 = -.75$

Estacionariedad

La condición para que un AR(2) sea estacionario es que sus raíces sean en módulo menor que la unidad. Bajo esta condición, la ecuación en diferencias (2.12) asegura que la ACF del modelo converge a cero cuando se incrementa el "lag".

AR(p)

Es fácil generalizar a partir los resultados de AR(1) y AR(2) que la esperanza para un modelo autorregresivo de orden p está dada por

$$\mathbb{E}(r_t) = \frac{\phi_0}{1 - \phi_1 - \dots - \phi_p}$$

y que su polinomio asociado al modelo es

$$1 - \phi_1 x - \phi_1 x^2 - \dots - \phi_p x^p = 0.$$

La ecuación en diferencias dada por la ACF será

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \rho_l = 0, \quad l > 0.$$

2.2.2. Identificar un AR

La función de autocorrelación parcial (PACF) es una herramienta que nos ayuda a identificar de forma sencilla el orden p de un proceso AR, puesto que, en general, no es fácil detectar el orden p usando sólo la función de autocorrelación (muestral).

Función de autocorrelación parcial (PACF)

La PACF de una serie de tiempo es una función de su ACF. Una manera simple, pero efectiva, de representar a la PACF es considerando los siguientes modelos AR:

$$\begin{aligned}
 r_t &= \phi_{0,1} + \phi_{1,1}r_{t-1} + e_{1t} \\
 r_t &= \phi_{0,2} + \phi_{1,2}r_{t-1} + \phi_{2,2}r_{t-2} + e_{2t} \\
 r_t &= \phi_{0,3} + \phi_{1,3}r_{t-1} + \phi_{2,3}r_{t-2} + \phi_{3,3}r_{t-3} + e_{3t} \\
 r_t &= \phi_{0,4} + \phi_{1,4}r_{t-1} + \phi_{2,4}r_{t-2} + \phi_{3,4}r_{t-3} + \phi_{4,4}r_{t-4} + e_{4t} \\
 &\vdots \\
 &\vdots
 \end{aligned}$$

donde $\phi_{0,j}$, $\phi_{i,j}$ y e_{jt} son una constante, el coeficiente de r_{t-i} y el error de un AR(j) respectivamente y $\phi_{k,k}$ es el coeficiente de correlación parcial de orden k que es una medida de relación lineal entre las observaciones separadas por k periodos, eliminando el efecto de las variables intermedias. Luego para el cálculo de la PACF tenemos que $\phi_{1,1} = \rho_1$ y para $k > 0$

$$\phi_{2,2} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}, \phi_{3,3} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}, \dots, \phi_{k,k} = \frac{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \dots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-2} & \rho_{k-3} & \dots & 1 & \rho_{k-1} \\ \rho_{k-1} & \rho_{k-2} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-2} & \rho_{k-3} & \dots & 1 & \rho_1 \\ \rho_{k-1} & \rho_{k-2} & \dots & \rho_1 & 1 \end{vmatrix}}.$$

Con un proceso estacionario $\{X_t\}$ con PACF $\{\phi_{k,k} : k = 1, 2, 3, \dots\}$ obtendremos el estimador de $\phi_{k,k}$, denotado por $\hat{\phi}_{k,k}$, como

$$\hat{\phi}_{1,1} = \hat{\rho}_1, \hat{\phi}_{k,k} = \frac{\begin{vmatrix} 1 & \hat{\rho}_1 & \dots & \hat{\rho}_{k-2} & \hat{\rho}_1 \\ \hat{\rho}_1 & 1 & \dots & \hat{\rho}_{k-3} & \hat{\rho}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{\rho}_{k-2} & \hat{\rho}_{k-3} & \dots & 1 & \hat{\rho}_{k-1} \\ \hat{\rho}_{k-1} & \hat{\rho}_{k-2} & \dots & \hat{\rho}_1 & \hat{\rho}_k \end{vmatrix}}{\begin{vmatrix} 1 & \hat{\rho}_1 & \dots & \hat{\rho}_{k-2} & \hat{\rho}_{k-1} \\ \hat{\rho}_1 & 1 & \dots & \hat{\rho}_{k-3} & \hat{\rho}_{k-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{\rho}_{k-2} & \hat{\rho}_{k-3} & \dots & 1 & \hat{\rho}_1 \\ \hat{\rho}_{k-1} & \hat{\rho}_{k-2} & \dots & \hat{\rho}_1 & 1 \end{vmatrix}}.$$

Si se tiene un proceso estacionario, se puede demostrar que la PACF muestral tiene las siguientes propiedades:

- $\hat{\phi}_{p,p}$ converge a ϕ_p si el tamaño de la muestra crece a infinito.
- $\hat{\phi}_{i,i}$ converge a cero para toda $i > p$.
- La varianza asintótica de $\hat{\phi}_{i,i}$ es $1/T$ para $i > p$.

Como podemos observar en las figuras (2.2) y (2.3) no se puede determinar bien el orden de un modelo AR con la ACF, sin embargo, al fijarnos en la PACF es más fácil observar qué orden podría tener la serie. Fijándonos en las primeras 3 gráficas que corresponden a un AR(3), es claro que la ACF no nos ayuda a saber el orden de este modelo, sin embargo, como ya lo mencionamos, para la PACF a partir del “lag 3” las autocorrelaciones ya no son significativas.

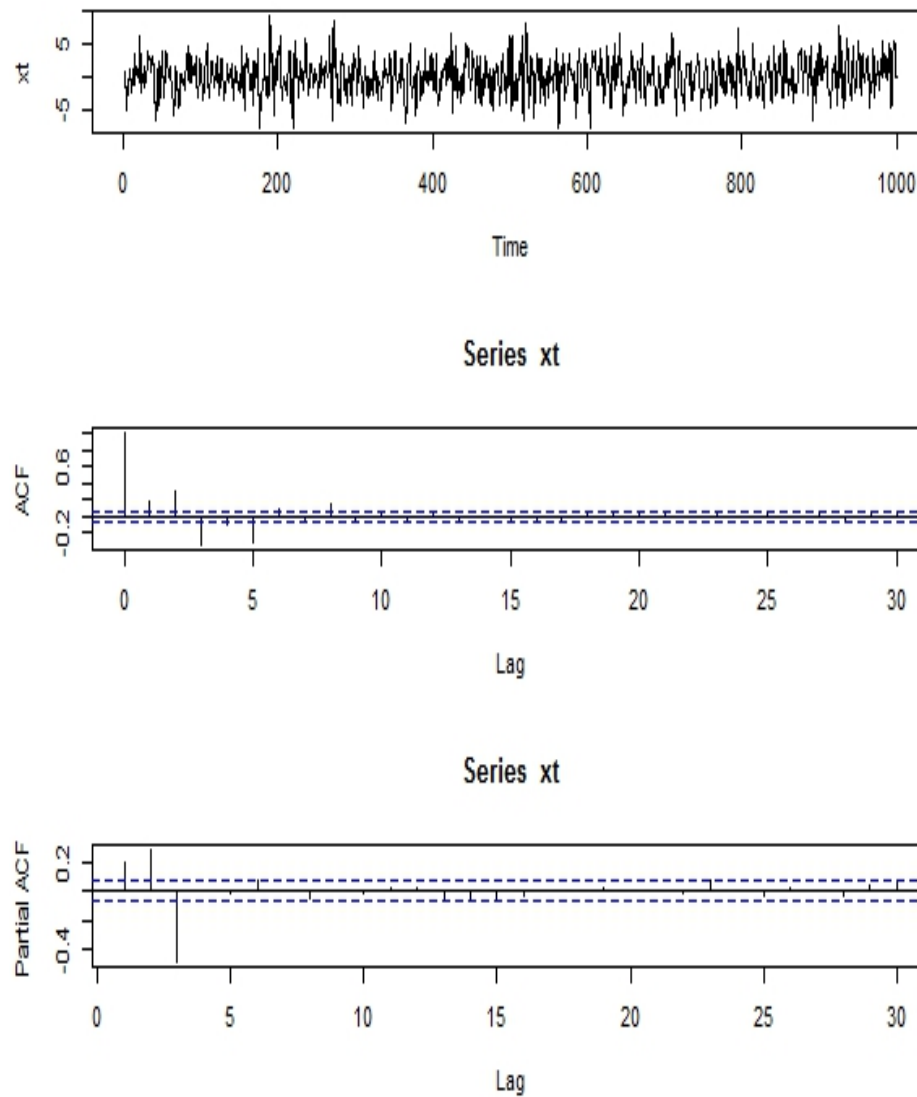


Figura 2.2: Simulación de un AR(3) y AR(2) con sus ACF's y PACF's respectivamente

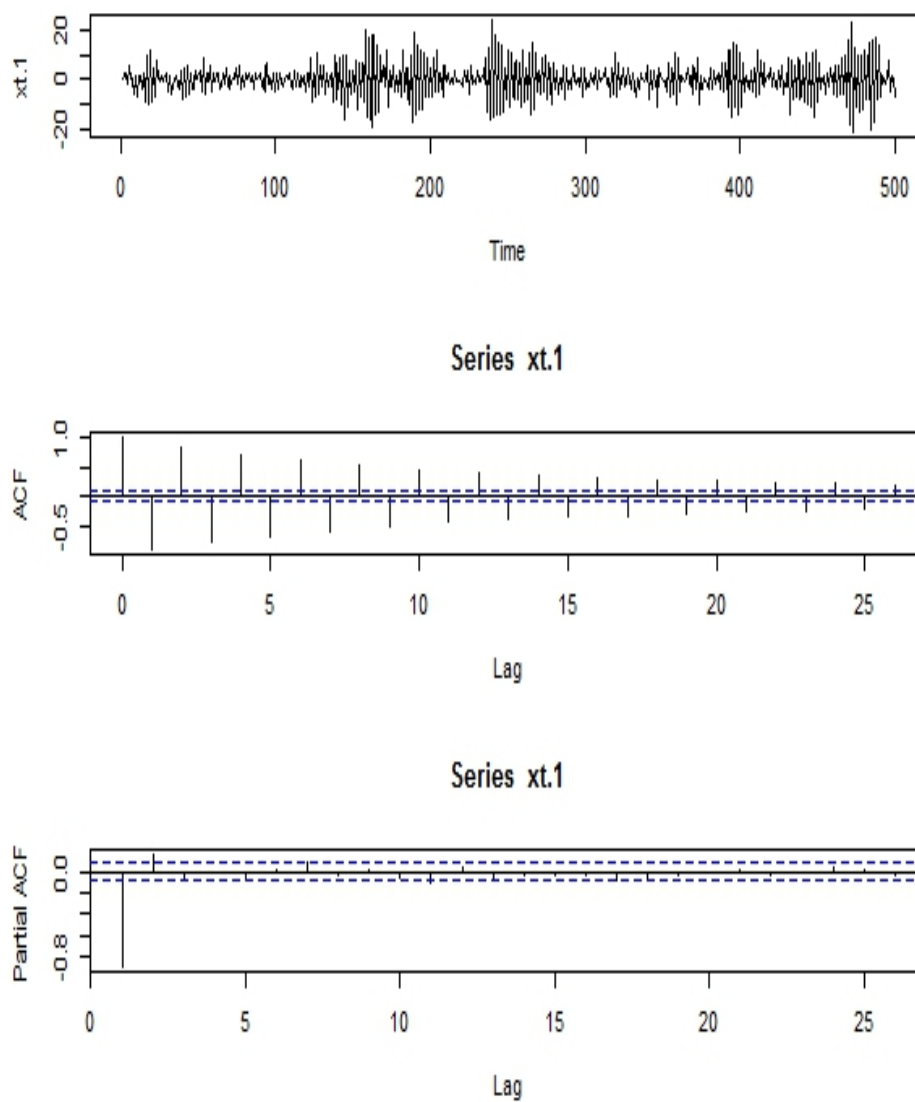


Figura 2.3: Simulación de un AR(2) y sus respectivas ACF's y PACF's

2.3. Modelo de Promedios Móviles (MA)

Los procesos de medias o promedios móviles de orden q , denotado por $MA(q)$, se utilizan generalmente para medir el impulso y definir las áreas de apoyo posible y la resistencia. Frecuentemente es un indicador utilizado en el análisis técnico que muestra el valor promedio del precio de un valor en un período determinado.

Se dice que algunos corredores tienen la creencia de que el precio de ciertos bienes financieros (commodities) tienden a repetir valores fijos. Esto es, cuando el precio actual es menor que ese valor, el precio tiende a incrementarse y, cuando el precio es más grande, el precio decrece; a este fenómeno se le conoce como regreso a la media.

