



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

---

FACULTAD DE CIENCIAS

SERIES DE TIEMPO DE MEMORIA LARGA CON  
VOLATILIDAD ESTOCÁSTICA Y SU APLICACIÓN A LA  
INFLACIÓN MEXICANA

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

ACTUARIO

PRESENTA:

DANIEL CERVANTES FILOTEO

DIRECTOR DE TESIS:

DRA. MARÍA ASUNCIÓN BEGOÑA FERNÁNDEZ FERNÁNDEZ



2014



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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.

## Hoja de Datos del Jurado

### 1. Datos del alumno

Apellido paterno  
Apellido materno  
Nombre(s)  
Teléfono  
Universidad Nacional Autónoma de México  
Facultad de Ciencias  
Carrera  
Número de cuenta

Cervantes  
Filoteo  
Daniel  
56077343  
Universidad Nacional Autónoma de México  
Facultad de Ciencias  
Actuaría  
30707267-2

### 2. Datos del tutor

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

Dra.  
María Asunción Begoña  
Fernández  
Fernández

### 3. Datos del sinodal 1

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

Dr.  
Federico  
O'Reilly  
Togno

### 4. Datos del sinodal 2

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

Dra.  
Ana  
Meda  
Guardiola

### 5. Datos del sinodal 3

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

Mat.  
Margarita Elvira  
Chávez  
Cano

### 6. Datos del sinodal 4

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

Dra.  
Ruth Selene  
Fuentes  
García

### 7. Datos del trabajo escrito.

Título  
Subtítulo  
Número de páginas  
Año

Series de tiempo de memoria larga con volatilidad  
estocástica y su aplicación a la inflación mexicana.  
141 p  
2014

# Índice general

<b>Introducción</b>	<b>V</b>
<b>1. Bases de series de tiempo</b>	<b>1</b>
1.1. Definiciones . . . . .	2
1.1.1. Series de tiempo estacionarias . . . . .	3
1.1.2. Autocorrelación . . . . .	5
1.1.3. Operador de retraso . . . . .	7
1.2. Procesos ARMA . . . . .	8
1.2.1. Procesos MA . . . . .	11
1.2.2. Procesos AR . . . . .	12
1.2.3. Proceso ARMA(1,1) . . . . .	14
1.3. Ajuste y predicción . . . . .	16
1.4. Procesos ARIMA . . . . .	19
<b>2. Volatilidad estocástica</b>	<b>21</b>
2.1. Hechos estilizados . . . . .	24
2.2. Procesos ARCH . . . . .	26
2.2.1. ARCH(1) . . . . .	28
2.3. Procesos GARCH . . . . .	32
2.3.1. GARCH(1,1) . . . . .	34
2.3.2. Una extensión del proceso GARCH . . . . .	37
2.4. Ajuste de un modelo GARCH . . . . .	39

<i>ÍNDICE GENERAL</i>	III
2.4.1. Estimación de parámetros . . . . .	39
2.4.2. Tipo de cambio MXN/EUR . . . . .	40
2.4.3. Estimación de medidas de riesgo . . . . .	49
<b>3. La memoria de un proceso</b>	<b>52</b>
3.1. Análisis espectral . . . . .	53
3.1.1. Densidad espectral y periodograma . . . . .	53
3.1.2. La densidad espectral de un proceso ARMA . . . . .	59
3.1.3. La densidad espectral de un proceso lineal . . . . .	61
3.1.4. Implementación en R . . . . .	62
3.2. Memoria . . . . .	63
3.2.1. Ejemplos de procesos con memoria corta . . . . .	66
3.3. Memoria larga . . . . .	68
3.3.1. Decremento exponencial e hiperbólico . . . . .	68
3.3.2. Estimadores del parámetro de memoria . . . . .	70
<b>4. Modelos de memoria larga</b>	<b>77</b>
4.1. Ruido fraccionario . . . . .	77
4.1.1. Procesos autosimilares . . . . .	78
4.1.2. Ruidos fraccionarios . . . . .	79
4.2. ARFIMA . . . . .	83
4.2.1. La densidad espectral de un proceso ARFIMA . . . . .	85
4.2.2. Ruidos fraccionales integrados . . . . .	87
4.3. Modelos de memoria larga con volatilidad estocástica . . . . .	88
4.3.1. LMSV . . . . .	90
4.3.2. ARFIMA–GARCH . . . . .	91
<b>5. Aplicación. La inflación mexicana</b>	<b>94</b>
5.1. Inflación . . . . .	95
5.1.1. La inflación en México . . . . .	96
5.2. Ajuste del modelo ARFIMA-GARCH . . . . .	102

5.2.1. Medidas de riesgo . . . . .	107
5.2.2. Back-testing para el VaR . . . . .	108
5.3. Conclusiones . . . . .	112
<b>A. Medidas de riesgo</b>	<b>114</b>
A.1. Gestión de riesgos . . . . .	114
A.2. Algunas definiciones . . . . .	117
A.2.1. Medidas de riesgo . . . . .	118
A.2.2. VaR . . . . .	119
A.2.3. Expected shortfall . . . . .	121
<b>B. Conceptos de estadística</b>	<b>122</b>
B.1. t-value . . . . .	122
B.2. p-value . . . . .	123
B.3. QQplot . . . . .	124
B.4. Back-testing para el VaR . . . . .	125

# Introducción

El reto más importante de todo lo que involucró la realización de este trabajo, una vez establecido el tema, fue comenzar a escribir. La principal razón de esto fue que no estaba seguro de lo que quería escribir, a pesar de ya tener el tema.

En matemáticas cuando se tiene un problema este motiva la creación de soluciones, este proceso es el que permite desarrollar teoría, permite que las matemáticas avancen.

Me fue difícil empezar a escribir porque estaba haciendo al revés el proceso, quería escribir sobre una solución y después buscarle el problema. El desarrollo del tema y su escritura se me facilitó mucho cuando encontré el problema y pensando en él estudié la teoría necesaria para atacarlo. Pude distinguir las características del tema que realmente necesitaba para escribir un trabajo coherente de él.

El presente trabajo es un primer acercamiento a los modelos avanzados de series de tiempo univariadas y a su aplicación. La principal motivación para desarrollar modelos de series de tiempo es la necesidad de entender los fenómenos afines a nuestra vida. Buscamos constantemente entender estos fenómenos porque seguramente tienen alguna consecuencia de nuestro interés.

La forma en que se comporte el fenómeno determinará cómo se comportan sus consecuencias. Cuando hablamos de "determinar el comportamiento de las

consecuencias de un fenómeno" estamos hablando de administración de riesgos. El desarrollo de este trabajo está motivado en 2 tipos de problemas, como los siguientes:

Supongamos que un productor de chocolate vive y produce en México, pero su principal mercado se encuentra en Europa, a donde envía sus productos a clientes europeos. El chocolatero tendrá que cubrir los gastos propios del negocio: materia prima, honorarios, exportación, entre otros.

No nos concentraremos en esos gastos. Supondremos que el negocio es rentable. Entonces, el chocolatero está interesado en sus ganancias, en el rendimiento que el negocio le deja (pues seguramente ha hecho fuertes inversiones en el negocio). Si el chocolatero vive en México, sería correcto suponer su ganancia en moneda nacional, en pesos.

El problema es que los clientes pagan en euros y el precio del peso (comprado en euros) no es constante en el tiempo; de hecho, es un proceso estocástico indexado al tiempo y su comportamiento determinará el rendimiento del chocolatero.

**PROBLEMA 1:** Modelar el comportamiento de los rendimientos del tipo de cambio MXN/EUR.

El segundo problema es, de hecho, una complicación del primero. El chocolatero, tan interesado en su rendimiento, tiene ahora el problema de querer modelar su rendimiento real. Una vez que él reciba su ganancia (en pesos) habrá pasado cierto tiempo desde su inversión y es probable que el valor del dinero (en pesos) no sea el mismo que entonces; este fenómeno se conoce como inflación y hablaremos a fondo del tema en el Capítulo 5.

Es evidente que ahora al chocolatero le interesa estudiar el proceso estocástico

que es la inflación en México y así saber qué comportamiento tiene su rendimiento real.

**PROBLEMA 2:** Modelar el comportamiento de las tasas de inflación en México.

Es clara la complicación e importancia que tendrían estos problemas cuando se traducen a instituciones financieras o comerciales de gran tamaño e influencia en la vida de muchas personas. Los problemas económicos que pueda presentar una importante institución de este tipo podría traer consecuencias graves en la sociedad.

Entonces la administración de riesgos es de vital importancia para un país. La determinación del comportamiento de las consecuencias de un fenómeno es la parte de la administración de riesgos que consiste en identificar y medir el riesgo al que se está expuesto por dichas posibles consecuencias.

Después de la identificación y la medición del riesgo sigue la implementación de acciones que ayuden a mitigar ese riesgo hasta un nivel aceptable. Entonces, a grandes rasgos, la administración de riesgo es la identificación, medición y mitigación del riesgo. Sobre la administración de riesgo hablaremos a detalle en el Apéndice A.

Este trabajo se concentra en la medición del riesgo, para lo cual es necesario tener medidas, que serán el VaR y el Expected-shortfall. También hablaremos a fondo de estas medidas en el Apéndice A.

Para calcular las medidas de riesgo será necesario entender el comportamiento de los procesos estocásticos a los que queremos calculárselas. Con el fin de entender dicho comportamiento creamos modelos matemáticos que se aproximen a los

procesos reales.

Lo primero es saber cuáles son las características que queremos que cumplan esos modelos, que deben ser las características de los fenómenos que vamos a estudiar. Una de las principales cualidades que se presentan en los fenómenos financieros es la volatilidad estocástica, esto es, la varianza parece no ser constante en el tiempo.

Además de esto existe una serie de características que normalmente se observan en las finanzas y se conocen con "hechos estilizados". Hablaremos de ellos en el Capítulo 2.

Uno de estos puntos habla sobre la poca correlación que existe entre las variables del proceso, pero esto suele no cumplirse en algunos fenómenos, como en la inflación mexicana donde veremos que hay una fuerte correlación serial. Esto es lo que llamaremos memoria larga.

La presencia de esta característica en la inflación mexicana es lo que nos motiva a buscar un modelo que tenga esa correlación fuerte y además no pierda las características financieras.

En los capítulos 1, 2 y 4 expondremos algunos modelos de series de tiempo que usan para aproximar los fenómenos de interés. En el Capítulo 1 comenzaremos por establecer las definiciones de los conceptos principales para el estudio de las series de tiempo. Hablaremos de los modelos más sencillos pero de gran importancia, los modelos ARMA (Autoregressive Moving Average), que son la base de modelos más generales. Estudiaremos los aspectos teóricos de estos modelos y realizaremos simulaciones de ellos con las cuales revisaremos las características de las realizaciones de estos modelos. Finalmente hablaremos un poco sobre ajuste y predicción con estos modelos y definiremos el modelo ARIMA (Autoregressive

Integrated Moving Average).

En el Capítulo 2 comenzaremos el estudio de los modelos de volatilidad estocástica; definiremos y estudiaremos los modelos ARCH (Autoregressive Conditional Heteroskedasticity) y GARCH (Generalized Autoregressive Conditional Heteroskedasticity), hablaremos de sus características matemáticas, un poco de su historia y veremos simulaciones de ellos. También definiremos el modelo ARMA–GARCH. Para terminar ese capítulo veremos una aplicación del modelo GARCH; primero estudiaremos la serie de tiempo del tipo de cambio MXN/EUR y le ajustaremos un modelo GARCH al comportamiento de sus rendimientos.

El Capítulo 3 será el preámbulo de los modelos de memoria larga. Definiremos el concepto de memoria en una serie de tiempo; para ello comenzaremos por estudiar algunos conceptos de análisis espectral. Concluiremos el capítulo definiendo series de tiempo de memoria larga y sus propiedades, hablaremos sobre el parámetro de memoria de una serie de tiempo y enunciaremos estimadores para este parámetro.

El Capítulo 4 tratará de los modelos con memoria larga y sus características, definiremos los ruidos fraccionarios y los modelos ARFIMA (Autoregressive Fractionally Integrated Moving Average). Finalmente estudiaremos modelos con memoria larga y volatilidad estocástica el LMSV y el ARFIMA–GARCH. Veremos simulaciones de estos procesos y demostraremos que tienen memoria larga, en algunos casos encontraremos el parámetro de memoria.

Para concluir el trabajo veremos una aplicación de los procesos con memoria larga y volatilidad estocástica en el Capítulo 5. Estudiaremos la serie de tiempo de la inflación mexicana y veremos las características que nos harán pensar que la serie tiene memoria larga y estimaremos su parámetro de memoria. Modelaremos la serie con un proceso ARFIMA–GARCH y una vez estimados todos los pará-

metros del modelo calcularemos las medidas de riesgo para la tasa de inflación mexicana. Finalmente calificaremos el modelo aplicando un Back-testing.

Los datos de las series de tiempo que estudiaremos fueron obtenidos de la base de datos del Banco de México en la página <http://www.banxico.org.mx/>.

# Capítulo 1

## Bases de series de tiempo

Comenzaremos por establecer la definición de serie de tiempo y su diferencia con una realización. Diremos que una serie de tiempo es un proceso estocástico indexado al tiempo (discretizado), es decir, una sucesión de variables aleatorias, y la realización de la serie de tiempo es el conjunto de datos que las variables aleatorias arrojaron al ir pasando el tiempo o al realizarse una simulación.

Las series de tiempo son un concepto fundamental en el estudio científico de casi cualquier cosa. Necesitamos observar el fenómeno que vamos a estudiar pues con lo que observamos hacemos inferencia de cómo creemos que se comportará el fenómeno más adelante.

Las observaciones que tenemos -los datos- son la realización de la serie, y la idea que proponemos de cómo se comportará después, es el modelo de la serie de tiempo (un concepto más); el modelo de la serie de tiempo es la propuesta de cómo, según nosotros, se distribuyen las variables aleatorias de la serie de tiempo.

Es importante no confundir la serie de tiempo con la realización, aunque son conceptos que van de la mano naturalmente, pero es más importante no confundir estos conceptos con el modelo de la serie.

El modelo es algo creado por nosotros que, con conocimientos básicos de probabilidad, sabemos que jamás será exactamente la distribución de la serie, por lo que no es la serie misma; solo es una aproximación.

Los modelos que se proponen tratan de ser lo más parecido posible a la serie de tiempo, y con la gran variedad de datos que hay para "jugar" se genera una enorme cantidad de modelos que enriquecen las posibilidades para estudiar datos. Este permanente desarrollo permite tener modelos mejores y más complejos.

En este capítulo vamos a estudiar algunos conceptos básicos de series de tiempo y algunos modelos con diferentes características que serán la base de los modelos que estudiaremos en el capítulo 2.

El contenido de este capítulo está basado en BROCKWELL P.J. y DAVIS R.A. [5], MCNEIL A.J., FREY R. y EMBRECHTS P. [11], ZONGWU, C. [22] y WEI, W.W.S. [18].

## 1.1. Definiciones

Definimos a una serie de tiempo como un proceso estocástico  $(X_t)_{t \in \mathbb{Z}}$ , es decir, una familia de variables aleatorias, sobre un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ . Si suponemos que los momentos de las variables aleatorias existen, podemos definir las funciones media  $(\mu_t)$  y autocovarianza  $(\gamma_{(t,s)})$  como:

$$\begin{aligned}\mu_t &= \mathbb{E}[X_t] \\ \gamma_{(t,s)} &= \mathbb{E}[(X_t - \mu_t)(X_s - \mu_s)].\end{aligned}$$

Podemos observar que la función de autocovarianza satisface que  $\gamma_{(t,s)} = \gamma_{(s,t)}$  y que  $\gamma_{(t,t)} = \text{Var}[X_t]$ .

Otra propiedad que nos interesará estudiar es la estacionariedad de la serie, y definiremos tres tipos.

### 1.1.1. Series de tiempo estacionarias

#### Definición 1.

Diremos que una serie de tiempo,  $(X_t)_{t \in \mathbb{Z}}$ , es **estrictamente estacionaria** si:

$$(X_{t_1}, \dots, X_{t_n}) =_D (X_{t_1+k}, \dots, X_{t_n+k})$$

para cualesquiera valores  $t_1, t_2, \dots, t_n, k \in \mathbb{Z}$  y con  $n \in \mathbb{N}$ .

Donde " $=_D$ " significa igualdad en distribución.

Es decir, una serie con esta propiedad será tal que la distribución de cualquier vector  $(X_{t_1}, \dots, X_{t_n})$  es invariante bajo traslaciones.

Otra forma de estacionariedad es cuando pedimos estacionariedad de las características numéricas de la serie.

#### Definición 2.

Diremos que una serie de tiempo, cuyas variables tengan segundo momento finito, es **débilmente estacionaria** (estacionaria de segundo orden) si:

$$\begin{aligned} \mu_t &= \mu \\ \gamma_{(t,s)} &= \gamma_{(t+k,s+k)} \end{aligned}$$

para todos  $t, s, k \in \mathbb{Z}$ .

Podemos ver, directamente de las definiciones, que si una serie de tiempo es estrictamente estacionaria y con varianza finita, entonces es débilmente estacionaria. Es importante señalar que hay procesos, con varianza infinita, que son estrictamente estacionarios pero no son débilmente estacionarios.