Definición 2.3.1.

Diremos que el proceso $\{r_t\}$ es un proceso de promedios móviles de orden q , $MA(q)$, si para q un entero mayor a 1 y $\theta_1, \dots, \theta_q, \mu$ son constantes reales tales que $\forall t \in T$ tenemos

$$\begin{aligned} r_t &= \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \\ &= \mu + (1 + \theta_1 B + \dots + \theta_q B^q) \epsilon_t \\ &= \mu + \theta_q(B) \epsilon_t \end{aligned} \quad (2.14)$$

donde ϵ_t es ruido blanco y $\theta_q(B)$ será el polinomio de media móvil dado por $1 + \theta_1 B + \dots + \theta_q B^q$.

2.3.1. Propiedades

Al igual que con los modelos AR, nos fijaremos en los $MA(1)$ y $MA(2)$ para después generalizar.

Estacionariedad

Tomando la esperanza y varianza a un $MA(1)$ tenemos que

$$\begin{aligned} \mathbb{E}(r_t) &= c_0 \\ \text{Var}(r_t) &= (1 + \theta_1^2) \sigma_\epsilon^2, \end{aligned}$$

son invariantes en el tiempo. Este resultado se puede generalizar a un $MA(q)$ en donde podemos observar dos propiedades: la primera, que la

constante c_0 es la media del proceso (μ en la ecuación 2.14) y, la segunda, que la varianza para un MA(q) está dada por

$$\text{Var}(r_t) = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)\sigma_\varepsilon^2$$

2.3.2. Identificar un MA

A diferencia de los AR que utilizan la PACF para obtener el orden, veremos que la ACF es una herramienta útil para calcular el orden de un MA.

ACF

Supongamos que para un MA(1) $c_0 = 0$, al Multiplicarlo por r_{t-1} y tomando esperanza obtenemos

$$\begin{aligned}\gamma_1 &= -\theta_1\sigma_\varepsilon^2 \\ \gamma_l &= 0 \text{ para } l > 1\end{aligned}$$

y a partir de este resultado se obtiene

$$\begin{aligned}\rho_0 &= 1, \\ \rho_1 &= \frac{-\theta_1}{1 + \theta_1^2}, \\ \rho_l &= 0 \text{ para } l > 1\end{aligned}$$

Para un MA(2) los coeficientes de autocorrelación son

$$\begin{aligned}\rho_1 &= \frac{-\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2}, \\ \rho_2 &= \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} \\ \rho_l &= 0 \text{ para } l > 2\end{aligned}$$

Observemos que en ambos casos, el MA(1) y MA(2), la ACF es cero a partir de $lag1$ y $lag2$ respectivamente, entonces para un MA(q) su ACF será cero a partir de $lagq$, es decir $\rho_l = 0$ para $l > q$. De esta propiedad podemos decir que un MA(q) es un modelo con memoria finita ya que está linealmente relacionado con sus primeros q valores.

Para ejemplificar lo antes dicho, véase la figura (2.4) donde se muestran dos modelos de promedios móviles. El primero corresponde a un MA(3) porque si nos fijamos en su ACF, de ésta sobresalen 3 autocorrelaciones. Si nos fijamos en el lag 4, también sobre sale, sin embargo las bandas que se pintan en la ACF son al 95 % y, si quisieramos ser más estrictos, entonces podemos hacer las bandas al 99 % y ésta observación ya no será significativa. Las bandas de confianza al 95 % están dadas por $\pm 1.96 / \sqrt{n}$, donde n es el tamaño de la muestra. Esto se debe gracias al hecho de que la estimación de la ρ_i tiende a ser una $N(\rho_i, 1 / \sqrt{n})$, ($\hat{\rho}_i \xrightarrow{n \rightarrow \infty} N(\rho_i, 1 / \sqrt{n})$). Para ver este resultado más a fondo, véase [2].

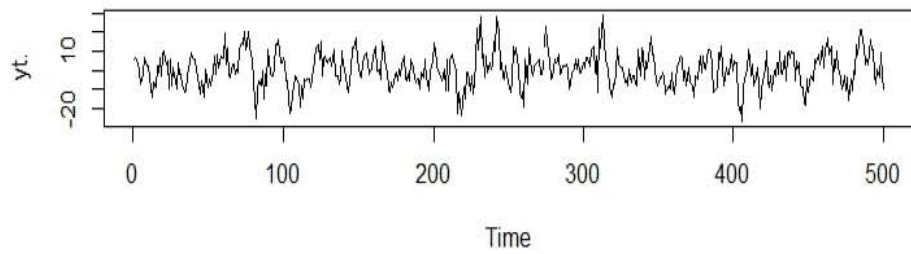
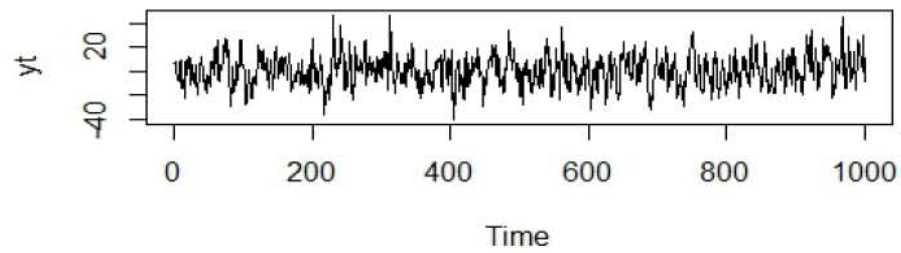
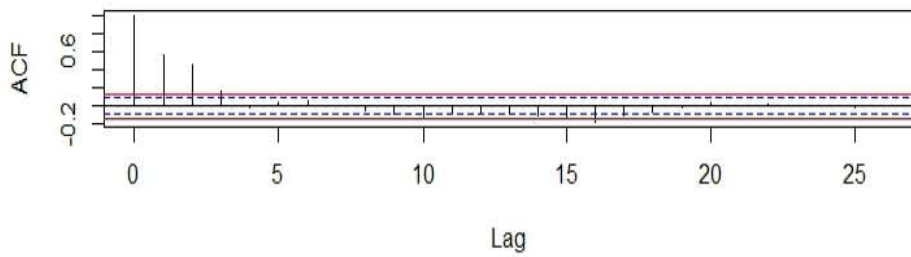
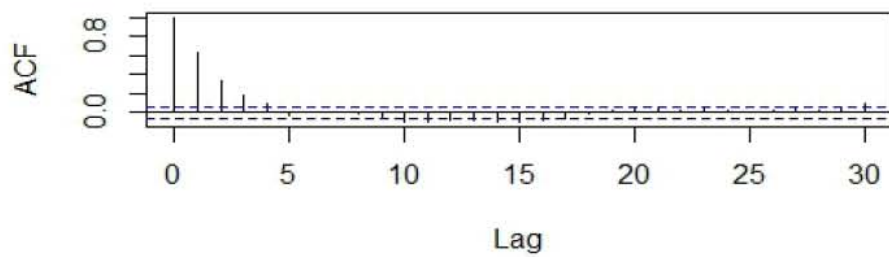
**Series yt.****Series yt**

Figura 2.4: Simulación de un MA(3) y MA(4) con su ACF respectivamente

2.4. Modelos ARMA

En algunas aplicaciones los modelos AR y MA pueden no resultar muy convenientes pues uno puede necesitar un orden muy grande y por ello tendrá muchos parámetros para poder ajustarse a los datos. Una posible solución a esto es combinar estos modelos para obtener una generalización de los procesos anteriores, a estos nuevos procesos se les conoce como modelos autorregresivos de promedios móviles (ARMA).

En el área de finanzas, para el caso de los rendimientos de activos, este modelo no se ocupa con frecuencia. Sin embargo, para la modelación de volatilidad es un modelo relevante. En esta sección veremos sólo el caso para un ARMA(1,1).

Un ARMA(1,1) está dado como

$$r_t - \phi_1 r_{t-1} = \epsilon_t - \theta_1 \epsilon_{t-1}. \quad (2.15)$$

Para que este modelo tenga sentido es necesario que $\phi_1 \neq \theta_1$, de no ser así, lo que obtendríamos de esto sería solo ruido blanco.

2.4.1. Propiedades de un ARMA(1,1)

Las propiedades de un ARMA(1,1) son generalizaciones de un AR(1) con algunas modificaciones por el impacto del MA(1). Haciendo un procedimiento parecido al realizado con AR y MA y, suponiendo que $\phi_0 = 0$, tenemos

$$\begin{aligned} \mathbb{E}(r_t) &= \mu = \frac{\phi_0}{1 - \phi_1}, \\ \text{Var}(r_t) &= \frac{(1 - 2\phi_1\theta_1 + \phi_1^2)\sigma_\epsilon^2}{1 - \phi_1^2}, \end{aligned}$$

lo que nos indica que la serie es débilmente estacionaria.

Para que este proceso sea estacionario debe cumplir $|\phi_1| < 1$, propiedad que también tiene el AR(1).

Las funciones de autocovarianza y autocorrelación están dadas por:

$$\begin{aligned} \gamma_l - \phi_1 \gamma_{l-1} &= 0 \quad l > 1, \\ \rho_1 &= \phi_1 - \frac{\theta_1 \sigma_\epsilon^2}{\gamma_0}, \\ \rho_l &= \phi_1 \rho_{l-1}; \quad l > 1. \end{aligned}$$

La ACF de un ARMA(1,1) estacionario decae exponencialmente a partir de lag2, por lo que la ACF no es de utilidad para saber el orden de un ARMA. Se puede demostrar que la PACF de un ARMA(1,1) no se anula después de un lag finito.

Un ARMA(p,q) está representado como

$$r_t = \phi_0 + \sum_{i=1}^p \phi_i r_{t-i} + \epsilon_t - \sum_{i=1}^q \theta_i \epsilon_{t-i}.$$

Usando el operador de retraso (B), podemos escribir el modelo como

$$(1 - \phi_1 B - \dots - \phi_p B^p) r_t = \phi_0 + (1 - \theta_1 B - \dots - \theta_q B^q) \epsilon_t. \quad (2.16)$$

Para esta representación de un ARMA(p,q) tenemos que el ARMA(1,1) está dado por (2.15). Si todas las soluciones de la ecuación

$$1 - \phi_1 y - \dots - \phi_p y^p = 0$$

$$1 - \theta_1 x - \dots - \theta_q x^q = 0$$

son en modulo menores que la unidad, entonces el modelo ARMA(1,1) es débilmente estacionario.

Identificar un ARMA

Como mencionamos anteriormente la ACF y la PACF pueden ser herramientas no tan útiles para determinar el orden de un ARMA o bien un tanto complicadas para determinar el orden, aunque muy utilizadas. Sin embargo, Tsay y Tiao (1984) propusieron la función de autocorrelación extendida (EACF) para determinar el orden de un ARMA. En términos simples la idea de la eacf es muy sencilla, la salida de la EACF es una tabla donde las filas corresponden al orden p de un AR y las columnas al orden q de un MA. En estas se ven O's y X's y el orden estará determinado en donde se forme un triángulo de O's en la parte superior izquierda, i.e., el triángulo de O's debe tener el vértice superior izquierdo en la posición (p,q).

En seguida mostramos la gráfica de un ARMA en la figura 2.5 y, si nos fijamos tanto en su ACF como en su PACF, no es posible saber el orden de este modelo. Una vez visto esto, como se mencionó para poder determinar

el orden de este tipo de procesos nos fijamos en la EACF dada por la tabla 2.1 en donde se puede ver que en la entrada (1,1) empiezan las "O's".

La tabla se construye de la siguiente manera:

- El valor absoluto correspondiente a la EACF que es mayor o igual que $2/\sqrt{N}$ se denota como X.
- La O corresponde a los valores absolutos de la EACF que son menores a $2/\sqrt{N}$.

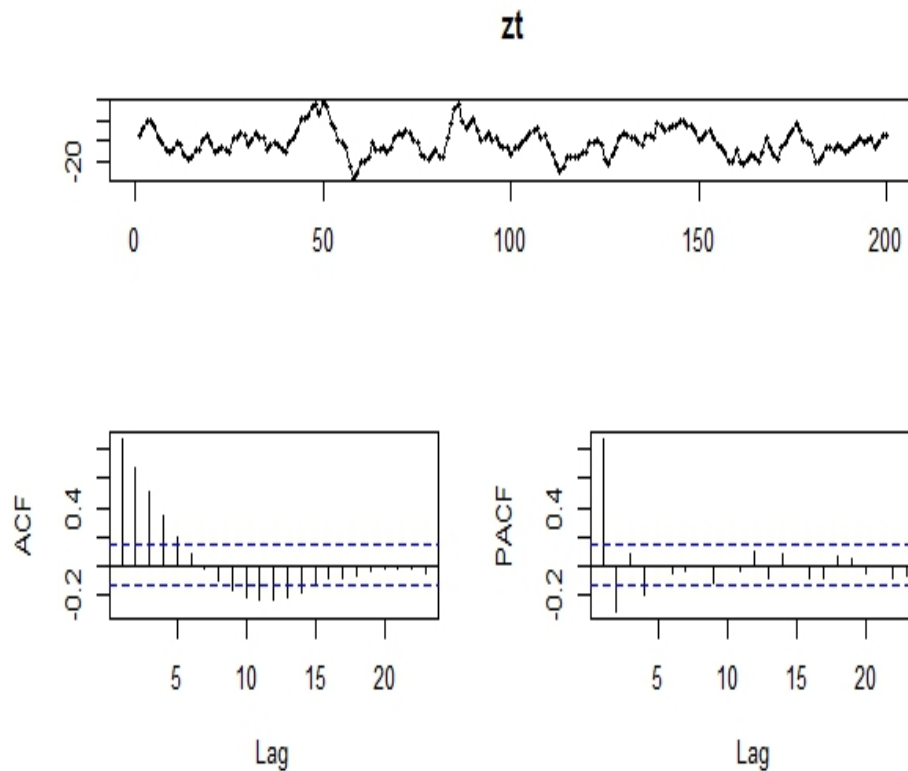


Figura 2.5: Simulación de un proceso ARMA(1,1) con $n = 200$, $\phi_1 = 0.8$ y $\theta_1 = 2$

AR\MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	o	o	o	o	o	x	x	x	x	x	x
1	x	o	o	o	o	o	o	o	o	x	o	o	o	o
2	x	x	o	o	x	o	o	o	o	o	o	o	o	o
3	x	x	x	o	o	o	o	o	o	o	o	o	o	o
4	x	x	x	o	o	x	o	o	o	o	o	o	o	o
5	o	x	x	x	o	o	o	o	o	o	o	o	o	o
6	o	x	x	x	o	x	o	o	o	o	o	o	o	o
7	x	x	x	x	o	x	x	o	x	o	o	o	o	o

Tabla 2.1: EACF de un ARMA(1,1)

Capítulo 3

Modelos no Lineales y Aplicación

En este capítulo nos apoyaremos de modelos econométricos y métodos estadísticos para la modelación de la volatilidad del rendimiento de un activo. A este tipo de modelos se les conoce como modelos con heteroscedasticidad condicional.

La volatilidad es un factor muy importante para la valuación de opciones. A ésta la podemos ver como la desviación estándar condicional del rendimiento de un activo. Una de las aplicaciones al modelar la volatilidad es que da un enfoque para el cálculo del VaR en la administración de riesgos. También ayuda a la eficiencia en la estimación de parámetros y a la precisión en los intervalos de predicción.

Series Financieras

Para comenzar daremos una introducción mencionando las propiedades de las series financieras y la volatilidad para después enfocarnos a los modelos que nos servirán para el modelado de la volatilidad.

El modelado de series financieras no es un problema sencillo de resolver. Esta complejidad no sólo se debe a la gran variedad de series que se utilizan hoy en día (acciones, tipos de cambio, tasas de interés, etc), a la importancia con la que se observan las frecuencias en las observaciones o a la disposición que se tiene de la información de datos, se debe mayormente a las regularidades estadísticas que son comunes en un gran número de series financieras y difíciles de reproducir con modelos estocásticos. Estas regularidades estadísticas se introdujeron por primera vez en Mandelbrot (1963) y desde entonces diversos autores han aportado información nueva.

A continuación daremos algunas propiedades bien conocidas en la literatura financiera.

- No estacionariedad en el precio. Las trayectorias de los precios generalmente son parecidas a caminatas aleatorias sin intercepto. Sin embargo, las trayectorias de los rendimientos generalmente son compatibles con los supuestos de los procesos estacionarios de segundo orden.
- En la variación de los rendimientos no hay autocorrelación. Estas variaciones suelen ser muy pequeñas e incluso parecidas a ruido blanco.
- Autocorrelación del cuadrado de los rendimientos (r_t^2). Estos rendimientos suelen estar altamente correlacionados.
- Agrupamiento de volatilidad. Periodos con altas perturbaciones les siguen periodos tranquilos (alta volatilidad seguida de baja volatilidad). Estos periodos suelen ser recurrentes pero no aparecen por periodos, esto puede contradecir a la estacionariedad.
- Colas pesadas. Diferentes estudios indican que la distribución de los rendimientos no es normal. Esta suele tener colas más pesadas y diferente kurtosis, a este efecto se le conoce como leptocúrtica y se mide con el coeficiente de curtosis. Asintóticamente este es 3 para muestras independientes e idénticamente distribuidas (iid) normales estándar pero este resulta ser mucho mayor a 3.
- Efecto de apalancamiento. Black (1976) fue quien lo notó, consiste en la asimetría del impacto de valores pasados tanto positivos como negativos que afectan a la volatilidad actual.
- Temporalidad. Hay que tener en cuenta el efecto de las fechas o el calendario, por ejemplo, días feriados, vacaciones, etc. ya que pueden tener efectos significativos en los rendimientos.

Características de la Volatilidad

Aunque la volatilidad no es observable, ésta tiene ciertas características:

- Puede ser alta o baja por periodos (agrupamiento de volatilidad).

- Los saltos son impredecibles.
- No diverge a infinito.
- Apalancamiento (diferente reacción a precios altos que a bajos).

3.1. ARCH

Los rendimientos son prácticamente impredecibles, tienen una gran cantidad de valores extremos y tanto los periodos de más agitación como los más tranquilos están agrupados en el tiempo. A estas características se les suele conocer como impredecibilidad, exceso de curtosis o colas pesadas y agrupamiento de la volatilidad. Un proceso ARCH produce patrones dinámicos de vuelta a la media que se pueden predecir, también produce más valores extremos de lo que se espera de una distribución normal estándar ya que estos valores durante el periodo de alta volatilidad son mayores de los que se pueden esperar con un proceso de volatilidad constante.

Para presentar este tema de una forma clara y no complicada seguiremos la dinámica que hemos venido haciendo a lo largo de este trabajo presentando un ARCH(1).

ARCH(1)

Un ARCH(1) está representado como,

$$\begin{aligned} a_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2, \end{aligned} \quad (3.1)$$

donde $\alpha_0 > 0$ y $\alpha_i \geq 0, \forall i > 0$ y ϵ_t es una variable aleatoria iid con media cero y varianza 1, que comunmente se distribuye como una normal estándar o una t de Student, en este caso supondremos que es Normal.

Los modelos ARCH nacieron para poder interpretar los ciclos económicos que se pensaba eran debido al carácter impredecible de la inflación. Friedman(1977) planteó la hipótesis de que el problema no era el nivel de inflación por sí mismo, sino que la incertidumbre sobre los costes y los precios futuros eran lo que desanimaba a los inversionistas y llevaba a una recesión. Ésto solo era posible si la incertidumbre cambiaba con el tiempo, a

esto se le conoce como heteroscedasticidad. En otras palabras, que la varianza no sea constante.

En esencia el modelo ARCH describe la predicción de la varianza en función de variables actuales observables, si el residual del rendimiento es grande, nuestra predicción para el siguiente periodo de volatilidad condicional también será grande. El modelo ARCH propone usar medias ponderadas de los cuadrados de los errores de predicción del pasado, esto se puede ver como una varianza ponderada.

Para ver al modelo ARCH(1), (3.1), como un AR(1) solo hay que tomar a_t^2 y restarlo a σ_t^2 . Dado esto, con $0 < \alpha < 1$ y con la propiedad de causalidad para un ARMA entonces, a_t^2 es un proceso AR(1) causal con la función de autocorrelación dada por $\rho_{a^2} = \alpha_1^h$, véase [2], propiedad que ya habíamos visto en el capítulo anterior.

3.1.1. Propiedades

Sea G_{t-1} la información que se tiene al tiempo $t - 1$, es decir, $G_u = \{r_u, r_{u-1}, \dots, r_0\}$. De la representación de un ARCH(1), (3.1), se tiene lo siguiente,

- $\mathbb{E}(a_t) = \mathbb{E}[\mathbb{E}(a_t|G_{t-1})] = \mathbb{E}[\sigma_t \mathbb{E}(\epsilon_t)] = 0.$
- $Cov(a_{t+h}, a_t) = \mathbb{E}(a_{t+h}a_t) = 0$
- $Var(a_t) = \alpha_0 / (1 - \alpha_1), 0 \leq \alpha_1 \leq 1.$
- $\mathbb{E}(a_t^4) = m_4 = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}.$

De los últimos dos puntos se puede obtener el valor estacionario de la curtosis

$$\frac{\mathbb{E}(a_t^4)}{(Var(a_t))^2} = 3 \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2},$$

donde $0 < \alpha_1 < 1/\sqrt{3}$, el cual es mayor o igual a 3, lo que significa que la densidad de las variables a_t tiene colas más pesadas que la distribución normal cuyo valor es 3.

3.1.2. Desventajas

- La suposición de que el cambio positivo y negativo en la información al tiempo t afecta de misma manera a la volatilidad. Sin embargo, se ha comprobado que en la práctica el precio de un activo financiero responde de manera diferente.
- El modelo no brinda nueva interpretación a la variación de una serie financiera. Este solo nos da una forma de describir el comportamiento de la dinámica de la varianza condicional.
- Suele sobrepredecir la volatilidad,

véase [11] o [5].

Identificar un ARCH

Como mostramos anteriormente, un modelo ARCH(1) se puede ver como un AR(1) y entonces la ACF de a_t^2 es la misma ACF que para un AR. Sin embargo, como se mostro el capítulo anterior, la PACF es una mejor forma de determinar el orden de un AR, por lo que la PACF nos ayudará obtener el orden de un ARCH.

3.1.3. Efecto ARCH

Al elevar al cuadrado a a_t , nos ayudará para revisar la heteroscedasticidad condicional, a lo que se le conoce como efecto ARCH.

Para poder comprobar los efectos ARCH existen dos pruebas o tests. El primero se refiere a la estadística Q de Ljung-Box aplicada a a_t^2 de un ARMA(p, q). La hipótesis nula para ésta estadística es que los primeros m lag de la ACF de a_t^2 son cero. $Q(m)$ es asintóticamente una Ji-cuadrada con $m - p - q$ grados de libertad y ésta es equivalente al test de los multiplicadores de Lagrange. Ésta estadística la presentaremos en la sección 3.3. La segunda prueba para la heteroscedasticidad condicional es la prueba de los multiplicadores de Lagrange. Esta prueba es la equivalente a la estadística F para la regresión lineal,

$$a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \dots + \alpha_m a_{t-m}^2 + e_t, \quad t = m + 1, \dots, T,$$

donde e_t es el error, que prueba, bajo la hipótesis nula, que los coeficientes son cero, i.e., $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$. La estadística F está dada por

$$F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T - 2m - 1)},$$

donde $SSR_0 = \sum_{t=m+1}^T (a_t^2 - \bar{a}_t^2)^2$ y $SSR_1 = \sum_{t=m+1}^T \hat{e}_t^2$ donde \hat{e}_t son los residuales de la regresión lineal. Ésta F, bajo la hipótesis nula, se distribuye asintóticamente como un Ji-cuadrada con m grados de libertad. La regla de decisión será rechazar H_0 si $F > \chi_m^2(\alpha)$.

3.2. GARCH

Como resultado del éxito de los modelos ARCH y sus modificaciones en la descripción de algunas características “no comunes” de los log-rendimientos, como colas pesadas, Bollerslev (1986) generalizó el proceso ARCH (GARCH) como una forma análoga a la generalización de los procesos AR a los ARMA.

Este modelo resulta muy útil para describir la dinámica de volatilidad de casi todas las series de rendimientos financieros. Este sirve para las acciones negociadas en la mayoría de los mercados desarrollados, la mayoría de las acciones que se negocian en mercados emergentes y para la mayoría de los índices de los mercados de valores. También se utiliza en tipos de cambio, rendimientos de los commodities o bienes primarios y en el rendimiento de obligaciones.