Para las series débilmente estacionarias se tiene:

$$\begin{aligned}\gamma(t-s,0) &= \gamma(t,s) \\ &= \gamma(s,t) \\ &= \gamma(s-t,0)\end{aligned}$$

Por lo que concluimos que la autocovarianza va sólo a depender de la separación de las variables (el *lag* o rezago). Entonces, para este tipo de series definiremos la función de autocovarianza como:

$$\gamma(h) = \gamma(h,0)$$

Esta función será de gran ayuda para hablar de la memoria de un proceso.

Notemos que  $\gamma_{(0)} = \text{Var}[X_t], \forall t \in \mathbb{Z}$ . Con esta idea definiremos la función de autocorrelación para series débilmente estacionarias.

La última forma de estacionariedad que veremos es cuando la distribución de los incrementos se conserva.

**Definición 3.**

*Diremos que una serie de tiempo tiene **incrementos estacionarios** si cumple que:*

$$(X_t - X_s) =_D (X_{t+h} - X_{s+h})$$

*para todos  $t, s, h \in \mathbb{Z}$ .*

Este tipo de estacionariedad es una propiedad de una serie de tiempo estrictamente estacionaria, pero una serie con incrementos estacionarios no necesariamente será una estrictamente estacionaria. Esto ocurre, por ejemplo, un proceso Poisson,  $\{X_t\}$ , donde  $X_t \sim Poisson(t\lambda)$ , por lo que no es estrictamente estacionario pero  $(X_t - X_s) \sim Poisson((t - s)\lambda)$  entonces, tiene incrementos estacionarios.

### 1.1.2. Autocorrelación

#### Definición 4.

Definimos a la **función de autocorrelación (ACF)** de un proceso débilmente estacionario como

$$\rho_{(h)} = \rho(X_h, X_0) = \frac{\gamma_{(h)}}{\gamma_{(0)}}$$

para toda  $h \in \mathbb{Z}$ .

La función de autocorrelación o la correlación serial es muy importante en el análisis clásico de las series de tiempo.

Ahora pasaremos a estudiar otro concepto, este será la clave que nos permitirá definir los modelos de series de tiempo.

#### Definición 5.

Diremos que un proceso estocástico  $(\epsilon_t)_{t \in \mathbb{Z}}$  es un **ruido blanco** si es débilmente estacionario y su función de autocorrelación es:

$$\rho_{(h)} = \begin{cases} 1, & \text{si } h = 0, \\ 0, & \text{si } h \neq 0. \end{cases}$$

Notación:  $\epsilon_t$  es un  $WN(\mu, \sigma^2)$  donde  $\mu$  es la media del proceso y  $\sigma^2$  su varianza.

Un ruido blanco centrado tiene media cero,  $WN(0, \sigma^2)$ . También podemos pedirle más propiedades al ruido blanco para definir un tipo de variable distinta.

**Definición 6.**

Un proceso estocástico  $(\epsilon_t)_{t \in \mathbb{Z}}$  es un **ruido blanco estricto** si es una sucesión de variables aleatorias independientes e idénticamente distribuidas con varianza finita. Notación:  $\epsilon_t$  es un  $SWN(\mu, \sigma^2)$ ,  $\mu$  es la media y  $\sigma^2$  la varianza de  $\epsilon_t$ .

Otro concepto que usaremos después es la diferencia-martingala. Para estudiarlo debemos suponer que la serie de tiempo  $(X_t)_{t \in \mathbb{Z}}$  es adaptada a cierta filtración <sup>1</sup>  $(\mathcal{F}_t)_{t \in \mathbb{Z}}$  que representa la información obtenida hasta el momento  $t$ .

Lo más común es tomar la filtración natural del proceso, esto es,  $\mathcal{F}_t = \sigma(X_s : s \leq t)$ . Esta filtración nos da la información generada por lo que ya ha ocurrido en el proceso que estamos estudiando y es la mínima  $\sigma$ -álgebra que contiene esa información.

**Definición 7.**

Diremos que un proceso  $(X_t)_{t \in \mathbb{Z}}$  es una **diferencia - martingala**, respecto a una filtración  $(\mathcal{F}_t)_{t \in \mathbb{Z}}$  (normalmente la natural) si:

1. Es integrable ( $\mathbb{E}[|X_t|] < \infty$ ),
2. Adaptada a la filtración ( $X_t$  es v.a. respecto a  $\mathcal{F}_t$ ) y
3.  $\mathbb{E}[X_t | \mathcal{F}_{t-1}] = 0$

para toda  $t \in \mathbb{Z}$ .

---

<sup>1</sup>Una filtración es una familia de  $\sigma$ -álgebras,  $(\mathcal{F}_t)_{t \in \mathbb{Z}}$ , tales que  $\mathcal{F}_n \subseteq \mathcal{F}_m$  cuando  $n \leq m$ .

Podemos ver directamente que:

$$\mathbb{E}[X_t] = 0, \forall t \in \mathbb{Z}.$$

Además si  $\mathbb{E}[X_t^2] < \infty$ , podemos dar información sobre la función de autocovarianza (sin pérdida de generalidad supongamos que  $s < t$ ):

$$\begin{aligned} \gamma_{(t,s)} &= \mathbb{E}[X_t X_s] \\ &= \mathbb{E}[\mathbb{E}[X_t X_s | \mathcal{F}_{s-1}]] \\ &= \mathbb{E}[X_t \mathbb{E}[X_s | \mathcal{F}_{s-1}]] \\ &= 0. \end{aligned}$$

Por lo que si una diferencia-martingala tiene varianza constante en el tiempo, entonces será un ruido blanco.

### 1.1.3. Operador de retraso

El ultimo concepto que veremos en esta sección es el operador de retraso. Todos los modelos de series de tiempo que veremos pueden definirse (y suelen definirse) en términos de operadores de retaso.

#### **Definición 8.**

Sea  $(X_t)_{t \in \mathbb{Z}}$  una serie de tiempo entonces, el **operador de retraso**  $B$  es tal que:

$$BX_t = X_{t-1}$$

para toda  $t \in \mathbb{Z}$ .

Directamente de la definición se tiene que:

$$B^k X_t = X_{t-k}$$

El operador de retaso compuesto con polinomios resulta muy útil para definir los modelos de series de tiempo que veremos. Esta composición es:

Sea  $\psi(z) = \sum_j \psi_j z^j$  un polinomio dado entonces, podemos abreviar:

$$\sum_j \psi_j X_{t-j} = \psi(B)X_t$$

Ahora vamos definir algunas familias paramétricas de modelos de series de tiempo que son muy usados en la teoría y en la práctica. Estos modelos suelen ser lo primeros que se estudian en los cursos de series de tiempo. Son bastante sencillos pero de vital importancia para el estudio de modelos más avanzados.

## 1.2. Procesos ARMA

Comezaremos a estudiar la familia de procesos ARMA (Autoregressive Moving Average). Esta es una familia clásica en el estudio de series de tiempo. Estos procesos son débilmente estacionarios y se construyen usando ruidos blancos.

### Definición 9.

Sea  $(\epsilon_t)_{t \in \mathbb{Z}}$  un  $WN(0, \sigma_\epsilon^2)$ . Decimos que  $\{X_t\}_{t \in \mathbb{Z}}$  es un **proceso ARMA**( $p, q$ ) de media cero si es débilmente estacionario y satisface:

$$X_t - \sum_{i=1}^p \phi_i X_{t-i} = \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}, \quad \forall t \in \mathbb{Z}.$$

para  $t \in \mathbb{Z}$ ,  $p, q \in \mathbb{N} \cup \{0\}$ .

Y diremos que  $\{X_t\}_{t \in \mathbb{Z}}$  es un proceso ARMA( $p, q$ ) de media  $\mu$  si el proceso  $(X_t - \mu)_{t \in \mathbb{Z}}$  es un proceso ARMA( $p, q$ ) de media cero.

La forma resumida de escribir la ecuación que define al proceso ARMA es:

$$\tilde{\phi}_p(B)X_t = \tilde{\theta}_q(B)\epsilon_t, \quad \forall t \in \mathbb{Z}.$$

Al ruido blanco de la definición lo llamamos innovación.

Notemos que si las innovaciones son independientes e idénticamente distribuidas o son un proceso estrictamente estacionario, entonces el proceso ARMA( $p, q$ ) será un proceso estrictamente estacionario.

Para propósitos prácticos restringiremos nuestro estudio de procesos ARMA a procesos ARMA causales. Estos satisfacen:

$$X_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \tag{1.1}$$

$$\sum_{i=0}^{\infty} |\psi_i|^2 < \infty \tag{1.2}$$

Donde  $(\epsilon_t)_{t \in \mathbb{Z}}$  es un  $WN(0, \sigma_\epsilon^2)$ .

El estudio de los procesos ARMA( $p, q$ ) causales se convierte en el estudio de los 2 polinomios que están dados en términos de los parámetros del proceso.

$$\tilde{\phi}_p(z) = 1 - \phi_1 z - \dots - \phi_p z^p$$

$$\tilde{\theta}_q(z) = 1 + \theta_1 z + \dots + \theta_q z^q$$

Por ejemplo, si condicionamos a que los polinomios no tengan raíces en común, entonces el proceso ARMA es causal si y sólo si el polinomio  $\tilde{\phi}$  tiene todas sus raíces fuera del círculo unitario  $|z| \leq 1$ .

Con esta notación los coeficientes de (1.1) se pueden determinar por:

$$\sum_{i=0}^{\infty} \psi_i z^i = \frac{\tilde{\theta}_q(z)}{\tilde{\phi}_p(z)}, \quad |z| \leq 1$$

**Lema 1.**

Cualquier proceso que satisfaga (1.1) y (1.2) es débilmente estacionario y su función de autocorrelación es

$$\rho_{(h)} = \left( \sum_{i=0}^{\infty} \psi_i \psi_{i+|h|} \right) \left( \sum_{i=0}^{\infty} \psi_i^2 \right)^{-1}, \quad h \in \mathbb{Z}.$$

*Demostración.* Primero notemos que:

$$\begin{aligned} \mathbb{E}[X_t] &= \mathbb{E} \left[ \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \right] \\ &= \sum_{i=0}^{\infty} \psi_i \mathbb{E}[\epsilon_{t-i}] = 0 \end{aligned}$$

y que

$$\begin{aligned} \text{Var}(X_t) &= \mathbb{E}[X_t^2] \\ &= \sum_{i=0}^{\infty} \psi_i^2 \mathbb{E}[\epsilon_{t-i}^2] \\ &= \sigma_\epsilon^2 \sum_{i=0}^{\infty} \psi_i^2 \end{aligned}$$

por lo que se tiene:

$$\begin{aligned} \gamma_{(h)} &= \mathbb{E}[X_t X_{t+h}] \\ &= \mathbb{E} \left[ \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \sum_{j=0}^{\infty} \psi_j \epsilon_{t+h-j} \right] \\ &= \sigma_\epsilon^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+|h|} \end{aligned}$$

finalmente usamos la definición de autocorrelación para tener:

$$\begin{aligned}\rho_{(h)} &= \frac{\gamma_{(h)}}{\text{Var}(X_t)} \\ &= \left( \sum_{i=0}^{\infty} \psi_i \psi_{i+|h|} \right) \left( \sum_{i=0}^{\infty} \psi_i^2 \right)^{-1}\end{aligned}$$

□

Si tenemos un proceso ARMA y anulamos uno de sus parámetros (es decir, hacemos  $p = 0$  ó  $q = 0$ ), se obtienen procesos que son parte de la familia ARMA pero con su nombre propio, ya sea un MA o un AR.

### 1.2.1. Procesos MA

Los procesos MA son llamados así por sus siglas en inglés (Moving Average). Un proceso MA( $q$ ) es un proceso ARMA(0,  $q$ ). Es decir, un MA( $q$ ) satisface la ecuación:

$$X_t = \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i}, \quad \forall t \in \mathbb{Z}.$$

A partir del Lema 1 podemos ver que:

$$\rho_{(h)} = \left( \sum_{i=0}^{q-|h|} \theta_i \theta_{i+|h|} \right) \left( \sum_{i=0}^q \theta_i^2 \right)^{-1}, \quad |h| \in \{0, 1, 2, \dots, q\}.$$

El paquete estadístico  $R$  nos permite hacer simulaciones de este tipo de series de tiempo a partir de un proceso ARMA( $p, q$ ).

La simulación de este proceso MA( $q$ ) requiere especificar las constantes que lo definen -las  $\theta_i$ - tantas como se quiera el orden del modelo.

Por ejemplo, un MA(2) con  $\theta_1 = -0.7$  y  $\theta_2 = 0.1$ ; simularemos 100 datos. El código para la simulación, para ver la gráfica de la serie de tiempo y su ACF es:

```
library(tseries)
ma.sim<-arima.sim(model=list(ma=c(-.7,.1)),n=150)
plot(ma.sim)
acf(ma.sim, main="ACF")
```

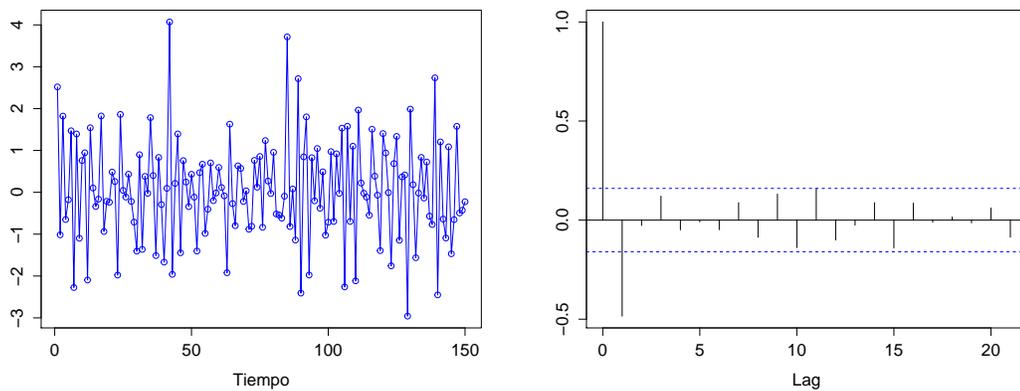


Figura 1.1: Simulación de un proceso MA(2) con  $\theta_1 = -0.7$  y  $\theta_2 = 0.1$  y su ACF.

### 1.2.2. Procesos AR

El otro proceso miembro de la familia ARMA es el AR, proceso autoregresivo. Un AR( $p$ ) es un proceso ARMA( $p, 0$ ). Es decir, satisface:

$$X_t = \epsilon_t + \sum_{i=1}^p \phi_i X_{t-i}, \quad \forall t \in \mathbb{Z}.$$

Ahora, analizaremos un caso particular, un proceso AR(1), este debe cumplir

la ecuación:

$$X_t = \epsilon_t + \phi_1 X_{t-1}$$

si sustituimos el valor de  $X_{t-1}$ , tenemos:

$$X_t = \epsilon_t + \phi_1(\epsilon_{t-1} + \phi_1 X_{t-2})$$

y si hacemos esto  $k$  veces, tendremos:

$$X_t = \epsilon_t + \phi_1^{k+1} X_{t-k-1} + \sum_{i=0}^k \phi_1^i \epsilon_{t-i}$$

y si además agregamos la condición  $|\phi_1| < 1$  (esto nos permitirá que el proceso sea causal), tendremos:

$$X_t = \sum_{i=0}^{\infty} \phi_1^i \epsilon_{t-i}$$

Podemos notar que un proceso AR(1) puede verse como un proceso MA( $\infty$ ), un proceso de medias móviles de orden infinito.

Para este proceso podemos calcular fácilmente sus funciones de autocovarianza y de autocorrelación:

$$\begin{aligned} \gamma(h) &= \frac{\phi_1^{|h|} \sigma_\epsilon^2}{1 - \phi_1^2} \\ \rho(h) &= \phi_1^{|h|}, \quad \forall t \in \mathbb{Z} \end{aligned}$$

Su ACF también es bastante especial. El siguiente código genera una simulación de 100 datos de un AR(1) con parámetro 0.8.

```
library(tseries)
ar.sim <- arima.sim(model=list(ar=c(.8)),n=150)
plot(ar.sim)
acf(ar.sim, main="FAC")
```

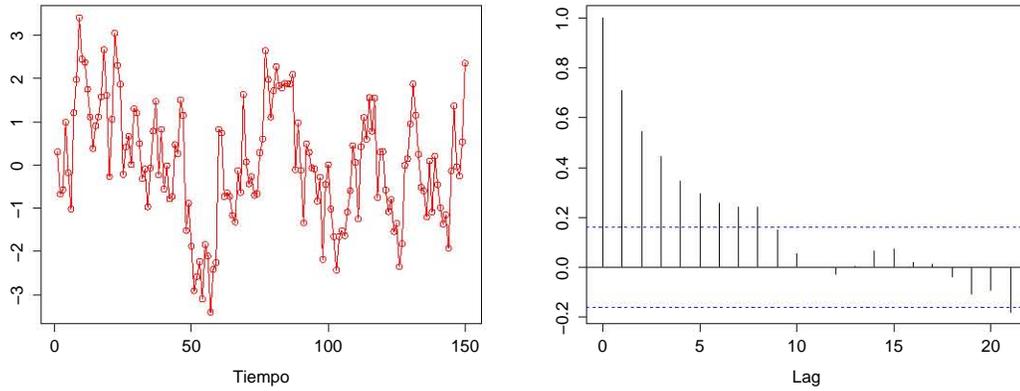


Figura 1.2: Simulación de un proceso AR(1) con  $\phi_1 = 0.8$  y su ACF.

Ahora veremos el caso particular del proceso ARMA(1,1), algunas de sus características y una simulación.

### 1.2.3. Proceso ARMA(1,1)

El ARMA(1,1) esta dado por la siguiente ecuación:

$$X_t - \phi X_{t-1} = \epsilon_t + \theta \epsilon_{t-1}, \quad \forall t \in \mathbb{Z}.$$

En este caso los polinomios del proceso serán:

$$\tilde{\phi}(z) = 1 - \phi z, \quad \tilde{\theta}(z) = 1 + \theta z$$

por lo que:

$$\sum_{i=0}^{\infty} \psi_i z^i = \frac{1 + \theta z}{1 - \phi z} = (1 + \theta z)(1 + \phi z + \phi^2 z^2 + \dots)$$

entonces tendremos:

$$X_t = \epsilon_t + (\theta + \phi) \sum_{i=1}^{\infty} \phi^{i-1} \epsilon_{t-i}$$

Podemos decir entonces que un proceso ARMA(1, 1) se puede ver en términos de un MA( $\infty$ ) e incluso si condicionamos a que  $|\theta| < 1$  entonces, podemos escribir al ARMA(1, 1) como un AR( $\infty$ ) de la forma:

$$X_t = \epsilon_t + (\theta + \phi) \sum_{i=1}^{\infty} (-\theta)^{i-1} X_{t-i}$$

Veremos ahora la simulación de un proceso ARMA(1, 1) con parámetros  $\phi = 0.6$  y  $\theta = 0.5$ . De esto se tiene:

```
arma.sim<-arima.sim(n=100, list(ar=c(0.6), ma=c(0.5)))
plot(arma.sim,ylim=c(-4,4),main="ARMA(1,1)")
acf(arma.sim, main="ACF",ylim=c(-1,1))
```

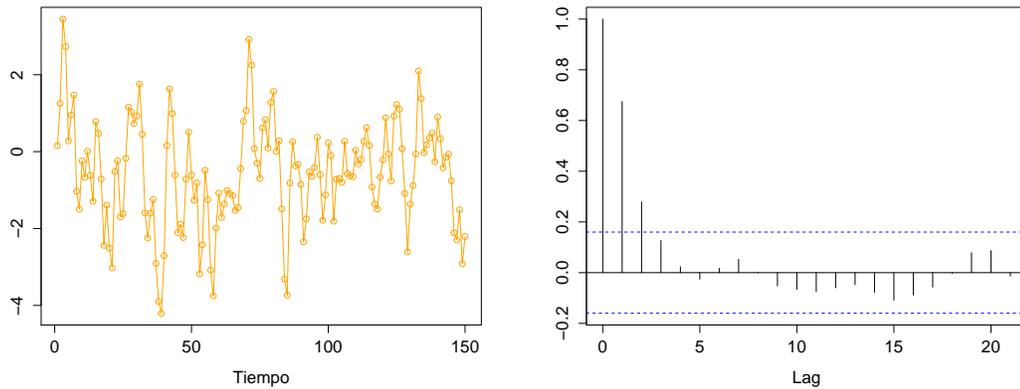


Figura 1.3: Simulación de un proceso ARMA(1,1) con parámetros  $\phi = 0.6$  y  $\theta = 0.5$  y su ACF.

### 1.3. Ajuste y predicción

Los modelos de series de tiempo que creamos nos ayudan a entender mejor los fenómenos de nuestro interés. Los modelos entonces, son aproximaciones de lo que ocurre en la realidad por lo que serán determinados por ella misma.

Los modelos que hemos visto hasta ahora, y también los que veremos más adelante son modelos paramétricos, depende de parámetros. Esto quiere decir que para hacer un modelo que se aproxime al fenómeno que estudiamos debemos calcular los parámetros del modelo. El cálculo o estimación de dichos parámetros es el ajuste del modelo.

Para estimar los parámetros de un modelo se usan métodos estadísticos que proponen a los parámetros como una función de las observaciones del fenómeno.

La cantidad de datos a usar y su ubicación debe determinarlo el analista. Si la mitad de un año de datos, un año, cinco años o 10 años son apropiados para modelar, dependerá de la situación (de lo que se quiere hacer).

Sin duda, es buena idea hacer varios análisis con diferentes rangos de datos para observar la sensibilidad de la inferencia estadística para la cantidad de datos.

Después de un análisis que nos permita decidir qué datos usaremos los colocamos en un vector  $\mathbb{X} = (X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$  para hacer la estimación de los parámetros. Donde  $X_i$  es la variable aleatoria que denota el valor de la serie a tiempo  $i$  y  $x_i$  el valor que tomó.

Comencemos con el ajuste del modelo; no indagaremos en cómo se llega a la decisión de ajustar un modelo ARMA a los datos que estamos analizando. Sólo supongamos que un confiable amigo nos observa trabajando con esos datos y, al

ver graficada la serie de tiempo, nos dice: "Estoy seguro de que eso se comporta como un ARMA". U

Una vez que definimos el modelo que vamos a ajustar podemos dar la **función de verosimilitud** que será la función de densidad del vector  $\mathbb{X}$  en el punto  $(x_1, x_2, \dots, x_n)$ . Esta es función de los parámetros del modelo que definen la distribución del vector  $\mathbb{X}$ .

Definimos a  $\Theta$  el vector de parámetros del modelo. La estimación de parámetros consiste en buscar el vector  $\hat{\Theta}$  que maximice la función de verosimilitud,  $L(\mathbb{X}, \Theta)$ .

Pensemos en un ARMA causal para el ajuste. Supongamos, por ahora, que ya estimamos los parámetros; entonces tenemos 2 cosas: el modelo ajustado  $\hat{\Theta} = (\hat{\mu}, \hat{\phi}_p, \hat{\theta}_q)$  y los datos  $(X_t)$ . Con esto calculamos los residuales del ajuste. Esto es:

$$\hat{\epsilon}_t = X_t - \hat{x}_t$$

donde

$$\hat{x}_t = \hat{\mu} + \sum_{i=0}^p \hat{\phi}_i (X_{t-i} - \hat{\mu}) + \sum_{j=0}^q \hat{\theta}_j \hat{\epsilon}_{t-j}$$

Es evidente que tendremos problemas para calcular el primer valor de  $\hat{\epsilon}_t$ , por lo que una respuesta simple para este problemas será simplemente usar  $\hat{\epsilon}_0 = 0$ .

Los residuales  $\hat{\epsilon}_t$  deben comportarse como un proceso de ruido blanco, ya que esta es nuestra hipótesis de modelo para las innovaciones, y esto se puede evaluar mediante la construcción de su correlograma.

Cuando no hay pruebas de correlación serial en la función de autocorrelación esto sugiere que se ha encontrado un modelo ARMA correcto.

Otra de las bondades de R es que nos permite ajustar un modelo de la familia ARMA a datos reales que estemos estudiando. Calculando por máxima verosimilitud los parámetros del modelo. Además, cuenta con algunas bases de datos precargados para hacer estudios sobre ellos.

La forma de ajustar un modelo ARMA a datos es ejecutando el código *arma*, el comando *summary* nos da los resultados del ajuste, la estimación de los parámetros y datos afines al ajuste como el error estándar del estimador o el p-valor. El comando *plot* nos arroja los resultados gráficos del ajuste y los residuales pueden obtenerse con el comando *residuals(r.arma)*.

## Pronóstico

Cuando ajustamos modelos a datos no solo lo hacemos por diversión sino también para dar un pronóstico de lo que ocurrirá con la serie de tiempo en el futuro; supongamos que tenemos datos a los que se les ha ajustado un modelo ARMA(1, 1).

Tenemos una estimación de los parámetros, entonces la predicción a un paso (un tiempo hacia adelante) la obtendremos dando el valor esperado de la serie, dado todo lo que ha pasado, es decir:

$$\mathbb{E}[X_{t+1}|\mathcal{F}_t] = \mu_{t+1} = \mu + \phi(X_t - \mu) + \theta\epsilon_t$$

Por lo que una predicción a  $k$  pasos será entonces:

$$\mathbb{E}[X_{t+k}|\mathcal{F}_t] = \mu_{t+k} = \mu + \phi^k(X_t - \mu) + \phi^{k-1}\theta\epsilon_t$$

y bajo la hipótesis de que  $|\phi| < 1$  se tiene entonces que:

$$\lim_{k \rightarrow \infty} \mathbb{E}[X_{t+k}|\mathcal{F}_t] = \mu, \quad c.s.$$

## 1.4. Procesos ARIMA

En los estudios que haremos en el Capítulo 4 y 5, de los procesos de memoria larga, usaremos una generalización de los modelos ARMA, los ARFIMA, procesos autoregresivos fraccionalmente integrados y de medias móviles.

Antes vamos a definir a los procesos ARIMA que son procesos más sencillos de estudiar. Ahora, los procesos ARIMA se definen en términos de los ARMA. Un proceso  $X_t$  es un procesos ARIMA si  $(1 - B)^d X_t$  es un proceso ARMA, es decir:

### Definición 10.

Sea  $(\epsilon_t)_{t \in \mathbb{Z}}$  un  $WN(0, \sigma_\epsilon^2)$ . Decimos que  $(X_t)_{t \in \mathbb{Z}}$  es un **proceso ARIMA**  $(p, d, q)$  si satisface:

$$\tilde{\phi}_p(B)(1 - B)^d X_t = \tilde{\theta}_q(B)\epsilon_t$$

para toda  $t \in \mathbb{Z}$  y con  $p, q, d \in \mathbb{N} \cup \{0\}$ .

La propiedad de que  $(1 - B)^d X_t$  es un proceso ARMA nos sirve mucho para transformar un proceso ARIMA en un ARMA y después realizar un ajuste.

Lo que suele hacerse cuando tenemos una serie de tiempo que se cree que es de la familia ARIMA es aplicar el operador  $(1 - B)^d$  para transformarla en una serie ARMA y después trabajar con la serie transformada. Este proceso se conoce como diferenciación de la serie.

No hablaremos mucho de los procesos ARIMA, solo es necesario definirlos para estudiar procesos similares en el Capítulo 4. Por ahora veremos solo una simulación de este proceso. El código es:

```
arima <- arima.sim(list(order=c(1,1,1), ar=.4, ma =.9), n=100)
```

```
plot(arima,type="o",col="orange", main="ARIMA(1,1,1)")
```

La gráfica del proceso ARIMA simulado se observa en la figura 1.4.

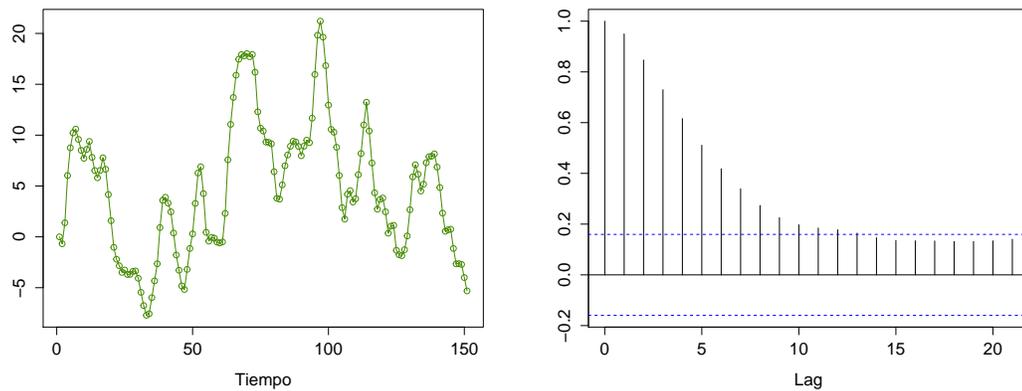


Figura 1.4: Simulación de un proceso ARIMA(1,1,1) con parámetros  $\phi = 0.4$ ,  $\theta = 0.9$  y  $d = 1$  y su ACF.

Con este proceso terminamos los conceptos y modelos básicos de series de tiempo. En el siguiente capítulo introduciremos algunos modelos avanzados.

# Capítulo 2

## Volatilidad estocástica

En este capítulo veremos algunos de los modelos para el estudio de series de factores de riesgo financiero. Veremos las definiciones de los modelos ARCH (autorregresivos de heteroscedasticidad condicional) y los GARCH (generalizaciones del ARCH), discutiremos algunas de sus propiedades matemáticas y su uso en la práctica.

El contenido de este capítulo está basado en ENGLE, R. [7], BOLLERSLEV, T.P. [3], MCNEIL A.J., FREY R. y EMBRECHTS P. [11] y WEI, W.W.S. [18].

Los modelos ARCH aparecen en los años 80 (ENGLE, R. [7]) con el objetivo de modelar los raros agrupamientos temporales de volatilidad que suelen observarse en las series de rendimiento de casi todo el mercado financiero. Desde entonces su variedad y aplicación en la práctica ha crecido tremendamente.

Las series de tiempo que producen los factores de riesgo financiero, en realidad, las diferencias logarítmicas (log-rendimientos) de los precios o tasas de cambio, presentan características peculiares que se estudian con modelos específicos. Daremos una breve introducción a estos modelos, los de volatilidad estocástica.

Se ha notado empíricamente que los datos de rendimiento financiero diario muestran una Kurtosis mucho más alta que una coherente a la distribución normal; su distribución se dice que es leptocúrtica, lo que significa que es más estrecha en el centro, pero tiene colas más largas y más pesadas que una distribución normal.

Además de esta observación, las series de tiempo financieras presentan características muy específicas. Estas cualidades se conocen como "Hechos estilizados".

En el artículo de los modelos ARCH (ENGLE, R. [7]), Engle cita tres situaciones que motivan y justifican la modelización con el uso de Heterocedasticidad Condicional Autorregresiva. Estas fueron las siguientes:

*"... The ARCH regression model has a variety of characteristics which make it attractive for econometric applications. Econometric forecasters have found that their ability to predict the future varies from one period to another.*

*"McNees suggests that, "the inherent uncertainty or randomness associated with different forecast periods seems to vary widely over time." He also documents that, "large and small errors tend to cluster together (in contiguous time periods)." This analysis immediately suggests the usefulness of the ARCH model where the underlying forecast variance may change over time and is predicted by past forecast errors. The results presented by McNees also show some serial correlation during the episodes of large variance.*

*"A second example is found in monetary theory and the theory of finance. By the simplest assumptions, portfolios of financial assets are held as functions of the expected means and variances of the rates of return. Any shifts in asset demand must be associated with changes in expected means and variances of the rates of return. If the mean is assumed to follow a standard regression or time-series model, the variance is immediately constrained to be constant over time. The use*

*of an exogenous variable to explain changes in variance is usually not appropriate.*

*"A third interpretation is that the ARCH regression model is an approximation to a more complex regression which has non-ARCH disturbances. The ARCH specification might then be picking up the effect of variables omitted from the estimated model. The existence of an ARCH effect would be interpreted as evidence of misspecification, either by omitted variables or through structural change. If this is the case, ARCH may be a better approximation to reality than making standard assumptions about the disturbances, but trying to find the omitted variable or determine the nature of the structural change would be even better. "*

1. En primer lugar, la experiencia empírica nos lleva a contrastar períodos de amplia varianza seguidos de otros de varianza más pequeña. Se requiere un modelo que nos ayude a pronosticar la varianza pasada de la serie para hacer mejor predicción de la varianza futura.
2. El segundo ejemplo es sobre el mantenimiento o venta de activos financieros. Los agentes financieros deciden esta cuestión en función de la información proveniente del pasado: al valor medio de su rentabilidad y la volatilidad que ésta ha tenido.
3. Por último, el modelo de regresión ARCH puede ser una aproximación a un sistema más complejo. El modelo ARCH podría recoger el efecto de variables omitidas. El modelo ARCH puede ser una mejor aproximación a la realidad que hacer suposiciones sobre los disturbios.

En definitiva, la clave de estos modelos está en considerar la información pasada de la variable y su volatilidad observada como factor altamente explicativo de su comportamiento presente y, por extensión lógica, de su futuro predecible.

## 2.1. Hechos estilizados

Los hechos estilizados son una colección de observaciones empíricas e inferencias tomadas de estas observaciones que se pueden aplicar a la mayor parte de las series de tiempo diarias de cambios en factores de riesgo como son las series de log-rendimientos de índices, tasas de cambio, precios de commodities, entre otros.

También se pueden aplicar a series con otras temporalidades como semanales, mensuales o incluso intradía. Pero el hecho de que en una serie de datos se observen los hechos estilizados no quiere decir que al cambiar la temporalidad de la serie se sigan observando esos hechos; de hecho es muy probable que al cambiar la temporalidad ya no se observen los hechos estilizados. Una versión de estos hechos es la siguiente:

1. Los rendimientos no son independientes e idénticamente distribuidos, y muestran una pequeña correlación serial.
2. Los valores absolutos o las raíces de los rendimientos muestran una profunda correlación serial.
3. La esperanza condicional de los rendimientos es cercana a cero.
4. La volatilidad parece variar en el tiempo.
5. La serie de los rendimientos es leptocúrtica o de colas pesadas.
6. Los extremos de los rendimientos aparecen en *clusters*.

Un cluster es un grupo de observaciones en el que a las fuertes variaciones les siguen valores extremos, es decir, a grandes cambios les siguen grandes rendimientos (en valor absoluto).

Veamos un ejemplo de los clusters en una serie de tipo de cambio (tipo de cambio MXN/EUR) y la compararemos con una simulación de una serie Gaussiana.