3.2.1. Propiedades

El modelo GARCH(p,q) del proceso a_t es

$$\begin{aligned} a_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \end{aligned} \quad (3.2)$$

donde $\{\epsilon_t\}$ es una sucesión de variables aleatorias iid con media cero y varianza 1. $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$.

Una de las principales ventajas de los modelos GARCH sobre los ARCH es que los ordenes de los procesos GARCH(p, q) no necesitan ser grandes como en ocasiones es el orden p en el caso del proceso ARCH(p), en donde p frecuentemente está muy por encima de p y q .

El modelo GARCH para a_t^2 se puede ver como un ARMA si hacemos $\eta_t = a_t^2 - \sigma_t^2$. Bajo esta observación se tiene

$$\mathbb{E}(a_t^2) = \sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i)},$$

véase [11], donde $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$.

GARCH(1,1)

La representación de un GARCH está dada por una ponderación de tres predicciones diferentes de la varianza. Una de ellas es una varianza constante que corresponde a la media de largo plazo, la segunda es la predicción que se realizó en el periodo anterior y la tercera está relacionada con la nueva información que no estaba disponible cuando se hizo la segunda predicción.

Se ha observado que el GARCH(1,1) es un buen modelo para muchas series financieras, por esto vamos a analizar sus características.

El GARCH(1,1) está dado por:

$$\begin{aligned} a_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \end{aligned} \quad (3.3)$$

Para valores muy grandes de a_{t-1}^2 se tiene valores grandes para σ_t^2 dando lugar al agrupamiento de volatilidad en las series financieras. Para el proceso GARCH(1,1) se puede ver que la curtosis está dada por

$$\frac{\mathbb{E}(a_t^4)}{[\mathbb{E}(a_t^2)]^2} = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - \beta_1^2 - 2\alpha_1\beta_1 - 3\alpha_1^2} > 3,$$

donde $\alpha_1 + \beta_1 < 1$, que al igual que el modelo ARCH, el GARCH tiene colas más pesada que la normal.

La predicción de la volatilidad en l pasos está dada por

$$\sigma_h^2(l) = \frac{\alpha_0[1 - (\alpha_1 + \beta_1)^{l-1}]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{l-1} \sigma_h^2(1), \quad (3.4)$$

donde $\sigma_h^2(1)$ está dado por $\alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2$.

Se puede demostrar que, si se hace tender l a infinito, $\sigma_h^2(l)$ converge a la varianza de a_t .

Anteriormente mencionamos que se puede ver que si $\{r_t\}$ es un proceso GARCH, entonces $\{r_t^2\}$ es un proceso ARMA, y para poder calcular el orden de un proceso GARCH esta propiedad nos será de ayuda. A pesar de esta observación es de gran utilidad, la obtención del orden para un modelo GARCH no es sencilla. En muchas aplicaciones se suelen utilizar modelos con orden bajo como el GARCH(1,1), GARCH(2,1) y GARCH(1,2).

Gracias a la observación antes mencionada podemos asegurar que las ACF's serán positivas.

Si r_t es un proceso GARCH(p,q) y el cuarto momento existe, entonces

$$\gamma_{r_t^2}(h) \geq 0, \forall h.$$

Más aún, si $\alpha_1 > 0$, entonces

$$\gamma_{r_t^2}(h) > 0 \forall h.$$

Para un GARCH(1,1) se tiene que $\gamma_{r_t^2}(h)$ está dada por

$$\gamma_{r_t^2}(h) = \rho_{r_t^2}(1)(\alpha_1 + \beta_1)^{h-1}, \quad h \geq 1, \quad (3.5)$$

donde

$$\rho_{r_t^2}(1) = \frac{\alpha_1[1 - \beta_1(\alpha_1 + \beta_1)]}{1 - (\alpha_1 + \beta_1)^2 + \alpha_1^2} \quad (3.6)$$

Analizaremos un ejemplo con los rendimientos del NYSE (New York Stock Exchange) ajustándole un modelo GARCH(1,1). Al analizar las gráficas de la figura 3.2 podemos suponer que la distribución de los residuales no esta lejos de ser Normal.

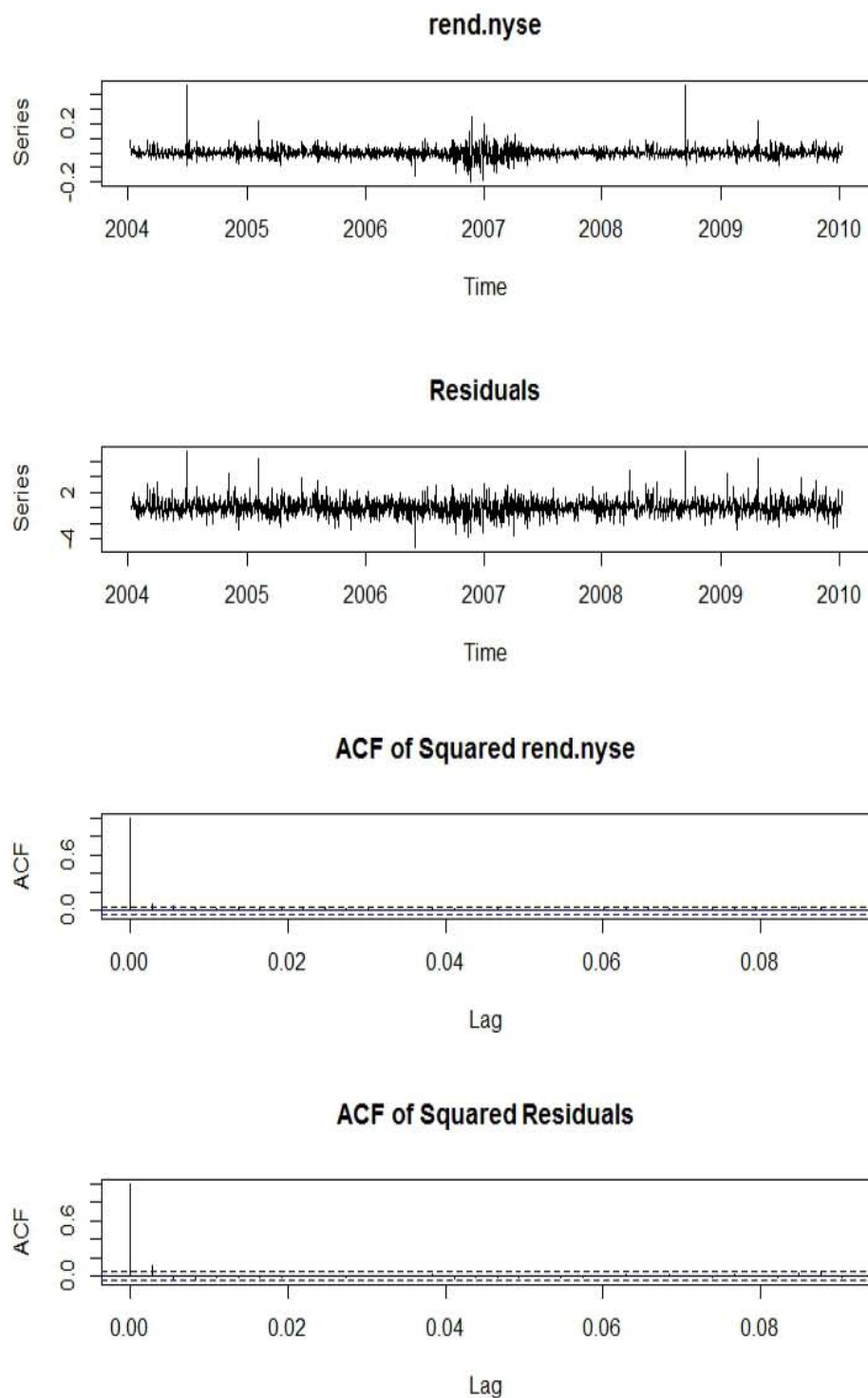


Figura 3.1: Log-rendimientos ACF de la serie NYSE y sus residuales en una ventana de 6 años(2004-2010).

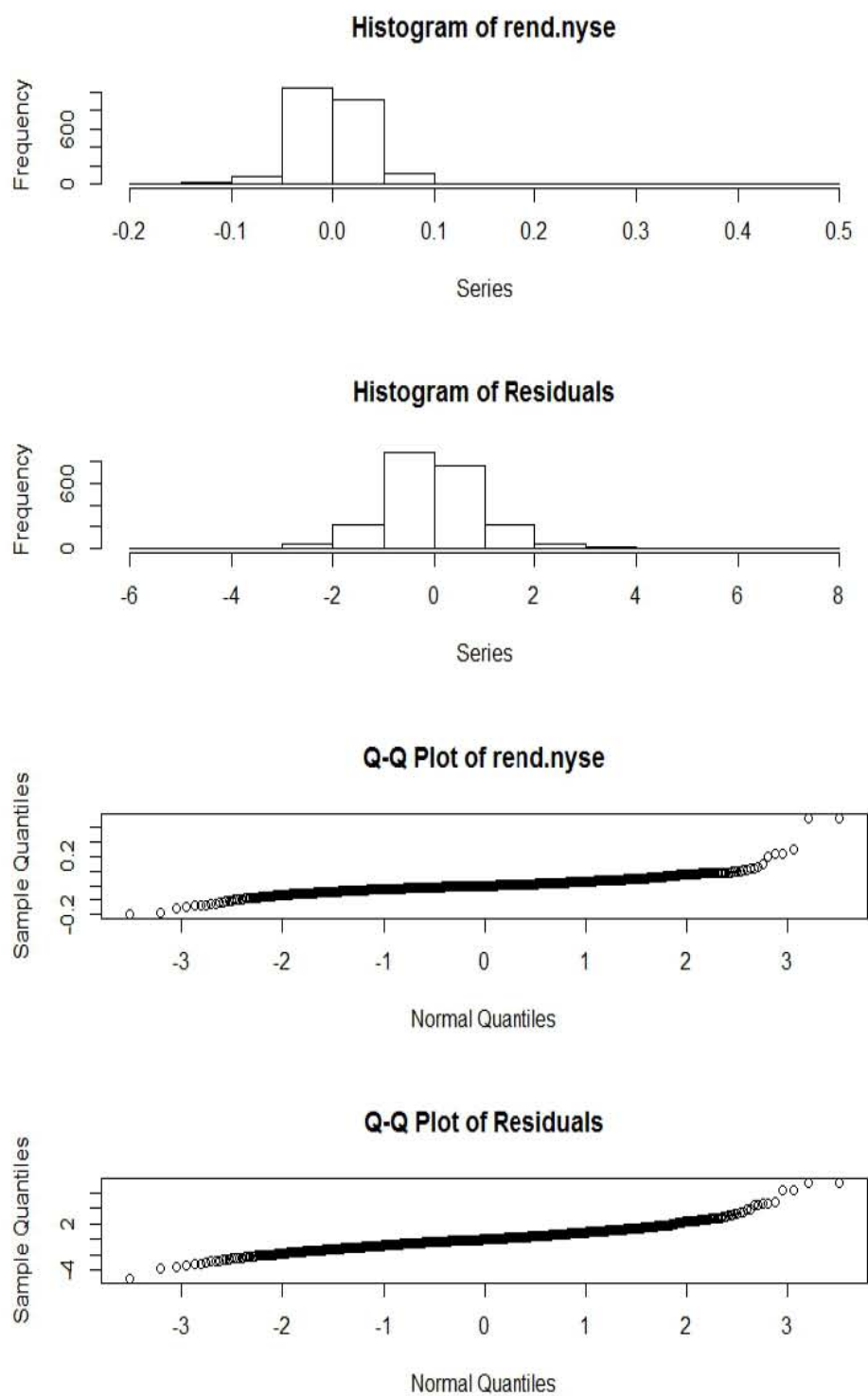


Figura 3.2: Histograma y QQ-plot de la serie NYSE y de los residuales

Para ver las predicciones del modelo GARCH aplicadas al ejemplo anterior, las calculamos como $\text{lag} = 1$, es decir, la ventana de tiempo fue de un día antes que se supiera el valor. Los resultados para los límites inferior y superior son $1.96 \pm \hat{\sigma}_t$ que son las líneas punteadas en la figura 3.3. Para poder ver más a detalle estas bandas se hicieron dos gráficas, la primera muestra el “zoom” de la ventana completa a diferencia de la segunda que es toda la ventana de tiempo.