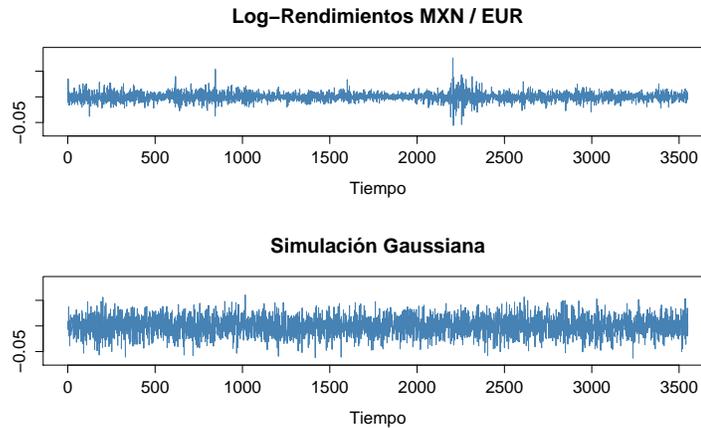


Figura 2.1: Simulación de un proceso Gaussiano comparado con la serie de log-rendimientos del tipo de cambio MXN/EUR.

Podemos ver cómo la serie financiera presenta momentos de gran volatilidad y otros donde es bastante pequeña, mientras que la simulación Gaussiana muestra una volatilidad constante. Además se observa que en los momentos en que la serie financiera tiene alta volatilidad es cuando se presentan sus valores más grandes.

Entonces para una serie que cumpla estas características podemos suponer que ajustar un modelo de volatilidad estocástica es lo más adecuado. A medida que aumente el periodo de los retornos (horizonte de tiempo en que medimos el rendimiento obtenido), al pasar de datos diarios a semanales, mensuales, trimestrales o anuales, los fenómenos que hemos identificado tienden a ser menos pronunciados.

## 2.2. Procesos ARCH

Como ya mencionamos en ENGLE, R. [7] se introducen los procesos ARCH (Autoregressive Conditional Heteroskedasticity, nombre dado por el mismo Engle), que definimos a continuación.

### Definición 11.

Sea  $Z_t$  un  $SWN(0, 1)$ . Decimos que  $X_t$  es **un proceso ARCH**( $p$ ) si es estrictamente estacionario y satisface (para algún proceso  $\sigma_t$ ):

$$X_t = \sigma_t Z_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2$$

para  $t \in \mathbb{Z}$ ,  $\alpha_0 > 0$  y  $\alpha_i \geq 0$ ,  $i = 1, 2, \dots, p$ .

Es importante que el proceso  $\sigma_t$  sea un proceso positivo ya que podemos notar fácilmente, de la definición de  $SWN(0, 1)$  y tomando nuevamente a  $\mathcal{F}_t = \sigma(X_s : s \leq t)$ , que

$$\mathbb{E}[X_t | \mathcal{F}_{t-1}] = \sigma_t \mathbb{E}[Z_t | \mathcal{F}_{t-1}] = \sigma_t \mathbb{E}[Z_t] = 0$$

$$\text{Var}[X_t | \mathcal{F}_{t-1}] = \sigma_t^2 \mathbb{E}[Z_t^2 | \mathcal{F}_{t-1}] = \sigma_t^2$$

Ya que la independencia entre  $Z_t$  y  $\mathcal{F}_{t-1}$  la podemos suponer directamente de la definición del proceso ARCH.

En R podemos hacer simulaciones de procesos ARCH. Para esto será necesario instalar la biblioteca **'bayesGARCH'** y la **'fGarch'**. Ahora, el código para la simulación (de un ARCH(2) con  $\alpha_1 = 0.2$ ,  $\alpha_2 = 0.4$  y  $\alpha_0 = 0$ ) es:

```
library(bayesGARCH)
```

```

library(fGarch)
spec = garchSpec(model=list(alpha=c(0.2,0.4), beta=0))
plot(garchSim(spec, n=150))

```

El resultado gráfico de esto es:

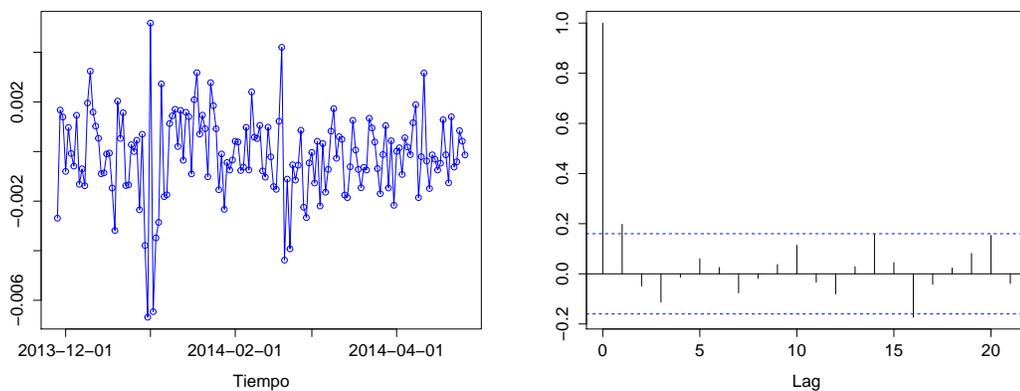


Figura 2.2: Simulación de un proceso ARCH(2) con  $\alpha_1 = 0.2$ ,  $\alpha_2 = 0.4$  y  $\alpha_0 = 0$  y su ACF

En este código genera un ARCH con innovaciones Gaussianas. Más adelante veremos cómo ajustar las innovaciones a nuestro gusto.

En esta simulación la especificación de los parámetros del modelo está dada en la construcción del vector *alpha* que, como su nombre lo indica, es el vector de parámetros  $\alpha_i$  del modelo. La cantidad de simulaciones está especificada en la *n*; para esta gráfica se realizaron 100, es decir, 100 tiempos.

### 2.2.1. ARCH(1)

El proceso ARCH(1) es muy importante porque es el primer modelo que surge de la familia ARCH y es la forma más simple de ellos, el modelo ARCH con menos parámetros.

En la Figura 2.3 podemos ver una simulación de un proceso ARCH(1), con  $\alpha_1 = 0.4$  y  $\alpha_0 = 0$ , para compararlo con el anterior. La escala de ambas gráficas es la misma.

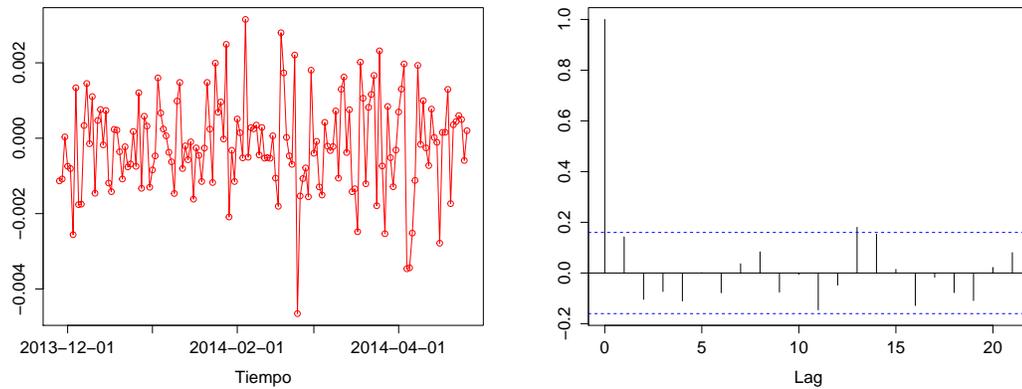


Figura 2.3: Simulación de un proceso ARCH(1) con  $\alpha_1 = 0.4$  y  $\alpha_0 = 0$  y su ACF

Ahora, usando la definición 8 y sustituyendo una ecuación en otra, tenemos que el proceso ARCH(1) puede ser escrito como:

$$X_t^2 = \alpha_0 Z_t^2 + \alpha_1 Z_t^2 X_{t-1}^2$$

Esta es una ecuación estocástica de recurrencia. Es importante notar que este modelo a pesar de ser estrictamente estacionario no es necesariamente un proceso

débilmente estacionario, pues se puede llegar a tener un proceso ARCH de varian-za infinita.

Una condición que se suele buscar (inferida de las ecuaciones estocásticas en recurrencia) para estos procesos es:

$$\mathbb{E}[\ln(\alpha_1 Z_t^2)] < 0$$

Esta es la condición para una solución estrictamente estacionaria de los ARCH(1) y se puede demostrar que es, de hecho, una condición necesaria y suficiente para la estacionariedad débil (ver Bougerol y Picard, 1992). La solución de la ecuación en recurrencia toma la forma:

$$X_t^2 = \alpha_0 \sum_{i=0}^{\infty} \alpha_1^i \prod_{j=0}^i Z_{t-j}^2$$

Si las innovaciones son Gaussianas estándar, la condición de estacionariedad se transforma, aproximadamente, en  $\alpha_1 < 3,562$  e incluso se puede decir que para innovaciones con distribución  $t$  con cuatro grados de libertad y con varianza uno la condición será  $\alpha < 5,437$ .

Como podemos ver de la ecuación  $\mathbb{E}[\ln(\alpha_1 Z_t^2)] < 0$ , la condición de estacionariedad estricta depende directamente de la distribución de las inovaciones; pero la estacionariedad débil, no.

La condición suficiente y necesaria para la estacionariedad débil, de un ARCH(1), es  $\alpha_1 < 1$ .

**Lema 2.**

*Un proceso ARCH(1) es débilmente estacionario si y sólo si  $\alpha_1 < 1$ .*

*Demostración.*

Primero supongamos que el proceso es covariantemente estacionario y que

$$\mathbb{E}[X_t^2] = 1$$

, entonces:

$$\sigma_t^2 = \mathbb{E}[X_t^2] = \alpha_0 + \alpha_1 \mathbb{E}[X_{t-1}^2] = \alpha_0 + \alpha_1 \sigma_t^2$$

de lo que se sigue que:

$$\sigma_t^2 = \frac{\alpha_0}{1 - \alpha_1}$$

y como  $\alpha_0 > 0$  y  $\sigma_t^2 > 0$  entonces

$$\alpha_1 < 1.$$

Ahora supongamos que  $\alpha_1 < 1$ , entonces

$$\mathbb{E}[\ln(\alpha_1 Z_t^2)] \leq \ln(\mathbb{E}[\alpha_1 Z_t^2]) = \ln(\alpha_1) < 0.$$

□

Otra de las propiedades que nos interesa de estos procesos es saber si tienen, o pueden llegar a tener, momentos finitos y cuántos tienen.

Veremos que bajo ciertas condiciones de las innovaciones podremos dar información de los momentos del proceso ARCH.

**Lema 3.**

*Sea  $m \geq 1$ . El proceso estrictamente estacionario ARCH(1) tiene momento  $2m$  finito si y sólo si*

$$\mathbb{E}[Z_t^{2m}] < \infty$$

y

$$\alpha_1 < (\mathbb{E}[Z_t^{2m}])^{-\frac{1}{m}}$$

Por ejemplo, para  $m = 2$ , en el caso Gaussiano estándar se necesita

$$\alpha_1 < \frac{1}{\sqrt{3}}$$

Y para el caso de una distribución  $t$  con 6 grados de libertad se necesitaría

$$\alpha_1 < \frac{1}{\sqrt{6}}$$

Por otro lado, si suponemos que el cuarto momento existe, es fácil calcularlo y, a partir de él, la Kurtosis del proceso,

$$\mathbb{E}[X_t^4] = \frac{\alpha_0^2 \mathbb{E}[Z_t^4] (1 - \alpha_1^2)}{(1 - \alpha_1)^2 (1 - \alpha_1^2 \mathbb{E}[Z_t^4])}$$

y entonces la Kurtosis será

$$\kappa_X = \frac{\mathbb{E}[X_t^4]}{\mathbb{E}[X_t^2]^2} = \frac{\kappa_Z (1 - \alpha_1^2)}{(1 - \alpha_1^2 \kappa_Z)}$$

Diremos entonces que una variable (o serie de tiempo) tiene distribución leptocúrtica cuando su Kurtosis sea más grande que la de una distribución Gaussiana estándar, que es de 3.

En algunas ocasiones se suele centrar la Kurtosis respecto a la de la Gaussiana estándar, tomando la Kurtosis como:

$$\kappa_X^* = \kappa_X - 3$$

En este caso se dice que tenemos una distribución leptocúrtica cuando su Kurtosis centrada es mayor a cero.

Ahora veremos la generalización del proceso ARCH.

### 2.3. Procesos GARCH

Como toda buena aportación matemática los modelos ARCH fueron atacados por una gran cantidad de matemáticos que buscaban mejorar el modelo y en algunos casos modificarlo para usarlo en problemas particulares donde el modelo funcionaría con un pequeño cambio.

Estas nuevas aportaciones al modelo generarían, en los siguientes años, una gran variedad de modelos de la familia de autorregresivos de heteroscedasticidad condicional.

En 1986 Tim Peter Bollerslev (BOLLERSLEV, T.P. [3]) hace una de las principales aportaciones a esta nueva familia de modelo al hacer una generalización del modelo de Engle.

La Figura 2.4 muestra una cronología las apariciones de algunos de los modelos de la familia ARCH, desde la aparición del ARCH en 1982 hasta 1996.

Como puede notarse, la familia de modelos autorregresivos de heteroscedasticidad condicional han crecido tremendamente en estos pocos años. En esta sección estudiaremos solamente al GARCH, una variante de él cuándo se combina con un proceso ARMA y un ejemplo de la aplicación de un GARCH para calcular medidas de riesgo a una serie financiera.

Comenzaremos por la definición, veremos el caso específico del GARCH(1,1) y estudiaremos algunas de sus propiedades más importantes que serán útiles para el estudio que haremos más adelante.

Los procesos GARCH (Generalized Autoregressive Conditional Heteroskedasticity) son definidos en BOLLERSLEV, T.P. [3].

<b>Año</b>	<b>Nombre</b>	<b>Autores</b>	<b>Aportación</b>
1982	ARCH	Engle	Primera aparición de un modelo de la familia ARCH.
1983	ARCH multivar.	Kraft y Engle	Incorporación de más variables y uso de la matriz de varianzas y covarianza en el modelo.
1986	ARCH-M	Engle, Lilien y Robins	Incorporación de la desviación estándar al modelo.
1986	GARCH	Bollerslev	Generalización del ARCH.
1986	LGARCH	Bollerslev y Taylor	GARCH lineal.
1986	MGARCH	Geweke y Pantula	Linealización con logaritmos.
1986	IGARCH	Engle y Bollerslev	Persistencia en varianza.
1989	EGARCH	Nelson	GARCH asimétrico con el uso de exponenciales.
1989	TS-GARCH	Schwert	Corrección de efectos asimétricos.
1992	T-GRCH	Gourieroux	Modelos dinámicos con la media y varianza funciones endógenas.
1993	GJR-GARCH	Glosten	Diferenciación del parámetro de subida y bajada.
1993	V-GARCH	Engle y Ng	Variación mayor en los parámetros asimétricos.
1993	A-PARCH	Ding	Modelar un valor potencial de la desviación estándar.
1995	Q-GARCH	Sentana	GARCH cuadrático.
1996	LST-GARCH	González-Rivera	Generalización del GLR-GARCH. Empleo de una ecuación de estado logística.

Figura 2.4: Algunos modelos de la familia ARCH y su fecha de aparición.

**Definición 12.**

Sea  $Z_t$  un  $SWN(0, 1)$ . Decimos que  $X_t$  es **un proceso GARCH**( $p, q$ ) si es estrictamente estacionario y satisface (para algún proceso  $\sigma_t$ ):

$$X_t = \sigma_t Z_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2$$

para  $t \in \mathbb{Z}$ ,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $i = 1, 2, \dots, p$  y  $\beta_j \geq 0$ ,  $j = 1, 2, \dots, q$

Los procesos GARCH son procesos ARCH generalizados; ahora, la volatilidad en algún tiempo va a depender de las volatilidades anteriores, así como de los valores anteriores del proceso (como el caso del ARCH).

**2.3.1. GARCH(1, 1)**

En la práctica los modelos GARCH de orden inferior (de parámetros pequeños) son los más usados ya que, como se mencionó antes, estos modelos no aportan mucha más información al aumentar el orden de los parámetros y además el buscar ajustar un modelo de algún orden grande aplicará también una mayor cantidad de valores que deben estimarse.

Supongamos que tenemos un modelo GARCH(1, 1). Directamente de la definición podemos ver que su volatilidad tendrá la forma:

$$\sigma_t^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta) \sigma_{t-1}^2$$

Nuevamente tenemos un sistema de ecuaciones recursivas, pero ahora la condición de estacionariedad se transforma en:

$$\mathbb{E}[\ln(\alpha_1 Z_t^2 + \beta)] < 0$$

y su solución general esta dada por:

$$\sigma_t^2 = \alpha_0 \left( 1 + \sum_{i=1}^{\infty} \prod_{j=1}^i (\alpha_1 Z_{t-j}^2 + \beta) \right)$$

Con esto, si el proceso  $(\sigma_t^2)$  es estrictamente estacionario, entonces también lo será el proceso GARCH. La solución del GARCH define la siguiente forma:

$$X_t = Z_t \sqrt{\alpha_0 \left( 1 + \sum_{i=1}^{\infty} \prod_{j=1}^i (\alpha_1 Z_{t-j}^2 + \beta) \right)}$$

tenemos un resultado que nos facilita saber si el proceso es estacionario o no, solo necesitamos conocer los parámetros del mismo.