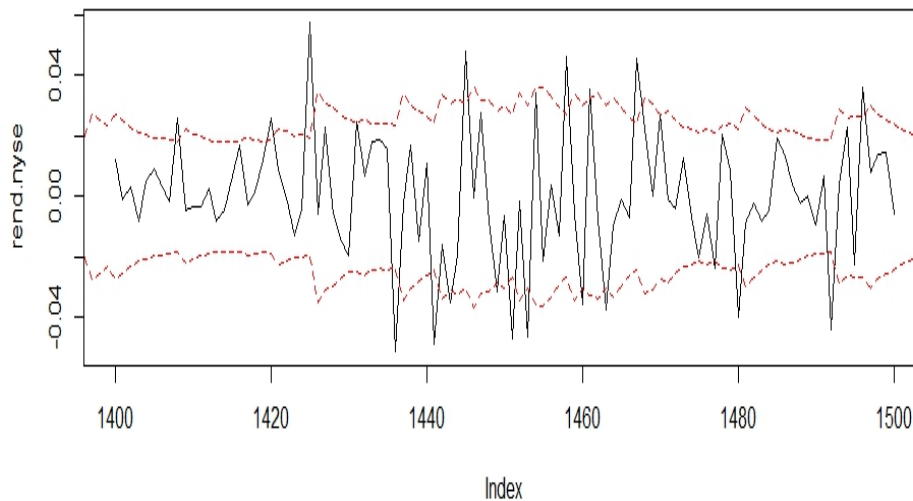


Figura 3.3: Predicción del modelo GARCH(1,1) ajustado a la serie NYSE en una ventana de tiempo de los últimos 100 datos.

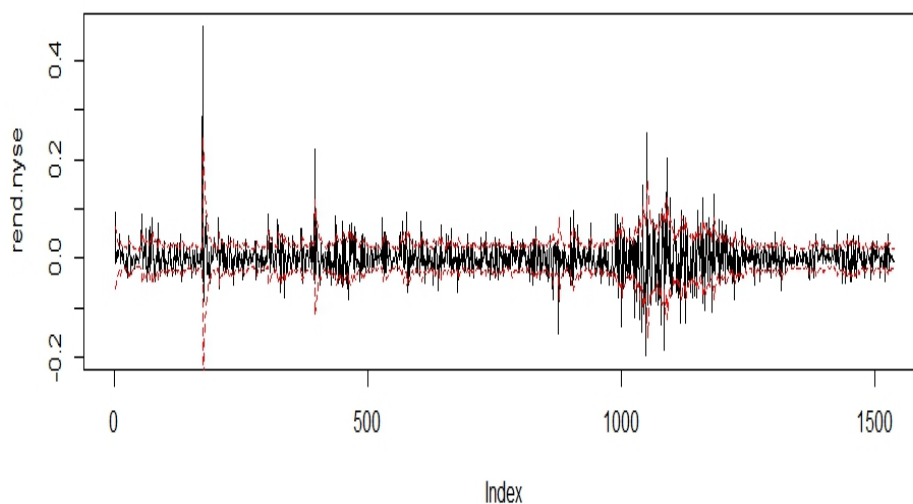


Figura 3.4: Predicción del modelo GARCH(1,1) ajustado a la serie NYSE.

El modelo GARCH también tiene defectos y por ello se han propuesto varias generalizaciones para sobrellevar estas desventajas. Por ejemplo, el modelo exponencial GARCH (EGARCH) se puede utilizar cuando los rendimientos son asimétricos. Para el caso en que la volatilidad sufre cambios accidentales o eventuales se utiliza el modelo GARCH integrado (IGARCH). Así como este tipo de modelos existen otros pero para este trabajo solo nos enfocaremos en los GARCH.

3.3. Contrastes de Estructura y Bondad de Ajuste

Box y Pierce, Portmanteau Test

La aplicación de estos modelos requiere realizar una comparación entre la correlación de r_t y cero. Box y Pierce (1970) propusieron la estadística de

Portmanteau

$$Q^*(m) = T \sum_{i=1}^m \hat{\rho}_i^2$$

para contrastar las hipótesis $H_0 : \rho_1 = \dots = \rho_m = 0$ vs $H_a : \rho_i \neq 0$ para alguna $i \in \{1, \dots, m\}$. Se puede demostrar que ésta estadística se distribuye como una Ji-cuadrada con m grados de libertad.

Ljung y Box

Ljung y Box (1978) modificaron la estadística de Box y Pierce para incrementar la prueba en muestras finitas como

$$Q(m) = T(T+2) \sum_{i=1}^m \frac{\hat{\rho}_i^2(a_i^2)}{T-i}.$$

La regla de decisión para esta estadística es rechazar H_0 si $Q(m) > \chi_{(m, 100(1-\alpha)\%)}^2$. La elección de m en base a estudios realizados es $m \approx \ln(T)$, sin embargo, no hay una decisión que sea la correcta aún. Como dijimos en la sección 3.1 la Q nos ayudará a checar heteroscedasticidad.

Jarque-Bera

Jarque y Bera (1987) propusieron una estadística para probar normalidad en los residuos, r_t , con la hipótesis nula como curtosis y sesgo igual a cero. La estadística propuesta es

$$JB = \frac{\hat{S}^2(r)}{6/T} + \frac{(\hat{K}(r) - 3)^2}{24/T}$$

que se distribuye asintóticamente como una Ji-cuadrada con 2 grados de libertad.

Criterio de Akaike (AIC)

Akaike (1973) definido como

$$AIC = \frac{-2}{T} \ln(\text{verosimilitud}) + \frac{2}{T}(p).$$

El primer término del AIC corresponde a la bondad de ajuste para un AR(p) y el segundo ajusta el modelo de acuerdo a cuántos parámetros se usaron. Simulaciones bajo el método de Monte Carlos sugiere que para el ajuste de modelos autorregresivos, el AIC tiende a sobre estimar a p.

Este criterio nos ayudará a la decisión entre uno o más modelos propuestos. Nos servirá para rechazar un modelo con más variables explicativas que otro cuando la incorporación de más variables no haga que el error disminuya. En general, siempre se tomará el modelo que tenga el menor AIC.

BIC

Este criterio trata de corregir la sobre estimación del AIC. Para un proceso ARMA(p,q) con media cero, causal e invertible se define el BIC como

$$BIC = (T - p - q) \ln \left[\frac{T\hat{\sigma}^2}{T - p - q} \right] + T(1 + \ln \sqrt{2\pi}) + \\ + (p + q) \ln \left[\left(\sum_{i=1}^T r_i^2 - T\hat{\sigma}_\epsilon^2 \right) / (p + q) \right],$$

donde $\hat{\sigma}_\epsilon^2$ es la estimación máximo verosímil del ruido blanco.

Una propiedad del BIC es que, si las observaciones en verdad son ARMA(p,q) y los órdenes estimados minimizando el BIC, entonces $\hat{p} \rightarrow p$ y $\hat{q} \rightarrow q$ cuando $n \rightarrow \infty$. Esta propiedad la cumple el AIC pero solo para procesos AR.

Capítulo 4

Aplicación

A continuación, para ejemplificar el uso de los modelos ARCH usaremos el índice Standard & Poor's (S&P500) y le ajustaremos un modelo GARCH verificando sus supuestos y haciendo una predicción del mismo. Todo esto se hará utilizando el programa estadístico R.

Se escogió éste índice porque es uno de los más importantes en el mercado de valores de Estados Unidos y se le considera el más representativo de la situación real del mercado, incluso expresa mejor al mercado accionario y de bonos que el Dow Jones y el Nasdaq. Se examinarán los precios diarios al cierre desde enero de 1980 a mayo del 2010, así se tendrá una ventana de tiempo bastante amplia, de 30 años, para poder ver algunas características como las crisis financieras. En la figura 4.1 vemos la historia del S&P500 y es fácil ver algunos puntos importantes como las crisis de 1987, 2001 y 2008.

Siguiendo con un análisis puramente gráfico, en la figura 4.2 podemos ver solo a los log-rendimientos de la serie en estudio. Lo que podemos observar son las "agrupaciones" que se tienen, esto se debe a que grandes (pequeños) cambios que son seguidos por cambios de misma magnitud o parecida. A este efecto se le conoce como agrupamiento de la volatilidad y es justo uno de los propósitos de los modelos tipo ARCH que son los encargados de cuantificar este efecto.

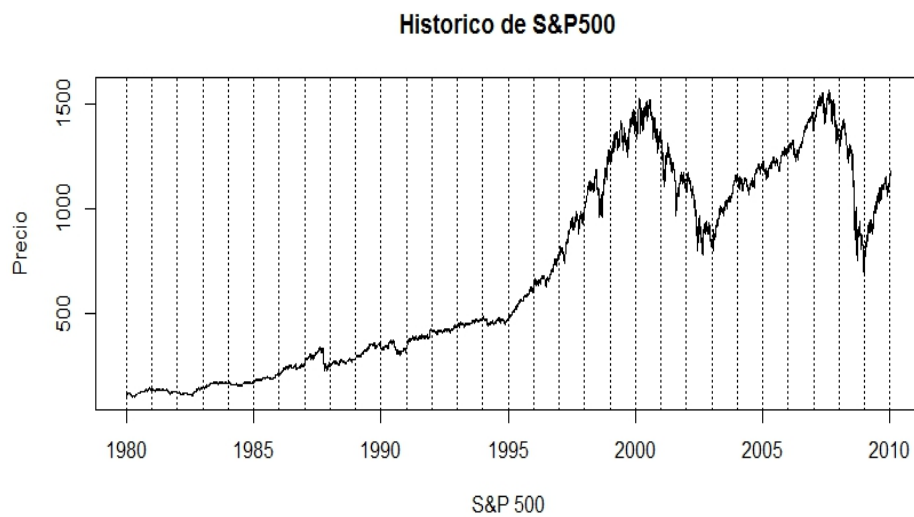


Figura 4.1: Histórico de S&P 500

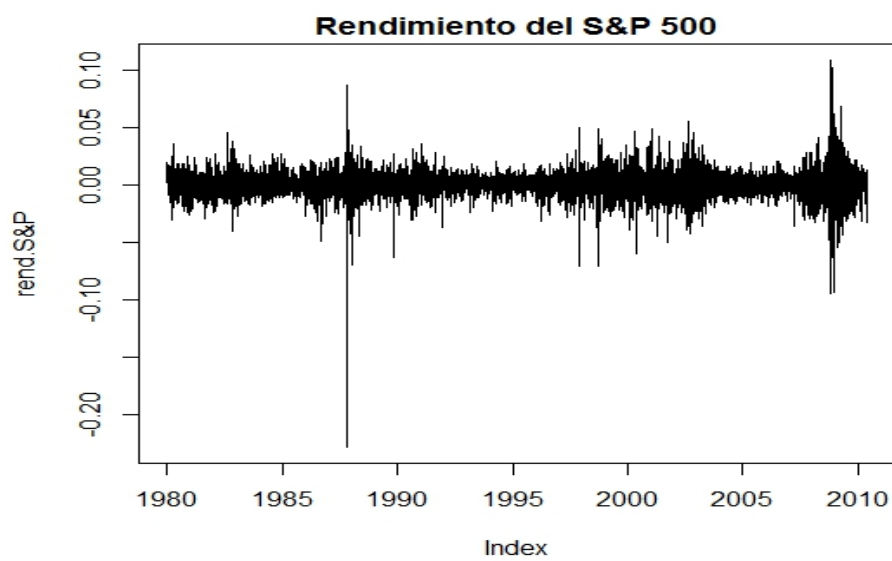


Figura 4.2: Rendimiento del S&P 500

Podemos empezar haciendo un pequeño análisis “económico” fijándonos en los rendimientos y el histórico de los precios, esto en la figura 4.3. De las primeras cosas a observar es el crecimiento del precio hasta el comienzo del nuevo milenio. Estos rendimientos están centrados en cero aunque a veces tienen variaciones más fuertes que otras, por ejemplo, hay dos periodos que sobresalen, uno es debido a la crisis de 1987 conocida como el “crash de 1987” o “lunes negro” y el otro periodo es la última crisis del 2008 conocida como “la crisis de las hipotecas subprime”. En la figura de la comparación del rendimiento contra el precio podemos observar que la volatilidad es alta cuando los precios bajan o van en caída, a esto se le conoce como mercado bull, con tendencia a la baja. Aquí la volatilidad tiende a ser superior. Este efecto (volatilidad asimétrica), como mencionamos antes, es modelado con el EGARCH.

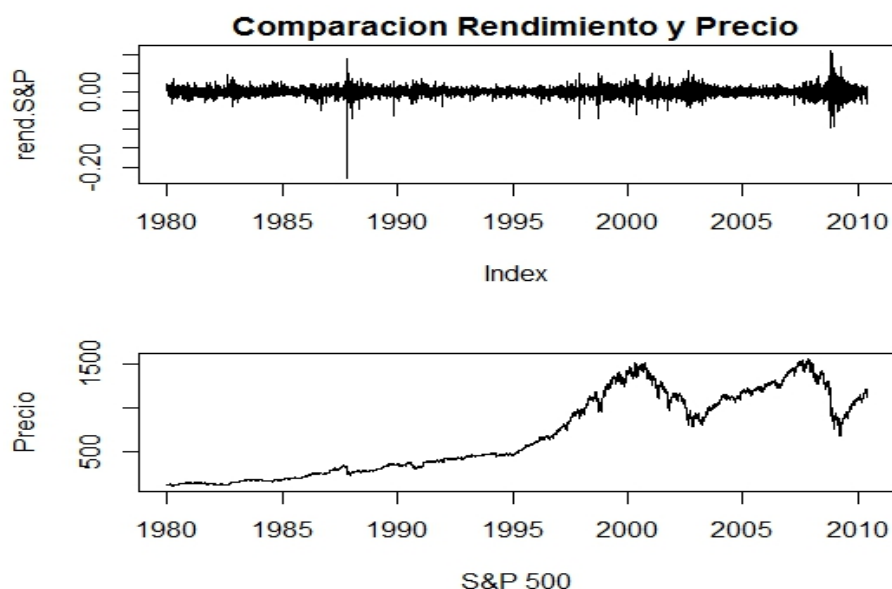


Figura 4.3: Comparativo entre el rendimiento y el precio del S&P 500

Veamos que podemos analizar esta serie con modelos tipo ARCH. En la figura que se muestra a continuación tenemos la ACF y PACF de los log-rendimientos de la serie original y los log-rendimientos al cuadrado respectivamente. Las correlaciones nos serán de mucha utilidad ya que las que sean significativas nos pueden dar una señal de que la serie es predecible.

ble mientras que las autocorrelaciones significativas de los rendimientos al cuadrado presentan la existencia de agrupamiento de volatilidad. En la gráfica a) de la figura 4.4 observamos que hay correlaciones que son significativas al 95 % (línea azul), aunque al 99 % (línea roja) existe menos significancia de las correlaciones, mientras que en la gráfica b) hay observaciones muy significativas, esto indica que los cambios en el precio no son independientes e idénticamente distribuibles y que hay presencia de efectos ARCH.

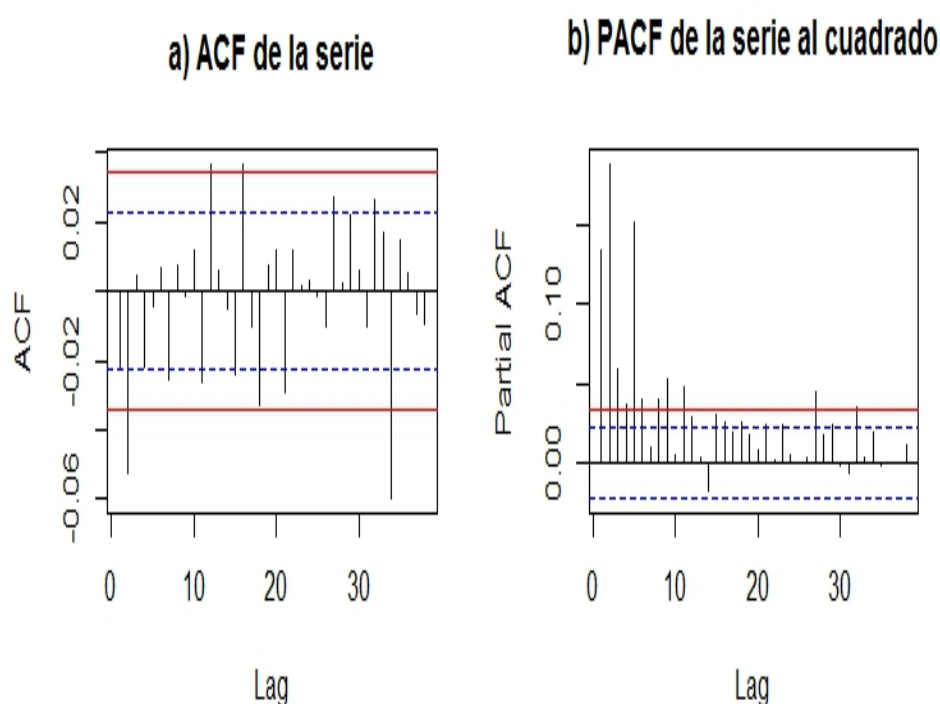


Figura 4.4: a) ACF de los log-rendimientos de la serie original. b) PACF de los log-rendimientos al cuadrado.

A pesar de que las gráficas siempre nos ayudarán de forma visual a ver ciertos aspectos del modelo para hacer inferencia existen pruebas formales que realizaremos más adelante.

Siguiendo con el análisis para revisar qué modelo GARCH le podemos ajustar a los datos, recordemos que, tomando los residuales al cuadrado de un GARCH(p,q), tenemos un ARMA(p,q). Entonces podemos aplicar el método antes mencionado, sección 2.4.1 de la EACF para modelos ARMA.

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1	x	x	o	x	x	x	o	o	o	o	x	o	o	o
2	x	o	o	x	x	o	o	o	o	o	x	o	o	o
3	x	o	x	o	x	x	o	o	o	o	x	o	o	x
4	x	x	x	x	x	o	o	o	o	o	o	o	o	x
5	x	x	x	x	x	o	o	o	o	o	o	o	o	x
6	x	x	x	x	o	x	x	o	o	o	o	o	o	x
7	x	x	o	o	o	x	x	x	o	o	o	o	o	x

Figura 4.5: EACF de los residuales al cuadrado

La figura 4.5 nos sugiere que podemos ajustar un GARCH(1,2), GARCH(2,1) y GARCH(2,2) a los datos del S&P 500.

Siguiendo con el análisis en la siguiente tabla, 4.1, mostraremos algunas estadísticas con respecto a lo que se ha hablado anteriormente.

A continuación los histogramas de los residuales para ver que en efecto estos presentan picos (leptocúrtica) y colas más pesadas:

Estadística	Rendimiento	GARCH(1,2)	GARCH(2,1)	GARCH(2,2)	Residuales
Media	0.00031	0.03411	0.033999	0.03448	
Desviación estándar	0.01146	0.99103	0.99927	0.99770	
Asimetría	-1.24727	-0.55262	-0.57171	-0.56941	
Curtosis	28.81395	4.42513	4.62765	4.64705	

Tabla 4.1: Tabla de estadísticas de rendimientos y residuales del S&P 500.

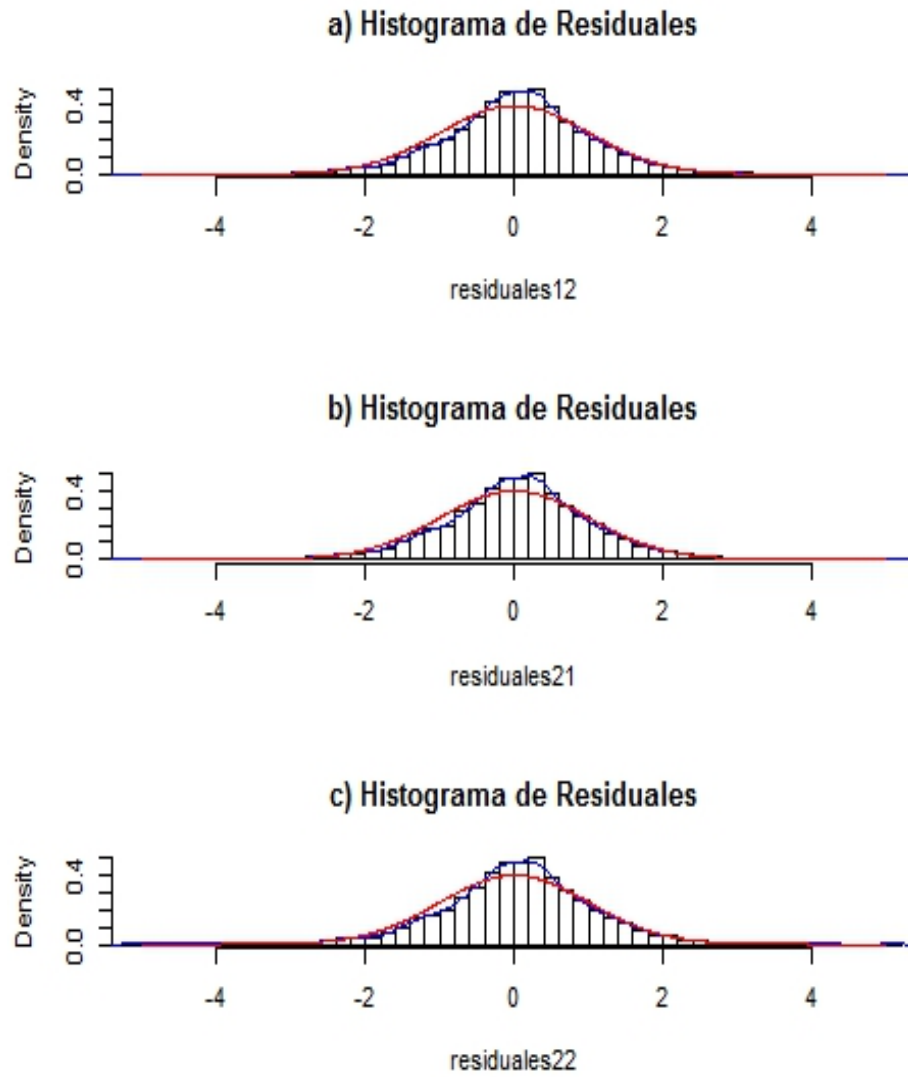


Figura 4.6: Histograma de los residuales. a) de un GARCH(1,2), b) GARCH(2,1) y c) GARCH(2,2).

En base a la tabla 4.1 no podemos suponer normalidad y antes de hacer una prueba formal hay más cuestiones que se puede analizar gráficamente. En seguida se muestran las gráficas de qqplot, los cuantiles de la muestra contra los de una normal, tomando en cuenta a la muestra como los residuales de los modelos ajustados. En los tres ajustes podemos ver que en el centro se ajustan bien a una línea recta, sin embargo, en los extremos no pasa esto lo que reafirma el hecho de colas pesadas, en este caso cola pesada izquierda para los tres modelos. Donde a), b) y c) son los cuantiles de los residuales de un GARCH(1,2), GARCH(2,1) y GARCH(2,2) respectivamente.

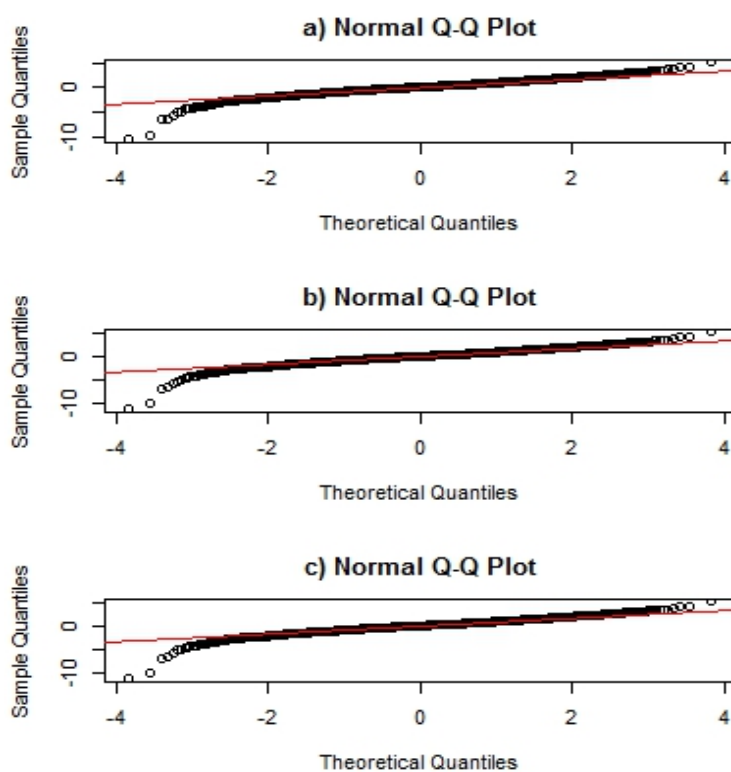


Figura 4.7: Cuantiles de los rendimientos del S&P 500 vs los de una normal.