**Lema 4.**

*Un proceso GARCH(1,1) es un ruido blanco débilmente estacionario si y sólo si*

$$\alpha_1 + \beta < 1$$

*La varianza de este proceso sería entonces*  $\text{Var}[X_t] = \frac{\alpha_0}{1 - \alpha_1 - \beta}$ .

Haremos un proceso similar al que hicimos con ARCH para encontrar la Kurtosis de un GARCH.

Podemos concluir que una condición suficiente y necesaria para la existencia del cuarto momento del proceso es

$$\mathbb{E}[(\alpha_1 Z_t^2 + \beta)^2] < 1$$

o equivalentemente que

$$(\alpha_1 + \beta)^2 < 1 - (\kappa_Z - 1)\alpha_1^2$$

Ahora, si suponemos la existencia del cuarto momento del proceso, podemos calcularlo fácilmente

$$\mathbb{E}[X_t^4] = \frac{\alpha_0^2 \kappa_Z (1 - (\alpha_1 + \beta)^2)}{(1 - \alpha_1 - \beta)^2 (1 - \alpha_1^2 \kappa_Z - \beta^2 - 2\alpha_1 \beta)}$$

De lo que se concluye que la Kurtosis del proceso es:

$$\kappa_X = \frac{\kappa_Z (1 - (\alpha_1 + \beta)^2)}{1 - (\alpha_1 + \beta)^2 - (\kappa_Z - 1)\alpha_1^2}$$

Generalizando esto, una condición suficiente y necesaria para la estacionariedad en covarianza de un proceso GARCH( $p, q$ ) es:

$$\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$$

En R también podemos simular este modelo.

```
spec=garchSpec(model=list(omega=.5,alpha=.4,beta=.2))
x<-garchSim(spec, n = 150)
plot(acf(x))
plot(x)
```

El resultado gráfico de esta simulación se muestra en la figura 2.5, donde también se observa la ACF de proceso que ilustra la propiedad que nos da el Lema 4.

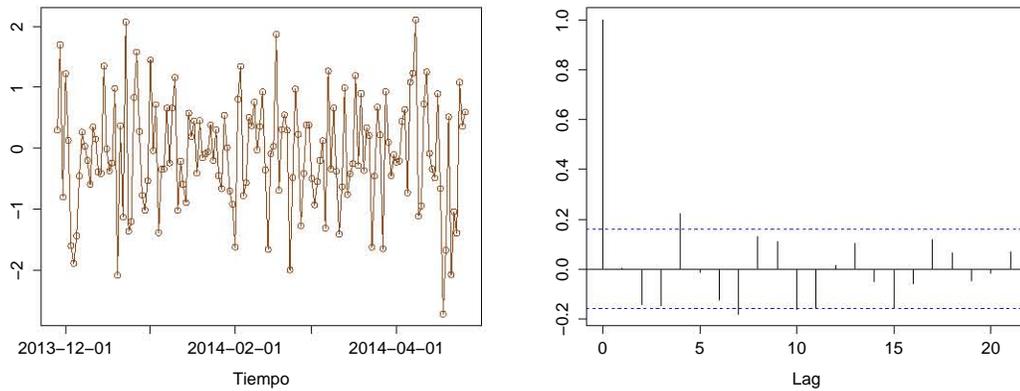


Figura 2.5: Simulación de un proceso GARCH(1,1) con  $\alpha_1 = 0.4$ ,  $\beta_1 = 0.2$  y  $\alpha_0 = 0.5$  y su ACF.

Ahora veremos una extensión simple del proceso GARCH al unirlo con la teoría de los procesos ARMA. Esto gracias a la información que obtuvimos del Lema 4 de cuando podemos considerar a un GARCH como un ruido blanco estrictamente estacionario, y posteriormente veremos cómo ajustar estos modelos a datos.

### 2.3.2. Una extensión del proceso GARCH

Se han propuesto muchas variantes y extensiones del modelo básico GARCH, como ya vimos en la figura 2.4. Aquí mencionamos sólo una de ellas para ejemplificar. El proceso que veremos solo es una combinación de dos modelos que ya estudiamos.

El proceso ARMA–GARCH es el modelo que nos permite combinar las características de ambos procesos en uno solo.

**Definición 13.**

Decimos que  $\{X_t\}$  es un proceso  $ARMA(p_1, q_1)$ – $GARCH(p_2, q_2)$  si es débilmente estacionario y cumple:

$$X_t = \sum_{i=1}^{p_1} \phi_i X_{t-i} + \epsilon_t + \sum_{j=1}^{q_1} \theta_j \epsilon_{t-j}$$

$$\epsilon_t = \sigma_t Z_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p_2} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q_2} \beta_j \sigma_{t-j}^2$$

para  $t \in \mathbb{Z}$ ,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $i = 1, 2, \dots, p_2$  y  $\beta_j \geq 0$ ,  $j = 1, 2, \dots, q_2$  y con la condición  $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$  y  $(Z_t)_{t \in \mathbb{Z}}$  un  $SWN(0, 1)$ .

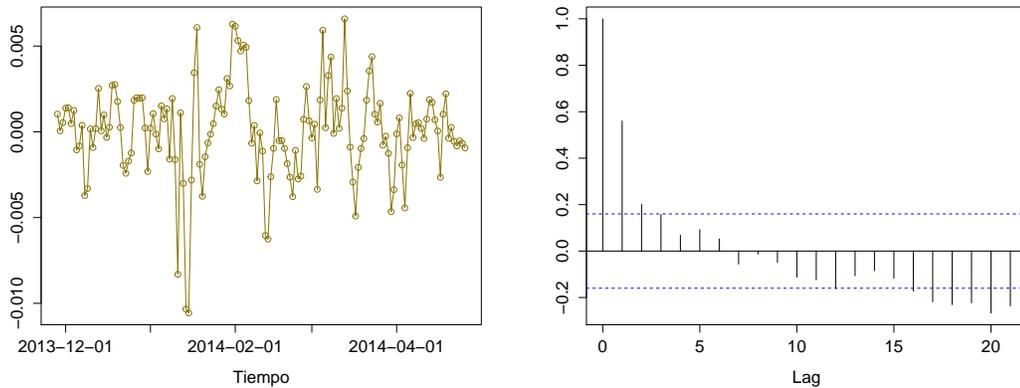


Figura 2.6: Simulación de un proceso  $ARMA(1, 2)$ – $GARCH(1, 1)$  con  $\alpha_1 = 0.4$ ,  $\beta_1 = 0.4$ ,  $\alpha_0 = 0.0$ ,  $\phi = 0.5$ ,  $\theta_1 = 0.3$  y  $\theta_2 = -0.3$  y su ACF.

En R será muy fácil simular este modelo. El código para la simulación es:

```
s=garchSpec(model=list(alpha=.4, beta=.4, ar=.5, ma=c(.3, -.3)))
plot(garchSim(s, n = 150), main="ARMA(1,2)-GARCH(1,1)")
```

La gráfica de esta simulación se puede observar en la figura 2.6.

## 2.4. Ajuste de un modelo GARCH

Veremos ahora cómo se comporta el modelo que hemos desarrollado al ajustarlo a datos reales. Utilizaremos datos financieros, tasas de cambio de monedas; estos tipos de datos suelen cumplir los hechos estilizados, y por eso son buenos candidatos a modelarse con procesos tipo GARCH.

Antes de realizar el ajuste vamos a revisar la estructura teórica de cómo se realiza el ajuste de los modelos GARCH a datos.

### 2.4.1. Estimación de parámetros

Como ya mencionamos en el Capítulo 1 la estimación de los parámetros de una serie de tiempo se suele realizar por máxima verosimilitud. Supongamos, nuevamente, que tenemos datos en un vector  $\mathbb{X} = (x_1, \dots, x_n)$ .

Vamos a considerar que queremos estimar los parámetros de un modelo GARCH(1, 1). Entonces  $\Theta = (\alpha_0, \alpha_1, \beta_1)$  y entonces, con la definición del modelo, la función de verosimilitud del modelo es:

$$L(\Theta, \mathbb{X}) = \prod_{t=1}^n \frac{1}{\sigma_t} g\left(\frac{x_t}{\sigma_t}\right)$$

con

$$\sigma_t = \sqrt{\alpha_0 + \alpha_1 x_{t-1} + \beta_1 \sigma_{t-1}^2}$$

Donde la función  $g(x)$  es la función de densidad de las innovaciones del modelo. Esta función también podría tener parámetros a estimar, en ese caso los parámetros que la distribución de las innovaciones también serán parte de las variables a maximizar.

El vector  $\Theta$  que estamos buscando es el que maximice la verosimilitud, la maximización puede hacerse directamente con la función de verosimilitud o con la log-verosimilitud,  $\ln(L(\Theta, \mathbb{X}))$ , con el fin de facilitar las cosas.

Entonces, el vector de parámetros estimados,  $\hat{\Theta}$ , satisface la ecuación:

$$\frac{\partial}{\partial \Theta} \ln(L(\hat{\Theta}, \mathbb{X})) = 0$$

Para hacer esta maximización se utilizan métodos numéricos que son modificación del método de Newton-Raphson. En particular se suele usar un método llamado BHHH.

Ahora haremos la estimación de los parámetros de una serie de tiempo financiera que suponemos se comporta como un GARCH. En este caso usaremos la información de la serie del tipo de cambio del peso-euro.

### 2.4.2. Tipo de cambio MXN/EUR

El tipo de cambio es el precio de una moneda en términos de otra. Lo expresamos habitualmente como el número de unidades de una moneda que hay que entregar a cambio de una unidad de la otra moneda. En el caso de México es la equivalencia del peso con respecto a la moneda extranjera.

El comercio internacional y la mayor parte de los movimientos registrados en los mercados financieros internacionales están constituidos por el mercado de divisas que es llamado *Forex*. El *Forex* es un mercado donde se realiza la compra-venta

de divisas buscando la mayor ganancia posible dada por la diferencia de precios en el tiempo, es decir, es un mercado de variaciones (de volatilidad).

En el Forex se miden las ganancias o pérdidas con los rendimientos o los log-rendimientos; estos valores son usados como referencia en cualquier estudio de pérdidas y ganancias. Se obtienen de la siguiente forma: supongamos que tenemos  $x_0, x_1, \dots, x_n$ , los datos históricos de algún tipo de cambio. Entonces, definimos:

1. Los rendimientos absolutos:  $y_i = x_i - x_{i-1}$
2. Los rendimientos:  $R_i = \frac{x_i - x_{i-1}}{x_{i-1}}$
3. Los log-rendimientos:  $r_i = \ln \left( \frac{x_i}{x_{i-1}} \right)$

Los log-rendimientos o *log-returns* es la medida de ganancia más usual en los mercados financieros.

Ahora estudiaremos el tipo de cambio del Euro respecto del Peso mexicano para evaluar si es conveniente ajustar un modelo de volatilidad estocástica y después (en todo caso) ajustárselo.

## El Euro

El Euro tiene sus inicios con el Tratado de la Unión Europea, también conocido como el Tratado de Maastricht, que entró en vigor desde 1993. Entre sus fines este acuerdo tenía la intención de crear una unión económica y monetaria que sería la culminación del proceso de convergencia de las políticas económicas de los estados miembros.

Al principio la moneda acordada sería la Unidad Monetaria Europea (ECU); pero los estados miembros de la Unión Europea acordaron en Madrid el 15 de

diciembre de 1995, la creación de una moneda común europea, bajo la denominación de "euro", con fecha de inicio de circulación el primero de enero del 2002 para los doce estados entonces.

El primer día de cotización de la nueva moneda europea con el Peso mexicano (a tipo de cambio FIX <sup>1</sup>), el 4 de enero de 1999 (cuando aún era la Unidad Monetaria Europea), 1 euro se cambiaba por 11.6054.

En julio de 2002, el euro sobrepasó la paridad con el dólar en el mercado de divisas, y se ha mantenido así. El 5 de octubre de 2009, el euro alcanzó su valor máximo, respecto al peso, hasta el momento, al cambiarse 1 euro por 19.9745 pesos.

La figura 2.6 es la gráfica del tipo de cambio MXN/EUR desde 2000 a la fecha y la serie de los log-rendimientos del tipo de cambio MXN/EUR en el mismo periodo.

Las líneas azules son  $(-0.08, 0.08)$ . Podemos ver que en el momento del máximo histórico se observa una volatilidad tremenda. Este periodo corresponde a las fluctuaciones de 2008 y sus secuelas hasta 2009.

También podemos ver en la figura 2.8 y la tabla 2.1 un análisis exploratorio de los datos. Primero vemos su histograma y calculemos sus variables descriptivas. Este análisis exploratorio se realiza con la intención de exhibir los hechos estilizados en la serie.

---

<sup>1</sup>El tipo de cambio FIX es determinado por el Banco de México con base en un promedio de cotizaciones del mercado de cambios al mayoreo para operaciones liquidables el segundo día hábil bancario siguiente y que son obtenidas de plataformas de transacción cambiaria y otros medios electrónicos con representatividad en el mercado de cambios.

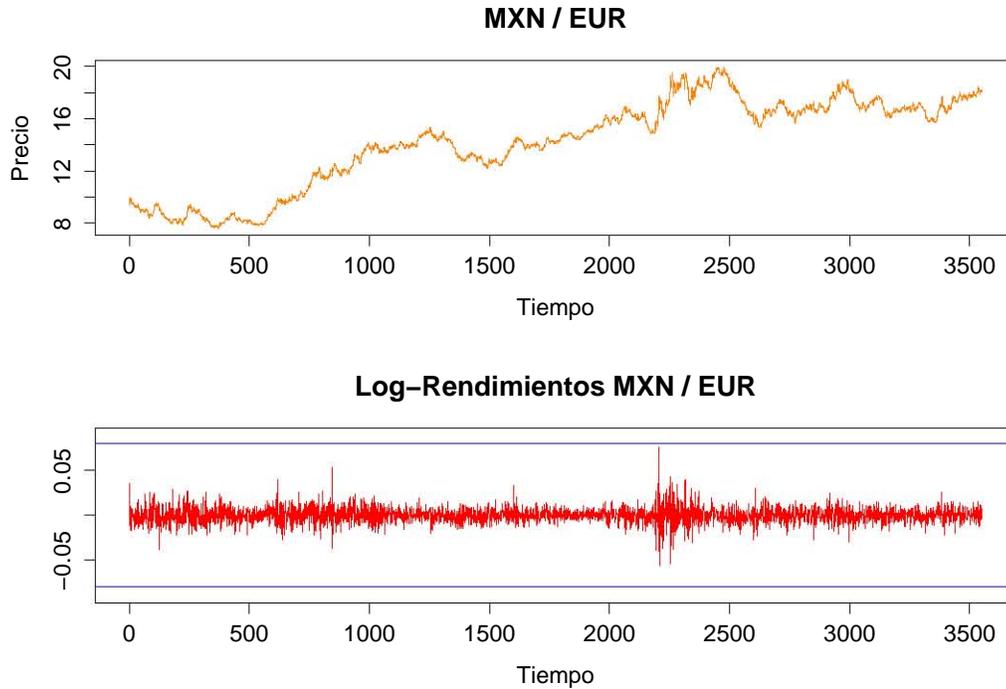


Figura 2.7: Tipo de cambio MXN/EUR desde 2000 y sus log-rendimientos

Observamos que la media es muy cercana a cero y que la kurtosis resulta mucho mayor a 3. Esto nos permite suponer que la distribución de los log-rendimientos es leptocúrtica. También vemos que la varianza muestral es muy cercana a cero, lo cual no tiene mucho sentido cuando observamos los puntos de alta volatilidad. Esa varianza no corresponde a la serie en esos momentos y por lo mismo no podemos suponer una varianza constante en el tiempo.

Veamos que un QQplot de los log-rendimientos nos muestra que no sería buena idea ajustar a los datos una normal. El valor que nos dio la Kurtosis nos dice que

Cuadro 2.1: Medidas descriptivas

Media	Kurtosis	Varianza
0.000180	7.80407	0.000072

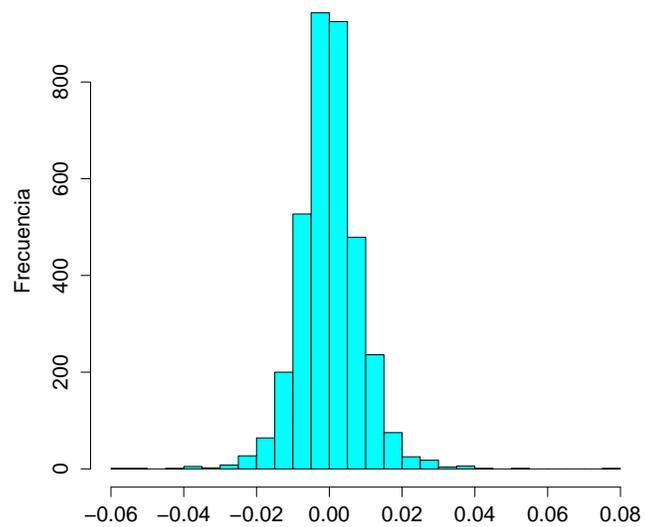


Figura 2.8: Histograma de los log-rendimientos del tipo de cambio MXN/EUR

las colas de la distribución son mucho más pesadas que las de una distribución normal y esto se puede ver fácilmente en el QQplot. El código para generar este QQplot es el siguiente:

```
qqnorm(x)
qqline(x,col="red")
qqplot(rt(length(x),df=3), x)
```

```
qqline(x,col="red")
```

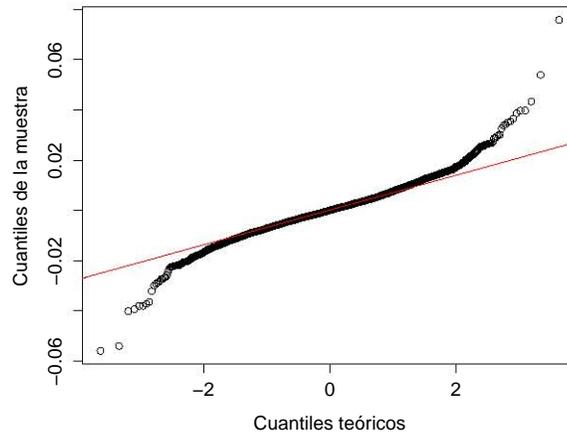


Figura 2.9: QQplot de los log-rendimientos del tipo de cambio MXN/EUR contra una Normal estándar

Con todo esto podemos ver que se cumplen gran parte de los hechos estilizados. Solo resta ver los primeros dos. Entonces calcularemos la ACF de los log-rendimientos, la de los cuadrados de estos y de sus valores absolutos.