En las gráficas de la función de autocorrelación simple (acf) de la figura 4.9 observamos que existe una correlación poco significativa para los residuales de la serie, mientras que los residuales al cuadrado ya se comportan como ruido blanco, i.e., son no correlacionados. En cambio la correlación es muy pobre entre los rendimientos de la serie, sin embargo, existe correlación entre los rendimientos al cuadrado. Esta correlación suele ser casi siempre positiva, con valores que no son muy grandes, pero estadísticamente significativos y que decrecen de forma lenta hacia cero. Esto es debido al agrupamiento que se produce en la volatilidad (lo que indica la presencia de heteroscedasticidad condicional en las series de rendimientos).

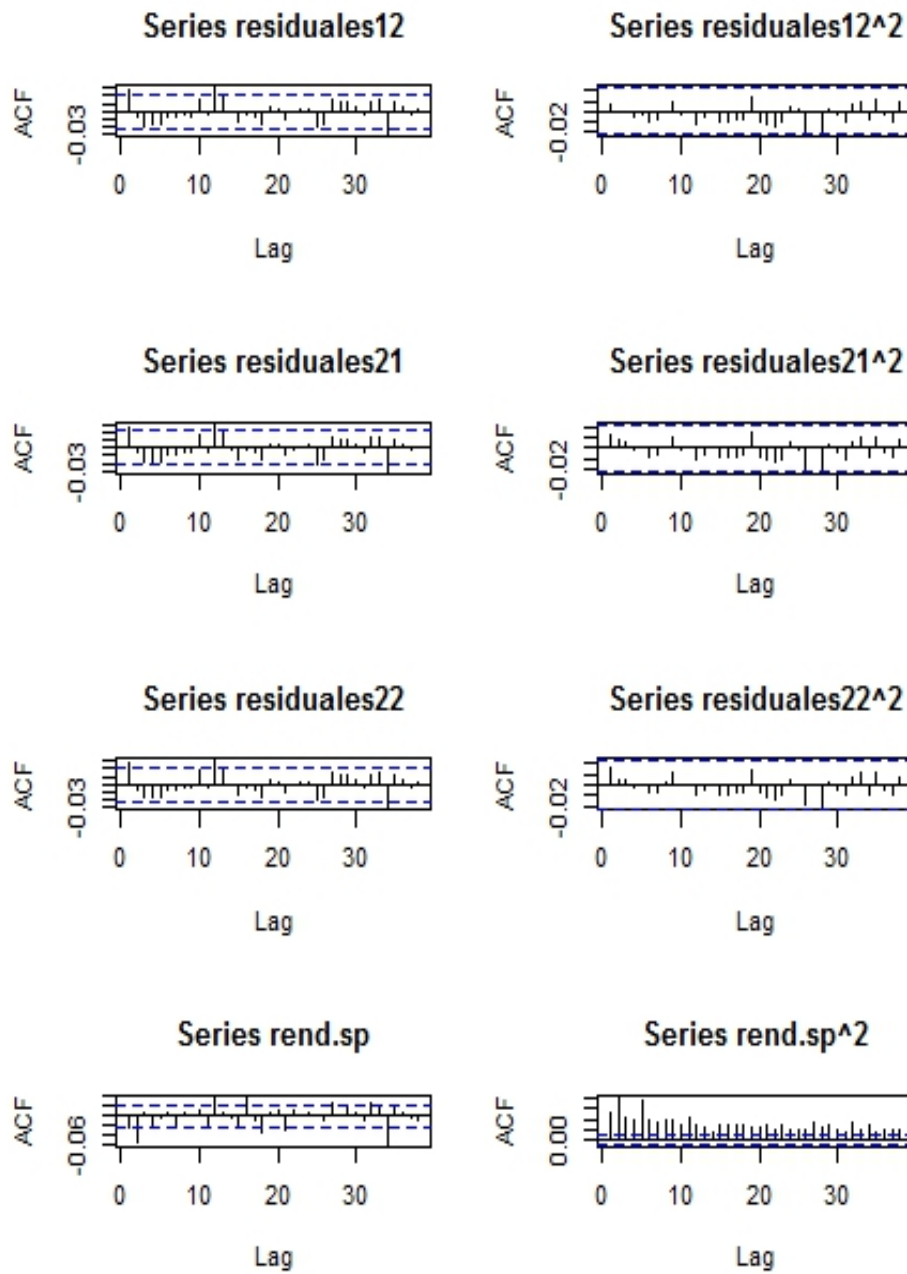


Figura 4.8: ACF de los residuales para los diferentes ajustes y log-rendimientos del S&P 500.

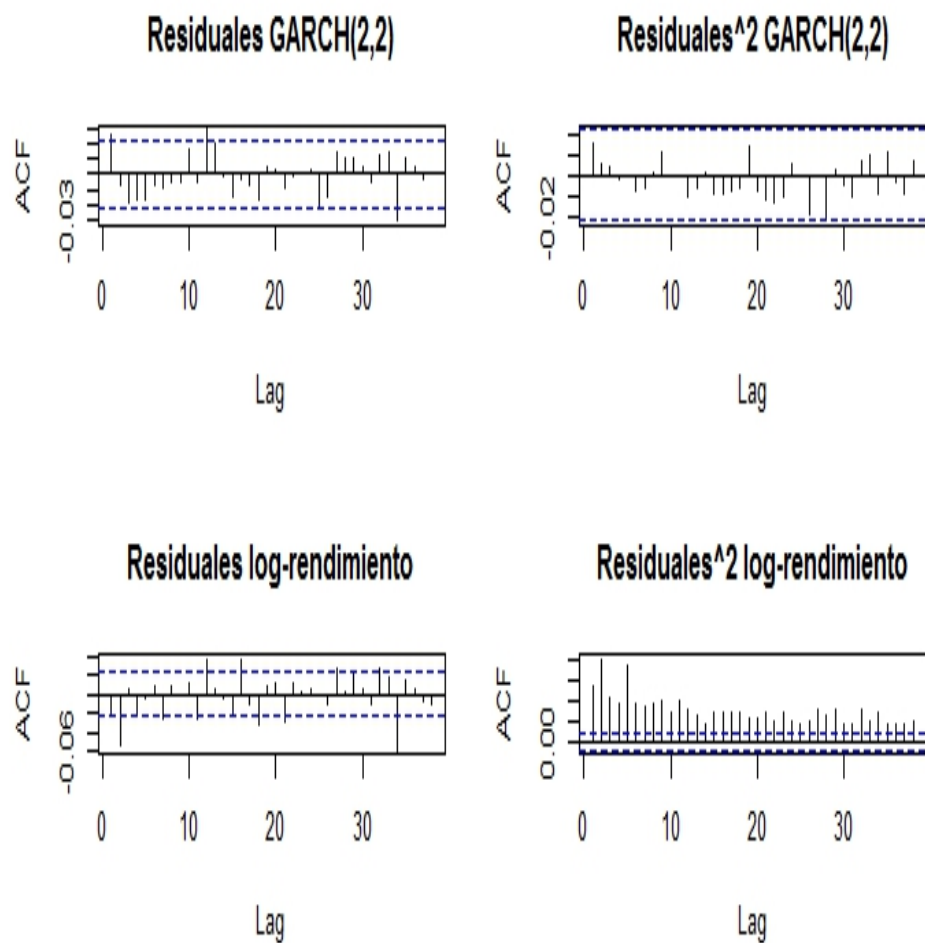


Figura 4.9: ACF de los residuales para los diferentes ajustes y log-rendimientos del S&P 500.

A continuación presentamos una tabla para los modelos ajustados en donde veremos sus estimaciones. Además se realizarán dos pruebas formales: La Jarque-Bera y la prueba de Box-Ljung. Como mencionamos en el capítulo anterior la Jarque-Bera sirve para probar normalidad. También se puede realizar la prueba de Shapiro-Wilk, que es más potente, para ver normalidad en los residuos. La prueba Q de Box-Ljung es para revisar la no correlación en los residuales y la presencia de heteroscedasticidad condicional.

Parámetro	GARCH(1,2)	GARCH(2,1)	GARCH(2,2)
α_0	1.336e-06	1.244e-06	1.336e-06
α_1	0.0817	0.0750	0.0817
α_2	NA	1.000e-08	1.000e-08
β_1	0.7873	0.9162	0.7873
β_2	0.1217	NA	0.1217

Tabla 4.2: Tabla de los parámetros estimados para los diferentes ajustes.

Veamos ahora las salidas presentadas en las figuras (4.10-4.12). Analizaremos solo la salida 4.10.

Podemos observar que todos los coeficientes son significativos a excepción del β_1 . Además tenemos diferentes pruebas tanto para la normalidad como para la correlación y la heteroscedasticidad. Antes de seguir con el análisis hay que resaltar el "NA" que se tiene en la prueba de Shapiro-Wilk y el 0 en la Jarque-Bera. Como mencionamos anteriormente, todo el análisis se realizó en el paquete estadístico R y éste, para la prueba de Sharpio-Wilk, solo puede tomar cierto número de valores. Tomando los valores necesarios para realizar la prueba se vio que el p-value es demasiado chico, entonces para más valores R pone por default el "NA". El caso de la prueba Jarque-Bera es similar.

Continuando con el análisis, si tomamos un nivel de significancia al 5 % no rechazamos muchas pruebas a excepción de la de normalidad que es algo que ya habíamos hecho con los histogramas y el qqplot. Si nos fijamos en la prueba de Ljung-Box, para $m = 10, 15$ tenemos estadísticas muy grandes, por lo que se rechazaría la hipótesis nula. La prueba de Ljung-

Box con los residuales al cuadrado es equivalente a la McLeod-Li para probar heteroscedasticidad condicional. Esto se comprueba con la figura 4.13 en donde podemos observar claramente que no hay heteroscedasticidad condicional en los residuales. Esto hace que el modelo parezca adecuado.

```

Title:
  GARCH Modelling

Call:
  garchFit(formula = ~garch(1, 2), data = rend.sp, trace = F, gamma = 0)

Mean and Variance Equation:
  data ~ garch(1, 2)
<environment: 0x000000014206610>
 [data = rend.sp]

Conditional Distribution:
  norm

Coefficient(s):
      mu      omega      alpha1      beta1
5.3381e-04 1.3361e-06 8.1706e-02 7.8728e-01
      beta2
1.2166e-01

Std. Errors:
  based on Hessian

Error Analysis:
      Estimate Std. Error t value Pr(>|t|)
mu      5.338e-04 9.488e-05 5.626 1.85e-08 ***
omega   1.336e-06 2.393e-07 5.584 2.35e-08 ***
alpha1  8.171e-02 9.005e-03 9.073 < 2e-16 ***
beta1   7.873e-01 1.329e-01 5.922 3.18e-09 ***
beta2   1.217e-01 1.257e-01 0.968 0.333
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
 24823.94      normalized: 3.24284

Description:
  Thu May 23 14:41:14 2013 by user: Administrador

Standardised Residuals Tests:
      Statistic p-Value
Jarque-Bera Test  R      Chi^2 6875.601 0
Shapiro-Wilk Test R      W      NA      NA
Ljung-Box Test   R      Q(10) 17.2232 0.06956883
Ljung-Box Test   R      Q(15) 28.89723 0.01658177
Ljung-Box Test   R      Q(20) 31.86102 0.04480189
Ljung-Box Test   R^2    Q(10) 3.399977 0.9703859
Ljung-Box Test   R^2    Q(15) 5.213338 0.9901608
Ljung-Box Test   R^2    Q(20) 8.594364 0.9871471
LM Arch Test     R      TR^2 4.25144 0.9784599

Information Criterion Statistics:
      AIC      BIC      SIC      HQIC
-6.484374 -6.479839 -6.484375 -6.482818