En este caso veremos las AFC para distancias de cero a cuarenta; este parámetro se puede modificar, pero aumentar demasiado las distancias en los correlogramas puede causar problemas, ya que tendremos cada vez menos datos para hacer las estimaciones.

Podemos ver que la correlación de los log-rendimientos es muy pequeña para todas distancias, prácticamente cero (obviamente, excepto en 1).

Para los valores absolutos podemos ver una correlación cercana a 0.2 para

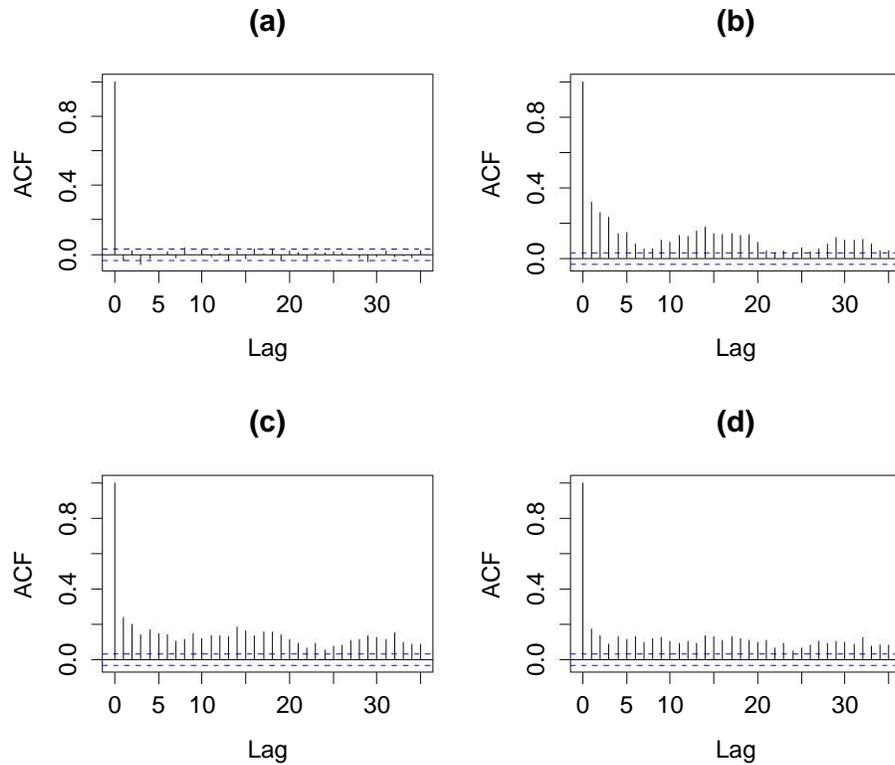


Figura 2.10: (a)ACF de los log-rendimientos del tipo de cambio MXN/EUR. (b)ACF de los cuadrados de los log-rendimientos del tipo de cambio MXN/EUR. (c)ACF de los valores absolutos de los log-rendimientos del tipo de cambio MXN/EUR. (d)ACF de las raices de los valores absolutos de los log-rendimientos del tipo de cambio MXN/EUR.

distancias, incluso mayores a 20. Para el caso de los cuadrados tenemos una correlación más fuerte al inicio, cercana a 0.3 antes de la distancia 10.

Ahora buscaremos ajustar un modelo de volatilidad estocástica. El modelo con mejor ajuste obtenido es el GARCH(1,1) con inovaciones  $t$ . Aquí presentamos los resultados.

**Ajuste de un GARCH(1,1) con innovaciones  $t$ .** El código es el siguiente, notemos que no es necesario especificar los grados de libertad de la  $t$  ya que R estimará también este parámetro. Usaremos la biblioteca '*fGarch*'.

```
library(fGarch)
x<-read.table("clipboard")
x<-ts(x)
f=garchFit(f= ~garch(1,1),d=x,cond.dist="std",trace=F)
plot(f)
```

El resultado del ajuste es:

Cuadro 2.2: Parámetros estimados				
	Estimación	Error estándar	t value	P[> t ]
$\mu$	0.000093	0.000116	0.803	0.4217
$\alpha_0$	0.0000005	0.00000018	2.834	0.0045
$\alpha_1$	0.04508	0.00717	6.288	$\approx 0$
$\beta_1$	0.9471	0.008422	112.452	$\approx 0$
$\nu$	9.48	1.358	6.986	$\approx 0$

Tenemos un proceso débilmente estacionario, la estimación que hizo R de las innovaciones fue una  $t$  con 9.48 grados de libertad.

En la Figura 2.11 podemos ver las características de los residuales del ajuste. Los residuales tienen una excelente forma en este ajuste. El QQplot respecto de una distribución  $t$  queda perfecto salvo por uno o dos datos ligeramente distanciados.

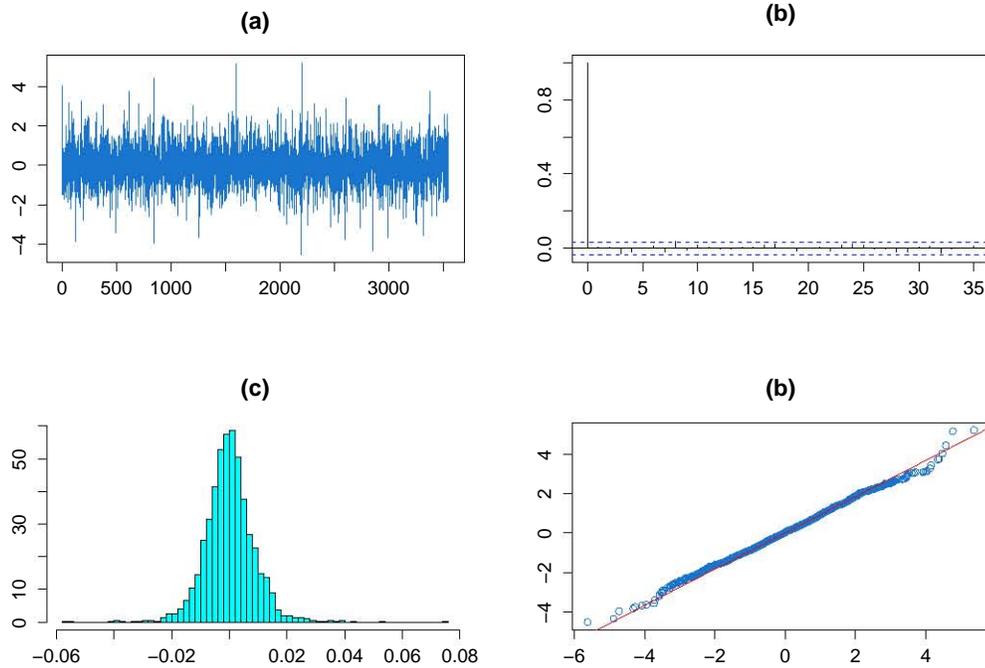


Figura 2.11: (a)Residuales del ajuste. (b)ACF de los residuales del ajuste. (c)Histograma de los residuales del ajuste. (d)QQplot de los residuales contra un distribución  $t$ .

Este ajuste es bastante satisfactorio por lo que sobre él hacemos la estimación, primero calculamos una estimación de la volatilidad, esto es de la forma:

$$\hat{\sigma}_{t+1}^2 = \hat{\alpha}_0 + \hat{\alpha}_1 X_t^2 + \hat{\beta}_1 \hat{\sigma}_t^2$$

Para calcular esto necesitaremos una varianza inicial que haremos cero, esta es una convención común en la práctica, se pueden tener diferencias cuando iniciamos con algún otro valor inicial, pero es una buena práctica iniciar con el valor de cero.

En este caso tenemos 3550 datos así que buscaremos la varianza estimada para

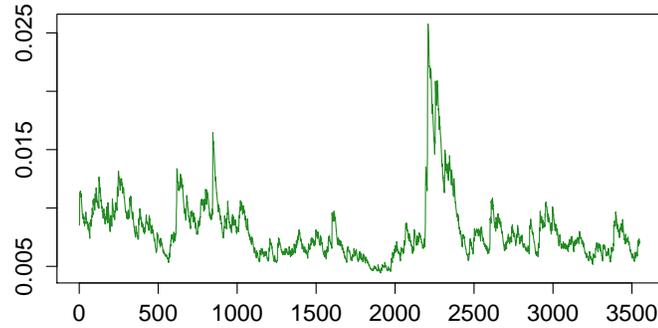


Figura 2.12: Desviación estándar estimada de los log-rendimientos del tipo de cambio MXN/EUR .

el día 3551, partiendo de que la varianza el día uno es cero, entonces:

$$\hat{\sigma}_{3551}^2 = 0.00004726$$

### 2.4.3. Estimación de medidas de riesgo

Para calcular la medidas de riesgo para un modelo GARCH vamos a conciderar que nuestra variable aleatoria de pérdida (ver A.2),  $L_t$ , está dada por un modelo GARCH con innovaciones  $Z_t$  que son un SWN( $\mu, \sigma$ ). Entonces:

$$L_t = \sigma_t Z_t$$

Si suponemos que la función de distribución de  $Z_t$  es  $G(x)$  tendrmos que la distribución de pérdida en el tiempo  $t + 1$  dado todo lo acurrido hasta  $t$  será:

$$F_{L_{t+1}|\mathcal{F}_t}(l) = P[\sigma_{t+1}Z_{t+1} \leq l|\mathcal{F}_t] = G\left(\frac{l}{\sigma_{t+1}}\right)$$

Usando esto tendremos que el VaR y el ES de la variable de pérdida son:

$$\begin{aligned}\text{VaR}_\alpha(L_{t+1}) &= \sigma_{t+1} * \text{VaR}_\alpha(Z) \\ \text{ES}_\alpha(L_{t+1}) &= \sigma_{t+1} * \text{ES}_\alpha(Z)\end{aligned}$$

Para  $\left(\frac{Z - \mu}{\sigma}\right) \sim t(\nu)$  (con media cero y varianza 1) podemos calcular ambas medidas de riesgo usando la densidad y distribución  $t$ .

$$\begin{aligned}\text{VaR}_\alpha(Z) &= \mu + \sigma t_\nu^{-1}(\alpha) \\ \text{ES}_\alpha(Z) &= \left(\frac{g_\nu(t_\nu^{-1}(\alpha))}{1 - \alpha}\right) \left(\frac{\nu + (t_\nu^{-1}(\alpha))^2}{\nu - 1}\right)\end{aligned}$$

Para realizar este cálculo solo se requiere la definición dada en A.2 y sustituir con la función de distribución de una  $t$  estándar.

Entonces el estimado del Valor en Riesgo y del Expected-shortfall de los log-rendimientos será:

$$\begin{aligned}\text{VaR}_\alpha(X_{3551}) &= \hat{\sigma}_{3551} * \text{VaR}_\alpha(Z) \\ \text{ES}_\alpha(X_{3551}) &= \hat{\sigma}_{3551} * \text{ES}_\alpha(Z)\end{aligned}$$

Para un nivel de confianza del 90 % y 95 %, usando las formulas anteriores y sustituyendo los grados de libertad estimados para las innovaciones  $t$ , tenemos el cálculo de las medidas de riesgo en la tabla 2.3.

Cuadro 2.3: Riesgo de los log-rendimientos

$\text{VaR}_{90\%}$	=	0.00947
$\text{ES}_{90\%}$	=	0.01378
$\text{VaR}_{95\%}$	=	0.01253
$\text{ES}_{95\%}$	=	0.01671

Además de este ajuste hemos utilizado el modelo GARCH para calcular medidas de riesgo a otras series de tiempo de log-rendimientos de otras series financieras obteniendo buenos resultados como se puede ver en la tesis "Modelos de volatilidad estocástica y su aplicación a la administración de riesgos financieros" de Jonathan Martínez Martínez quien fue mi compañero en el curso "Seminario de Matemáticas Actuariales Aplicadas" donde, bajo la dirección de la Dra. Begoña Fernández Fernández, estudiamos los modelos condicionalmente heterocedasticas. Estas series parecen reaccionar bien a ajusten de modelos GARCH con innovaciones  $t$ .

## Capítulo 3

# La memoria de un proceso

En este capítulo daremos las bases para definir los procesos de memoria larga. Comenzaremos por nombrar y desarrollar los conceptos que formarán la definición de memoria de los procesos estocásticos.

El concepto fundamental es la densidad espectral pero solo nos concentraremos en estudiarla en cero ya que su comportamiento en ese punto caracterizará a la memoria de la serie.

Veremos una forma de estimar a la densidad espectral, las propiedades de este estimador y cómo obtenerlo en  $\mathbb{R}$  a partir de datos.

Daremos un panorama general de memoria definiendo los 3 tipos de esta: corta, larga y de antipersistencia. Estudiaremos un poco a los procesos con memoria corta y veremos que los procesos que hemos visto hasta ahora son de esta familia.

Finalmente definiremos, parametrizaremos y estudiaremos lo que significa que un proceso tenga memoria larga. Desarrollaremos las características del parámetro de memoria y encontraremos estimadores para él de los cuales veremos sus características asintóticas y la forma de calcularlos en  $\mathbb{R}$  a partir de datos.

El contenido de este capítulo está basado principalmente en PALMA, W. [14] y ROBINSON, P. [16].

También se tomó como referencia: GRANGER C.W.J. y JOYEUX R. [9], GEWEKE, J. y PORTER-HUDAK S. [8], BREIDT, F.J., CRATO, N. y DE LIMA P. [4], CHRONOPOULOU, A. y VIENS, F.G. [6], GUEGAN, D. [10], QIAN, B. y RASHEED, K. [15] y WILKINS, N. [19].

### 3.1. Análisis espectral

El análisis espectral de los procesos estocásticos nos permite conocer sus características periódicas. No desarrollaremos mucho este tema, solo nos interesaremos en un par de conceptos que surgen de él.

#### 3.1.1. Densidad espectral y periodograma

Los siguientes dos conceptos son de gran importancia en el estudio de la memoria de los procesos.

**Definición 14.**

*Dadas las observaciones  $x_t, t = 1, \dots, n$ , de una serie de tiempo. Definimos a la transformada discreta de Fourier como:*

$$d_n(\nu_j) = \frac{1}{\sqrt{n}} \sum_{t=1}^n x_t e^{-i\nu_j t},$$

$$\text{con } \nu_j = \frac{2\pi j}{n}, \quad j = 0, 1, 2, \dots, (n-1).$$

Los coeficientes  $\nu_j$  se conocen como las frecuencias de Fourier.

**Definición 15.**

Sea  $\{X_t\}_{t \in \mathbb{Z}}$  una serie de tiempo covariantemente estacionaria, con función de autocovarianza  $\gamma_{(k)}$ . La **densidad espectral** o el espectro de  $\{X_t\}_{t \in \mathbb{Z}}$  es:

$$f(\nu) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_{(k)} e^{-ik\nu}$$

para  $-\pi < \nu \leq \pi$ .

Con ayuda de la fórmula de Euler

$$e^{ix} = \cos(x) + i \sin(x)$$

tendremos:

$$\begin{aligned} f(\nu) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_{(k)} e^{-ik\nu} \\ &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_{(k)} \cos(k\nu) \\ &= \frac{1}{2\pi} \gamma_{(0)} + \frac{1}{1\pi} \sum_{k=0}^{\infty} \gamma_{(k)} \cos(k\nu) \end{aligned}$$

para  $-\pi < \nu \leq \pi$ .

Existe otra forma de definir la densidad espectral de una serie de tiempo. Esta otra forma define al espectro de la siguiente forma, si existe una función  $f(\nu)$  definida en  $[-\pi, \pi]$  tal que su anti-trasformada de Fourier es la función de autocorrelación entonces, dicha función es la densidad espectral de la serie. Esto es:

**Lema 5.** Otra definición de la densidad espectral.

Sea  $\{X_t\}_{t \in \mathbb{Z}}$  una serie de tiempo covariantemente estacionaria, con función de autocovarianza  $\gamma_{(k)}$  y densidad espectral  $f(\nu)$ . Entonces:

$$\gamma_{(h)} = \int_{-\pi}^{\pi} e^{ih\nu} f(\nu) d\nu$$

***Demostración.***

Primero notemos que:

$$\int_{-\pi}^{\pi} e^{i\nu(h-k)} d\nu = \begin{cases} 2\pi, & \text{si } h = k \\ 0, & \text{si } h \neq k \end{cases}$$

Para el caso  $h = k$  es evidente la igualdad. Para el caso  $h \neq k$  tenemos:

$$\begin{aligned} \int_{-\pi}^{\pi} e^{i\nu(h-k)} d\nu &= \frac{e^{i\nu(h-k)}}{i(h-k)} \Big|_{\nu=-\pi}^{\nu=\pi} \\ &= \frac{\cos(\nu(h-k)) + i \sin(\nu(h-k))}{i(h-k)} \Big|_{\nu=-\pi}^{\nu=\pi} \\ &= \frac{(1-1) + i(0-0)}{i(h-k)} \\ &= 0 \end{aligned}$$

Con lo que tenemos:

$$\begin{aligned} \int_{-\pi}^{\pi} e^{ih\nu} f(\nu) d\nu &= \int_{-\pi}^{\pi} e^{ih\nu} \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_{(k)} e^{-ik\nu} d\nu \\ &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_{(k)} \int_{-\pi}^{\pi} e^{i\nu(h-k)} d\nu \\ &= \frac{1}{2\pi} \gamma_{(k)} \int_{-\pi}^{\pi} e^{i\nu(k-k)} d\nu \\ &= \gamma_{(k)} \end{aligned}$$

□

### Propiedades de la densidad espectral

Las primeras propiedades importantes de la densidad espectral son que es una función periódica con periodo  $2\pi$  y que es una función simétrica.

$$\begin{aligned} f(\nu) &= f(\nu + 2\pi) \\ &= f(-\nu) \end{aligned}$$

Por estas dos propiedades es que normalmente se presenta la gráfica de una densidad espectral solamente en  $0 \leq \nu \leq \pi$ .

### La densidad espectral de la suma de procesos independientes

Para otra propiedad necesitamos definir una generalización de la función de autocovarianza.

#### Definición 16.

Sea  $\{X_t\}$  débilmente estacionario, con autocovarianza  $\gamma_{(k)}$ . Definimos a la **función generadora de autocovarianza** como:

$$\gamma(B) = \sum_{k=-\infty}^{\infty} \gamma_{(k)} B^k$$

Con esta función podemos, fácilmente, ver que la densidad espectral de un proceso será:

$$f(\nu) = \frac{1}{2\pi} \gamma(e^{-i\nu}) \quad (3.1)$$

Consideremos al proceso  $Z_t = Y_t + X_t$ , con  $X_t$  y  $Y_t$  independientes y débilmente

estacionarios. Entonces:

$$\begin{aligned}
 \gamma_Z(B) &= \sum_{k=-\infty}^{\infty} \gamma_{Z(k)} B^k \\
 &= \sum_{k=-\infty}^{\infty} \text{Cov}(Y_t + X_t, Y_{t+k} + X_{t+k}) B^k \\
 &= \sum_{k=-\infty}^{\infty} [\text{Cov}(X_t, X_{t+k}) + \text{Cov}(Y_t, Y_{t+k})] B^k \\
 &= \sum_{k=-\infty}^{\infty} \gamma_{X(k)} B^k + \sum_{k=-\infty}^{\infty} \gamma_{Y(k)} B^k \\
 &= \gamma_X(B) + \gamma_Y(B)
 \end{aligned}$$

Con esto y con (3.1) tenemos que:

$$f_X(\nu) = f_X(\nu) + f_Y(\nu)$$

Ahora, el periodograma será el cuadrado de la transformada de Fourier.

**Definición 17.**

Supongamos que tenemos las observaciones  $x_t$ ,  $t = 1, 2, \dots, n$ , de alguna serie de tiempo. Definimos al **periodograma** de dichas observaciones como:

$$I_n(\nu_j) = (d_n(\nu_j))^2$$

con  $\nu_j = \frac{2\pi j}{n}$ ,  $j = 0, 1, 2, \dots, (n-1)$ .

En la práctica no es posible tener todos los valores de la función de autocorrelación por lo que se utiliza al periodograma como una aproximación de la densidad espectral. Esta aproximación es posible ya que podemos reescribir al periodograma de otra forma más conveniente.

Primero notemos que  $I_n(0) = n\bar{x}^2$ , donde  $\bar{x}$  es la media muestral, también podemos ver que para  $j \neq 0$ :

$$\begin{aligned} \sum_{t=1}^n e^{(-iv_j)t} &= e^{-iv_j} \left( \frac{1 - e^{-2\pi ji}}{1 - e^{-iv_j}} \right) \\ &= e^{-iv_j} \left( \frac{1 - \cos(2j\pi) + i \operatorname{sen}(2j\pi)}{1 - e^{-iv_j}} \right) \\ &= e^{-iv_j} \left( \frac{1 - 1 + 0}{1 - e^{-iv_j}} \right) \\ &= 0 \end{aligned}$$

Por lo que tenemos que:

$$d_n(\nu_j) = \frac{1}{\sqrt{n}} \sum_{t=1}^n (x_t - \bar{x}) e^{-iv_j t}, \quad j \neq 0.$$

Con esta relación podemos escribir al periodograma, para  $j \neq 0$ , como:

$$\begin{aligned} I_n(\nu_j) &= \frac{1}{n} \sum_{t=1}^n \sum_{s=1}^n (x_t - \bar{x})(x_s - \bar{x}) e^{-iv_j(t-s)} \\ &= \frac{1}{n} \sum_{h=-(n-1)}^{n-1} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}) e^{-iv_j(h)} \\ &= \sum_{h=-(n-1)}^{n-1} \hat{\gamma}(h) e^{-iv_j h} \end{aligned}$$

Esta última relación nos muestra que el periodograma se puede ver como una estimación de la densidad espectral. El siguiente resultado nos permite dar intervalos de confianza sobre la densidad espectral en algún punto.

**Teorema 1.**

Sea  $\{X_t\}_{t \in \mathbb{N}}$  una serie de tiempo causal, es decir, cumple (1.1) y (1.2). Si  $f(\nu_j) > 0$  Entonces:

$$\frac{2I_n(\nu_j)}{f(\nu_j)} \longrightarrow_d \chi_{(2)}^2$$

Ahora construiremos el intervalo de confianza. Sea  $\chi_{(2)}^2(\alpha)$  el cuantil  $\alpha$  de una  $\chi_{(2)}^2$ , es decir, si  $X \sim \chi_{(2)}^2$  entonces:

$$P[X \leq \chi_{(2)}^2(\alpha)] = \alpha.$$

Entonces, con un nivel de confianza del  $100(1 - \alpha)\%$  de confianza podemos escribir que:

$$\frac{2I_n(\nu_j)}{\chi_{(2)}^2\left(1 - \frac{\alpha}{2}\right)} \leq f(\nu_j) \leq \frac{2I_n(\nu_j)}{\chi_{(2)}^2\left(\frac{\alpha}{2}\right)}$$

Estos resultados nos servirán cuando veamos los modelos finales de este estudio.

### 3.1.2. La densidad espectral de un proceso ARMA

Sin pérdida de generalidad, pensemos en un proceso ARMA( $p, q$ ),  $\{X_t\}$ , causal con media cero.

$$\tilde{\phi}_p(B)X_t = \tilde{\theta}_q(B)\epsilon_t, \quad \forall t \in \mathbb{Z}.$$

En la demostración del Lema 1 encontramos que la autocovarianza de este proceso es:

$$\gamma_{(k)} = \sigma_\epsilon^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k} \quad (3.2)$$

Usaremos la relación (3.2) para encontrar la densidad espectral de un ARMA. Primero veamos que, para el proceso ARMA, su función de generación de auto-

covarianza será:

$$\begin{aligned}
 \gamma(B) &= \sum_{k=-\infty}^{\infty} \sigma_{\epsilon}^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k} B^k \\
 &= \sigma_{\epsilon}^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j B^{j-i} \\
 &= \sigma_{\epsilon}^2 \sum_{i=0}^{\infty} \psi_i B^{-i} \sum_{j=0}^{\infty} \psi_j B^j \\
 &= \sigma_{\epsilon}^2 \psi(B) \psi(B^{-1}) \\
 &= \sigma_{\epsilon}^2 \frac{\tilde{\theta}_q(B) \tilde{\theta}_q(B^{-1})}{\tilde{\phi}_p(B) \tilde{\phi}_p(B^{-1})}
 \end{aligned}$$

Finalmente, tendremos un resultado que nos muestra la densidad espectral del proceso ARMA.

**Lema 6.**

Sea  $\{X_t\}_{t \in \mathbb{Z}}$  un proceso ARMA( $p, q$ ) causal. Entonces, su densidad espectral es:

$$f(\nu) = \frac{\sigma^2}{2\pi} \left( \frac{|\tilde{\theta}(e^{-i\nu})|}{|\tilde{\phi}(e^{-i\nu})|} \right)^2 \tag{3.3}$$

***Demostración.***

Usando (3.1) y (3.3) tenemos:

$$\begin{aligned}
 f(\nu) &= \frac{1}{2\pi} \gamma(e^{-i\nu}) \\
 &= \frac{\sigma_{\epsilon}^2}{2\pi} \frac{\tilde{\theta}_q(e^{-i\nu}) \tilde{\theta}_q(e^{i\nu})}{\tilde{\phi}_p(e^{-i\nu}) \tilde{\phi}_p(e^{i\nu})} \\
 &= \frac{\sigma^2}{2\pi} \left( \frac{|\tilde{\theta}(e^{-i\nu})|}{|\tilde{\phi}(e^{-i\nu})|} \right)^2
 \end{aligned}$$

□

### 3.1.3. La densidad espectral de un proceso lineal

Pensemos en un proceso lineal de media cero,  $\{X_t\}$ , esto es:

$$\begin{aligned} X_t &= \sum_{h=0}^{\infty} \alpha_h Y_{t-h} \\ &= \alpha(B)Y_t \end{aligned}$$

donde  $\{Y_t\}$  es cualquier proceso estocástico (absolutamente sumable) con densidad espectral  $f_Y(\nu)$  y autocovarianza  $\gamma_Y(k)$ . Los coeficientes  $\alpha_h$  serán tales que:

$$\sum_{h=0}^{\infty} |\alpha_h|^2 < \infty$$

Con la misma construcción de la demostración del Lema 1 y con (3.2) tenemos que la autocovarianza de  $X_t$  será:

$$\begin{aligned} \gamma(k) &= \mathbb{E}[X_t X_{t+k}] \\ &= \mathbb{E} \left[ \sum_{h=0}^{\infty} \alpha_h Y_{t-h} \sum_{j=0}^{\infty} \alpha_j Y_{t+k-j} \right] \\ &= \sum_{h=0}^{\infty} \alpha_h \sum_{j=0}^{\infty} \alpha_{t+k-j} \mathbb{E}[Y_{t-h} Y_{t+k-j}] \\ &= \sum_{h=0}^{\infty} \alpha_h \sum_{j=0}^{\infty} \alpha_{t+k-j} \gamma_Y(k-j+h) \end{aligned}$$

Con lo que la densidad espectral de  $X_t$  es:

$$\begin{aligned}
 f(\nu) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_{(k)} e^{-ik\nu} \\
 &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \sum_{h=0}^{\infty} \alpha_h \sum_{j=0}^{\infty} \alpha_{t+k-j} \gamma_{Y(k-j+h)} e^{-ik\nu} \\
 &= \sum_{h=0}^{\infty} \alpha_h e^{ih\nu} \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_{Y(k)} e^{-ik\nu} \sum_{j=0}^{\infty} \alpha_j e^{-ij\nu} \\
 &= (\alpha(e^{i\nu}))^2 f_Y(\nu)
 \end{aligned}$$

Por lo tanto

$$f(\nu) = (\alpha(e^{i\nu}))^2 f_Y(\nu) \quad (3.4)$$

donde

$$(\alpha(e^{i\nu}))^2 = \alpha(e^{i\nu})\alpha(e^{-i\nu})$$

A partir de (3.4) podríamos obtener fácilmente (3.3).

Este último resultado es mucho más general y nos permitirá encontrar la densidad espectral de procesos lineales cuyas innovaciones no sean ruidos blancos, como en el caso de los procesos causales.

### 3.1.4. Implementación en R

Para calcular el periodograma en R usaremos una biblioteca que tiene una función que nos arroja el periodograma de los datos. La biblioteca es '*GeneCycle*' y la función *periodogram*.

Con estas bibliotecas crearemos una función que haga la gráfica a partir de una función *periodogram*. El código de esto es:

```

library("GeneCycle")
gperiodogram <- function(a,col) {
  y<-periodogram(a)$spec
  x<-periodogram(a)$freq *2*pi
  plot(x,y, main="Periodograma", ylab="densidad",
  xlab = "frecuencia", type = "l", col = col) }

```

Después usaremos esta función para ver el periodograma de datos y analizar características de esos datos a partir de la gráfica del periodograma.

## 3.2. Memoria

Ahora definiremos formalmente el concepto de memoria de un proceso; comenzaremos por dar una idea intuitiva y después daremos las definiciones usuales.

Primero pensemos en los modelos que se suelen estudiar inicialmente en el análisis de series de tiempo. Por ejemplo los modelos ARMA del Capítulo 1, estos presentan una caída muy rápida de sus funciones de autocorrelación. Esta es una de las características principales que se observa de los modelos ARMA, vemos que a partir de cierta distancia, en las variables a considerar, la ACF es cero.

Pero ¿Qué significa que la autocorrelación "desaparezca"?

Directamente de la definición de la ACF podemos ver que un dramático decaimiento a partir de cierta distancia, mostraría que las variables tienen una correlación nula, entonces las nuevas variables tendrán cada vez menos información de los acontecimientos más antiguos.

Esta es la idea intuitiva del concepto de memoria corta en procesos estocásticos.

Tomando esto, podemos pensar en exactamente lo contrario para hacer un concepto intuitivo de memoria larga; entonces un proceso de memoria larga tiene que tener un decaimiento de su ACF no muy rápido.

Evidentemente esto no nos sirve en matemáticas, debemos ser más precisos en eso de "rápido" y "no muy rápido".

Existen varias definiciones de procesos de memoria larga o corta y en general no son equivalentes. Primero veremos una definición general, una clasificación de los procesos estocásticos en 3 grupos, los de memoria corta, los de memoria larga y los que tienen antipersistencia.

**Definición 18.**

Sea  $\{X_t\}_{t \in \mathbb{N}}$  una serie de tiempo con densidad espectral  $f(\nu)$ . Definimos:

1.  $X_t$  tiene **Antipersistencia** si

$$f(0) = 0 \tag{3.5}$$

2.  $X_t$  tiene **Memoria corta** si

$$0 < f(0) < \infty \tag{3.6}$$

3.  $X_t$  tiene **Memoria larga** si

$$f(0) = \infty \tag{3.7}$$

No hablaremos de lo que significa la antipersistencia e indagaremos poco en los resultados de los procesos con memoria corta. Nos concentraremos, un poco más, en la memoria larga.

Notemos que la definición está, en el fondo, en términos de la autocorrelación del proceso, pues:

$$f(0) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma(j)$$

Por lo que, la memoria depende de la serie de correlaciones.

$$\sum_{j=-\infty}^{\infty} \gamma(j) < \infty \longrightarrow \text{memoria corta}$$

$$\sum_{j=-\infty}^{\infty} \gamma(j) = \infty \longrightarrow \text{memoria larga}$$

Ahora veremos un resultado, que tomando la varianza de la media muestral, nos da información sobre la densidad espectral en 0.

**Teorema 2.** *Teorema de Féjèr.*

Sea  $\{X_t\}_{t \in \mathbb{Z}}$  una serie de tiempo causal con  $f(\nu)$  continua y positiva en 0. Si tenemos una muestra aleatoria de la serie, digamos  $x_1, \dots, x_n$  y tomamos la media muestral,  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ . Entonces:

$$\text{Var}(\bar{x}) = \frac{1}{n} \sum_{j=1-n}^{n-1} \left(1 - \frac{|j|}{n}\right) \gamma(j) \quad (3.8)$$

$$\sim \frac{2\pi f(0)}{n}, \quad \text{cuando } n \rightarrow \infty \quad (3.9)$$

En este caso ' $\sim$ ' indica que el cociente de ambas partes tiende a 1.

Hay señales empíricas que muestran que algunas series de tiempo de agricultura tienen una convergencia parecida a la del Teorema 2 pero con  $n^{-\alpha}$ ,  $0 < \alpha < 1$  en lugar de  $n^{-1}$ .

Este resultado nos da una manera de aproximar la densidad espectral en cero a partir de un muestra aleatoria.

### 3.2.1. Ejemplos de procesos con memoria corta

Más adelante dedicaremos un capítulo a examinar procesos con memoria larga. Así que ahora veamos unos pocos ejemplos de procesos con memoria corta. Serán procesos que ya conocemos.

#### Ruido blanco

Empecemos con un proceso que podemos ver fácilmente que tiene memoria corta. Tomemos un  $\epsilon_t \sim WN(\mu, \sigma^2)$ , por definición sabemos que su función de autocorrelación es:

$$\rho(h) = \begin{cases} 1, & \text{si } h = 0, \\ 0, & \text{si } h \neq 0. \end{cases}$$

Entonces tenemos:

$$\begin{aligned} f(0) &= \frac{\sigma^2}{2\pi} \sum_{j=-\infty}^{\infty} \rho(j) \\ &= \frac{\sigma^2}{2\pi} \\ &< \infty \end{aligned}$$

Por lo tanto un ruido blanco tiene memoria corta.