```

Figura 4.10: Resumen del modelo GARCH(1,2) con sus pruebas de normalidad y heteroscedasticidad.

```

Title:
  GARCH Modelling

Call:
  garchFit(formula = ~garch(2, 1), data = rend.sp, trace = F)

Mean and Variance Equation:
  data ~ garch(2, 1)
<environment: 0x000000013cf6670>
  [data = rend.sp]

Conditional Distribution:
  norm

Coefficient(s):
      mu      omega      alpha1      alpha2
5.3421e-04 1.2445e-06 7.5081e-02 1.0000e-08
      beta1
9.1624e-01

Std. Errors:
  based on Hessian

Error Analysis:
      Estimate Std. Error t value Pr(>|t|)
mu      5.342e-04 9.477e-05  5.637 1.73e-08 ***
omega   1.244e-06 2.307e-07  5.394 6.90e-08 ***
alpha1  7.508e-02 1.055e-02  7.118 1.09e-12 ***
alpha2  1.000e-08 1.365e-02  0.000 1
beta1   9.162e-01 8.393e-03 109.169 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
 24823.5      normalized: 3.242783

Description:
  Thu May 23 13:18:26 2013 by user: Aministrador

Standardised Residuals Tests:
      Statistic p-Value
Jarque-Bera Test  R      Chi^2  7201.639  0
Shapiro-Wilk Test  R      W      NA      NA
Ljung-Box Test     R      Q(10)  17.18603 0.0703477
Ljung-Box Test     R      Q(15)  28.81084 0.0170107
Ljung-Box Test     R      Q(20)  31.78126 0.04568562
Ljung-Box Test     R^2    Q(10)  3.865924 0.9531927
Ljung-Box Test     R^2    Q(15)  5.704878 0.9842901
Ljung-Box Test     R^2    Q(20)  9.171329 0.9808365
LM Arch Test       R      TR^2   4.733238 0.9663007

Information Criterion Statistics:
      AIC      BIC      SIC      HQIC
-6.484259 -6.479724 -6.484260 -6.482703

```

Figura 4.11: Resumen del modelo GARCH(2,1) con sus pruebas de normalidad y heteroscedasticidad.


```

Title:
  GARCH Modelling

Call:
  garchFit(formula = ~garch(2, 2), data = rend.sp, trace = F)

Mean and Variance Equation:
  data ~ garch(2, 2)
<environment: 0x0000000013b44f80>
 [data = rend.sp]

Conditional Distribution:
  norm

Coefficient(s):
      mu      omega      alpha1      alpha2
5.3380e-04 1.3361e-06 8.1706e-02 1.0000e-08
      beta1      beta2
7.8728e-01 1.2166e-01

Std. Errors:
  based on Hessian

Error Analysis:
      Estimate Std. Error t value Pr(>|t|)
mu      5.338e-04 9.493e-05 5.623 1.87e-08 ***
omega   1.336e-06 9.732e-07 1.373 0.170
alpha1  8.171e-02 1.655e-02 4.936 7.97e-07 ***
alpha2  1.000e-08 7.196e-02 0.000 1.000
beta1   7.873e-01 7.267e-01 1.083 0.279
beta2   1.217e-01 6.621e-01 0.184 0.854
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
24823.94      normalized: 3.24284

Description:
Thu May 23 13:17:14 2013 by user: Administrador

Standardised Residuals Tests:
      Statistic p-Value
Jarque-Bera Test  R      Chi^2 6875.599 0
Shapiro-Wilk Test R      W      NA      NA
Ljung-Box Test   R      Q(10) 17.22319 0.06956889
Ljung-Box Test   R      Q(15) 28.89723 0.01658178
Ljung-Box Test   R      Q(20) 31.86102 0.04480194
Ljung-Box Test   R^2    Q(10) 3.399972 0.9703861
Ljung-Box Test   R^2    Q(15) 5.213332 0.9901609
Ljung-Box Test   R^2    Q(20) 8.594359 0.9871472
LM Arch Test     R      TR^2 4.251435 0.97846

Information Criterion Statistics:
      AIC      BIC      SIC      HQIC
-6.484113 -6.478671 -6.484114 -6.482246

```

Figura 4.12: Resumen del modelo GARCH(2,2) con sus pruebas de normalidad y heteroscedasticidad.

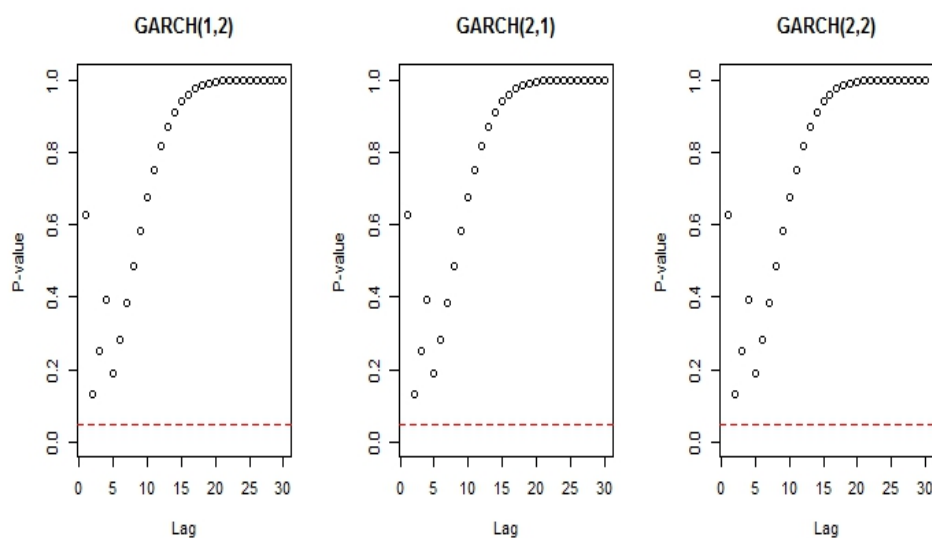


Figura 4.13: Prueba de McLeod-Li para los residuales al cuadrado.

Como se puede observar el AIC y el BIC de los 3 modelos no cambia mucho. Sin embargo, de los 3 modelos el más pequeño es el modelo 3 por lo que podemos escoger a este como el mejor modelo.

El ajuste del mejor modelo a la serie se presenta en la figura 4.14 el cual se ajusta bastante bien.

Entonces, para poder realizar la predicción de las varianzas condicionales para $2 \leq t \leq n$ de la serie en cuestión se debe de tomar en cuenta el valor inicial (σ_1^2), este lo podemos proponer como la varianza no condicional de un proceso estacionario dada por $\sigma^2 = \frac{\hat{\alpha}_0}{1 - \hat{\alpha}_1 - \hat{\beta}_1}$ y tomando los rendimientos al cuadrado podemos usar la ecuación (3.4). Luego, los intervalos de predicción de un paso adelante para la serie analizada están dados en la figura 4.15.

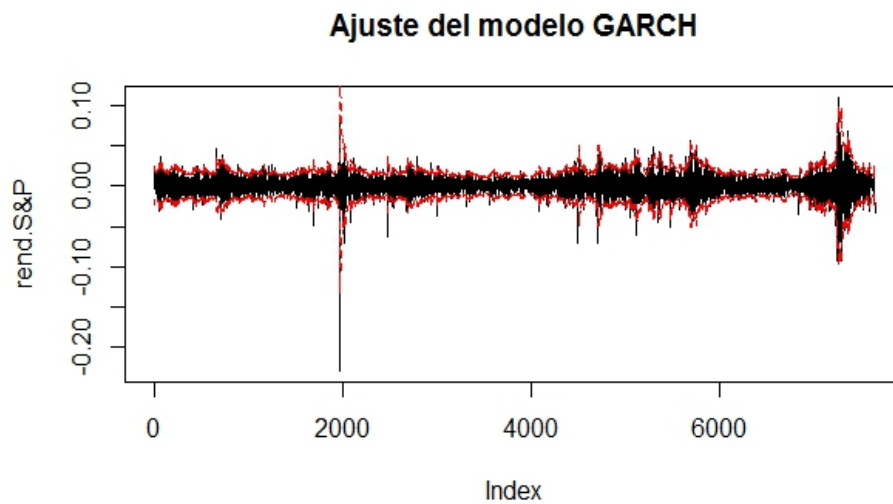


Figura 4.14: Ajuste del modelo 3 a la serie S&P500.

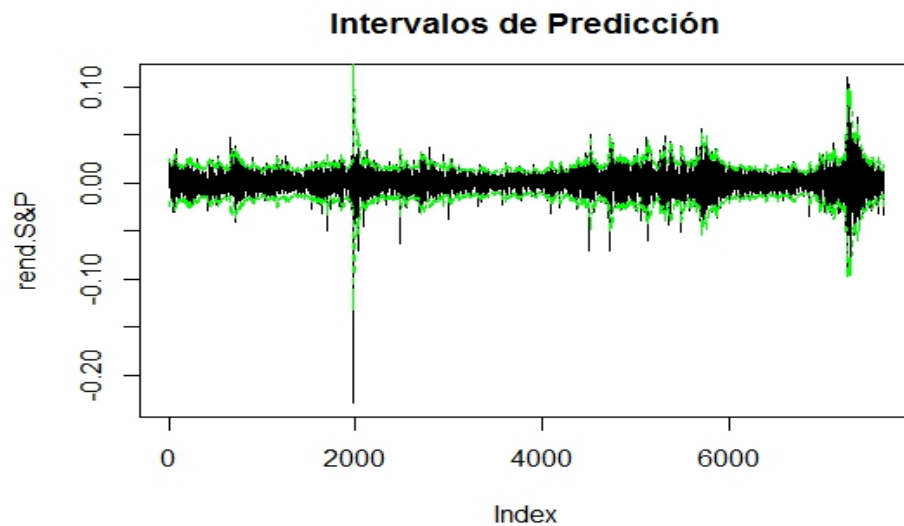


Figura 4.15: Intervalos de predicción para la serie S&P 500.

Conclusiones

En este trabajo se analizó principalmente el modelo GARCH con diferentes parámetros. Los modelos no diferían mucho uno de otro, en cuanto al AIC. Se observó que, a pesar de la sugerencia para la elección de modelos bajo la eacf, algunos parámetros no eran significativos por lo que la sugerencia de que el GARCH(1,1), véase [6], es un modelo óptimo. Por otro lado cada uno de los modelos propuestos cumplió con los supuestos establecidos, excepto el de normalidad, mostrando que estos modelos tienen un buen ajuste.

El supuesto de normalidad siempre es difícil de que se cumpla, y más en este caso por las características de los rendimientos financieros. Debido a esto, han habido diferentes estudios, como el de Eberlein y Keller (1995),¹ en el que hacen un estudio para ver que distribución se ajusta mejor a los rendimientos. En base a esto, hay que resaltar las colas pesadas que tienen este tipo de series, resaltando que la distribución normal para los rendimientos no es la más adecuada.

Gracias a estos modelos podemos hacer predicciones fijándonos en su historia y así tener una idea de que es lo que podría pasar con el activo en cuestión e incluso para carteras de inversión. El análisis de predicción que se realizó en este trabajo fue de un paso adelante el cual nos da una buena representación de los log-rendimientos de la serie. Estos análisis demuestran que a pesar del supuesto de normalidad, los modelos GARCH son buenos para la predicción de un paso.

¹Eberlein, E. and Keller, U. (1995) *Hyperbolic Distributions in Finance*, Bernoulli Vol.1 no.3, 281-99.

Bibliografía

- [1] Brockwell & Davis *Introduction to Time Series and Forecasting, Second Edition*, Springer.
- [2] Brockwell & Davis *Time Series: Theory and Methods*, Springer.
- [3] Campbell, J., Lo, A. & Mackinlay, A. C. *The Econometrics of Financial Markets*, Princeton University Press.
- [4] Gouriéroux, C. *ARCH Models and Financial Applications*, Springer.
- [5] Cryer, J.D. & Chan, K.S. *Time Series Analysis, Second Edition*, Springer.
- [6] Francq, C. & Zakoian, J.-M. *GARCH Models: Structure, Statistical Inference and Financial Applications*, John Wiley.
- [7] Franses, P.H. *Time Series Models for Business and Economic Forecasting*, Cambridge University Press.
- [8] Hull, J.C. *Fundamentals of Futures and Options Markets*, Pearson Education, Prentice Hall.
- [9] Hull, J.C. *Options, Futures and other Derivatives*, Pearson Education, Prentice Hall.
- [10] Shumway, R.H. & Stoffer, D.S. *Time Series Analysis and Its Applications*, Springer.
- [11] Tsay R.S. *Analysis of Financial Time Series, Second Edition*, Wiley.