Además sabemos que un  $SWN(\mu, \sigma^2)$  es un  $WN(\mu, \sigma^2)$  y con el resultado anterior tenemos que un ruido blanco estricto también será un proceso de memoria corta.

Ahora veremos que los primeros modelos que estudiamos también tienen memoria corta.

**AR(1)**

La forma de la autocorrelación de un proceso AR(1) causal ( $|\phi_1| < 1$ ) la vimos en 1.2.2, a partir de ella podemos encontrar el valor de la densidad espectral el cero con relativa facilidad.

$$\begin{aligned}
 f(0) &= \frac{\sigma^2}{2\pi} \sum_{j=-\infty}^{\infty} \rho(j) \\
 &= \frac{\sigma^2}{2\pi} \sum_{j=-\infty}^{\infty} \phi_1^{|j|} \\
 &= \frac{\sigma^2}{2\pi} \left( 2 \sum_{j=1}^{\infty} \phi_1^j + 1 \right) \\
 &= \frac{\sigma^2}{2\pi} \left( 2 \left( \frac{\phi_1}{1 - \phi_1} \right) + 1 \right) \\
 &< \infty
 \end{aligned}$$

Entonces concluimos que los procesos AR(1) son procesos de memoria corta. Ahora veremos un resultado más general.

**ARMA( $p, q$ )**

Utilizando el lema 6 podemos encontrar el valor de la densidad espectral en cero de un modelo ARMA causal.

$$\begin{aligned}
 f(0) &= \frac{\sigma^2}{2\pi} \left( \frac{|\tilde{\theta}(1)|}{|\tilde{\phi}(1)|} \right)^2 \\
 &= \frac{\sigma^2}{2\pi} \left( \frac{|1 + \theta_1 + \dots + \theta_q|}{|1 - \phi_1 - \dots - \phi_p|} \right)^2 \\
 &= \frac{\sigma^2}{2\pi} \left( \sum_{i=0}^{\infty} |\psi_i| \right)^2 \\
 &< \infty
 \end{aligned}$$

Por lo tanto un proceso ARMA( $p, q$ ) causal tiene memoria corta. Esto podíamos notarlo empíricamente desde que vimos las ACF de estos procesos.

### 3.3. Memoria larga

Ahora veremos una definición más general de memoria larga. Veremos no solo cuándo un proceso tiene memoria larga sino también una parametrización de su memoria. Definiremos el parámetro de memoria de un proceso de memoria larga.

#### 3.3.1. Decremento exponencial e hiperbólico

Un gran número de tipos de series de tiempo, incluida la familia de procesos ARMA, tienen un **decremento exponencial** de su autocorrelación, es decir:

$$|\rho_k| \leq Ac^k, \quad 0 < A, \quad 0 < c < 1, \quad \forall k > 0$$

Otros procesos presentan un decremento de su autocorrelación mucho más lento, un **decremento hiperbólico**, esto es:

$$|\rho_k| \sim Ak^{2d-1}, \quad 0 < A, \quad 0 < d < \frac{1}{2}, \quad \text{cuando } k \rightarrow \infty$$

El decremento hiperbólico será lo que nos permita definir una serie de tiempo de memoria larga con un parámetro de memoria que será  $d$ .

#### Definición 19.

Diremos que  $\{X_t\}_{t \in \mathbb{Z}}$  es un **proceso de memoria larga de orden  $d$** ,  $d \in (0, \frac{1}{2})$ , si

$$\gamma_{(k,s)} \sim ch^{2d-1}, \quad \text{cuando } h \rightarrow \infty \quad (3.10)$$

con  $h = k - s$ , y  $c \in \mathbb{R}$ .

A  $d$  lo llamaremos parámetro de memoria del proceso. Esta nueva definición de memoria larga es más general, pues (3.10) implica (3.7).

**Teorema 3.**

Sea  $\{X_t\}_{t \in \mathbb{Z}}$  es un proceso débilmente estacionario de memoria larga de orden  $d$  con densidad espectral  $f(\nu)$ . Entonces:

$$f(0) = \infty$$

**Demostración.**

Se tiene que:

$$\sum_{h=-\infty}^{\infty} \gamma_h \sim \frac{cn^{2d}}{2d}, \text{ para valores grandes de } n$$

$$\rightarrow \infty, \text{ cuando } n \rightarrow \infty$$

de lo que se sigue que

$$\sum_{h=-\infty}^{\infty} \gamma_h = \infty$$

Y por lo tanto

$$f(0) = \infty$$

□

Se puede demostrar que si en (3.10)  $d = 0$  entonces, el proceso tiene memoria corta.

Con esta definición también tenemos un resultado que involucra a la varianza de la media muestral y que nos ayudará para hacer una prueba de memoria.

Tomemos a  $X_t$  una serie de tiempo débilmente estacionaria de memoria larga con parámetro  $d$  y con función de autocovarianza  $\gamma_{(k)}$ . Tomemos una muestra

aleatoria de la serie,  $X_1, X_2, \dots, X_n$ ; con media muestral  $\bar{x}$ . A partir de (3.8) y usando (3.10), para  $n$  y  $j$  grandes, tenemos:

$$\begin{aligned}
 \text{Var}(\bar{x}) &= \frac{1}{n} \left[ 2 \sum_{j=1}^{n-1} \left( 1 - \frac{1}{j} \right) \gamma_{(j)} + \gamma_{(0)} \right] \\
 &\sim 2cn^{2d-1} \sum_{j=1}^{n-1} \left( 1 - \frac{j}{n} \right) \frac{j^{2d-1}}{n} \frac{1}{n} \\
 &\sim 2cn^{2d-1} \int_0^1 (1-t)t^{2d-1} dt \\
 &\sim \frac{c}{d(2d+1)} n^{2d-1} \\
 &= c_2 n^{2d-1}
 \end{aligned} \tag{3.11}$$

### 3.3.2. Estimadores del parámetro de memoria

Aquí veremos dos estimadores para el parámetro de memoria y una prueba de memoria. En adelante supondremos que tenemos datos,  $x_1, x_2, \dots, x_n$ , de alguna serie de tiempo de nuestro interés.

#### Estimador GPH

En GEWEKE, J. y PORTER-HUDAK S. [8] se construye, a partir de una regresión, un estimador para el parámetro de memoria. La suposición de la que parten es que la densidad espectral puede escribirse como:

$$f(\nu) = f_0(\nu_j) \left( 2 \sin \left( \frac{\nu}{2} \right) \right)^{-2d}$$

Los procesos ARFIMA y ARFIMA–GARCH que analizaremos en el Capítulo 4 cumplen esta propiedad.

Pensado en esta forma para la densidad espectral y evaluando en alguna frecuencia de Fourier  $\nu_j$ , tendremos:

$$\log(f(\nu_j)) = \log(f_0(0)) - d \log \left( 2 \sin \left( \frac{\nu_j}{2} \right) \right)^2 + \log \left( \frac{f_0(\nu_j)}{f_0(0)} \right)$$

Por otro lado el logaritmo del pediodograma en la misma frecuencia de Fourier lo podemos escribir como:

$$\log(I_n(\nu_j)) = \log \left( \frac{I_n(\nu_j)}{f(\nu_j)} \right) + \log(f(\nu_j))$$

Con estas 2 ecuaciones obtenemos:

$$\log(I_n(\nu_j)) = \log(f_0(0)) - d \log \left( 2 \sin \left( \frac{\nu_j}{2} \right) \right)^2 + \log \left( \frac{I_n(\nu_j) \left( 2 \sin \left( \frac{\nu_j}{2} \right) \right)^{2d}}{f_0(0)} \right)$$

Ahora construimos la ecuación de regresión:

$$y_j = \alpha + \beta x_j + \epsilon_j$$

donde:

$$y_i = \log(I_n(\nu_j))$$

$$\alpha = \log(f_0(0))$$

$$\beta = -d$$

$$x_j = \log \left( 2 \sin \left( \frac{\nu_j}{2} \right) \right)^2$$

$$\epsilon_j = \log \left( \frac{I_n(\nu_j) \left( 2 \sin \left( \frac{\nu_j}{2} \right) \right)^{2d}}{f_0(0)} \right)$$

Además podemos hacer la aproximación

$$f(\nu_j) \sim f_0(0) \left( 2 \sin \left( \frac{\nu_j}{2} \right) \right)^{-2d}$$

para frecuencias  $\nu_j$  cercanas a cero (es decir, para  $j = 1, \dots, m$  con  $m \ll n$ ). Por lo que también tendremos la aproximación:

$$\epsilon_j \sim \log \left( \frac{I_n(\nu_j)}{f(\nu_j)} \right)$$

Finalmente el estimador de  $d$  será el estimador de  $-\beta$  en la ecuación de regresión:

$${}_m\hat{d}_{GPH} = -\frac{S_{xy}}{S_{xx}}$$

donde:

$$S_{xy} = \sum_{j=1}^m (x_j - \bar{x})(y_j - \bar{y})$$

$$S_{xx} = \sum_{j=1}^m (x_j - \bar{x})^2$$

En ROBINSON, P. [16] se obtiene la distribución asintótica del estimador GPH.

**Teorema 4.**

*Para un proceso de memoria larga y bajo condiciones de regularidad, se tiene:*

$$\sqrt{m}({}_m\hat{d}_{GPH} - d) \rightsquigarrow N \left( 0, \frac{\pi^2}{6S_{xx}} \right) \quad (3.12)$$

para  $m = n^\alpha$  con  $0 < \alpha < 1$ .

Este resultado es válido cuando  $0 < d < \frac{1}{2}$ , cuando el proceso es de memoria larga, pero no tenemos ningún resultado para  $d = 0$ , procesos de memoria corta.

Deo y Hurvich (2003) hicieron un experimento de simulación Monte-Carlo para estudiar la distribución asintótica de  ${}_m\hat{d}_{GPH}$  cuando  $d = 0$ . Ellos concluyen, empíricamente, que (3.12) también es válido para el caso de memoria corta.

Esto nos permitirá construir intervalos de confianza para el parámetro de memoria, hacer una prueba de memoria a un proceso.

Lo más importante de este parámetro es siempre tener presente cómo fue construido, todo parte de la suposición de que la densidad espectral puede escribirse como:

$$f(\nu) = f_0(\nu_j) \left(2 \sin\left(\frac{\nu}{2}\right)\right)^{-2d}$$

El siguiente estimador usa (3.11) sin usar una forma explícita para la densidad del proceso del que se quiere estimar el parámetro de memoria.

### Variance Plot

El segundo estimador que veremos es resultado de otra regresión lineal. Sea  $X_t$  una serie de tiempo de memoria larga con parámetro  $d$ .

Tomemos una muestra aleatoria de la serie,  $X_1, X_2, \dots, X_n$ ; con media muestral  $\bar{x}$ . Sabemos que:

$$\text{Var}(\bar{x}) \sim cn^{2d-1}$$

Ahora, dividimos la muestra en  $k$  bloques de  $m$  variables, con  $km = n$  y  $\bar{x}_j$  la media de bloque  $j$ :  $\{X_{1+m(j-1)}, \dots, X_{m+m(j-1)}\}$  para  $j = 1, \dots, k$ . Entonces obtenemos:

$$\log(\text{Var}(\bar{x}_j)) \sim c + (2d - 1) \log(j)$$

Por lo que un estimador heurístico para el parámetro de memoria será:

$${}_k\hat{d}_V = \frac{S_{ab}}{2S_{aa}} - \frac{1}{2}$$

donde:

$$S_{ab} = \sum_{j=1}^k (\log(j) - a)(\log(\text{Var}(\bar{x}_j)) - b)$$

$$S_{aa} = \sum_{j=1}^k (\log(j) - a)^2$$

$$a = \frac{1}{k} \sum_{j=1}^k \log(j)$$

$$b = \frac{1}{k} \sum_{j=1}^k \log(\text{Var}(\bar{x}_j))$$

El Variance Plot es la gráfica de los puntos  $(\log(\text{Var}(\bar{x}_j)), \log(j))$ , sobre esos puntos se hace la regresión lineal que arroja el estimador.

### La prueba con la estadística $R/S$

La prueba de memoria que veremos será a partir de un reescalamiento de la estadística  $R/S$  (Hurst, 1951; Mandelbrot y Wallis, 1968), esta es:

$$R/S = \frac{1}{S_n} \left[ \max_i \sum_{j=1}^i (x_j - \bar{x}) - \min_i \sum_{j=1}^i (x_j - \bar{x}) \right]$$

donde:

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$$

$$S_n = \left( \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2 \right)^{\frac{1}{2}}$$

La estadística que usaremos en la prueba es:

$${}_n L_{R/S} = \frac{\log(R/S)}{\log(n)} - \frac{1}{2}$$

La prueba, en este caso, es muy sencilla:

Cuando el proceso es de memoria corta  $\Rightarrow {}_nL_{R/S} \approx 0$

Cuando el proceso es de memoria larga  $\Rightarrow {}_nL_{R/S} > 0$

### Implementación en R

Algunos de los estimadores que vimos en esta sección se pueden calcular directamente en R, como el estimador GPH que se puede calcular con la función *fdGPH* de la biblioteca '*fracdiff*'.

También para el estimador  $\widehat{d}_V$  tenemos una función, de la misma biblioteca, que nos permitirá calcularlo e incluso nos muestra el Variance Plot. Esta función es *hurstSlider*, pero esta función nos da la estimación del parámetro  $H = d + 0,5$ . Esta función toma  $k = 50$  por default, pero se puede modificar el parámetro.

Para la estadística  ${}_nL_{R/S}$  tendremos que hacer una función que la haga. El código para crear esta función es el siguiente:

```
LRS<-function(x){
  a<-mean(x)
  b<-sqrt(var(x))
  s<-c(0)
  for(i in 1:length(x)) { s[i]<-sum((x[1:i])-a) }
  R_S<-(max(s)- min(s))/b
  L<-((log(R_S))/(log(length(x))))-(0.5)
  return(L)}
```

Con estas tres funciones podremos hacer pruebas y estimaciones a datos. Con la estadística  ${}_nL_{R/S}$  podemos hacer un prueba de memoria y si obtenemos que los

datos tienen memoria larga entonces, calcular los estimadores que definimos.

Una vez que estimemos el parámetro de memoria podremos buscar un modelo que se ajuste a los datos. Entonces lo único que nos falta para poder trabajar con datos son modelos de memoria larga.

# Capítulo 4

## Modelos de memoria larga

En este capítulo estudiaremos modelos específicos de series de tiempo con memoria larga. Primero veremos los ruidos fraccionarios y estudiaremos cómo se generan y sus propiedades de memoria. Después estudiaremos una generalización de los procesos ARMA que tendrá memoria larga, el proceso ARFIMA.

Finalmente, definiremos un modelo de memoria larga con volatilidad estocástica que será el que usaremos para modelar los datos del Capítulo 5.

Este capítulo está basado en PALMA, W. [14], ROBINSON, P. [16], BAILLIE, R.T. [1], XIE, L. [20] y BAILLIE, R.T., BOLLERSLEV, T., y MIKKELSEN, H.O. [2] y SAMORODNITSKY, G. [13].

### 4.1. Ruido fraccionario

El primer modelo que veremos constituye una familia bastante amplia de procesos con memoria larga. Comenzaremos por definir un concepto preliminar.

### 4.1.1. Procesos autosimilares

El primer estudio, formal, que se hace de los procesos autosimilares lo hace J. Lamperti (1962). La definición que daremos de estos procesos es la siguiente:

**Definición 20.**

Decimos que  $\{X_t\}_{t \in \mathbb{Z}}$  es un **proceso autosimilar con parámetro  $H$**  (discreto),  $0 < H \leq 1$ , si para todo  $a > 0$  se tiene que:

$$X_{at} =_D a^H X_t$$

para  $t \in \mathbb{Z}$ .

Notación: Escribimos que  $\{X_t\}$  es un  $H - SSSI$  cuando  $\{X_t\}$  sea un proceso autosimilar de incrementos estacionarios con parámetro  $H$ .

**Lema 7.**

Sea  $X_t$  un  $H - SSSI$  con varianza finita,  $\mathbb{E}(X_t) = 0$  y  $X_0 = 0$ . Entonces para  $k > s \geq 0$  se tiene:

$$Cov(X_k, X_s) = \frac{\mathbb{E}(X_1^2)}{2} [k^{2H} + s^{2H} - (k - s)^{2H}]$$

**Demostración.**

$$\begin{aligned} Cov(X_k, X_s) &= \mathbb{E}(X_k X_s) \\ &= \mathbb{E} \left[ \frac{1}{2} (X_k^2 + X_s^2 - (X_k - X_s)^2) \right] \\ &= \frac{1}{2} \mathbb{E} [X_k^2 + X_s^2 - (X_{k-s} - X_0)^2] \\ &= \frac{1}{2} \mathbb{E} [X_k^2 + X_s^2 - X_{k-s}^2] \\ &= \frac{1}{2} \mathbb{E} [k^{2H} X_1^2 + s^{2H} X_1^2 - (k - s)^{2H} X_1^2] \\ &= \frac{\mathbb{E}(X_1^2)}{2} [k^{2H} + s^{2H} - (k - s)^{2H}] \end{aligned}$$

□

### 4.1.2. Ruidos fraccionarios

El proceso que realmente nos interesa, se deriva de los procesos autosimilares, es el proceso de incrementos de un  $H - SSSI$ .

**Definición 21.**

Sea  $X_t$  un  $H - SSSI$  con varianza finita,  $\mathbb{E}(X_t)$  constante en el tiempo y  $X_0 = 0$ . Al proceso  $r_t = X_{t+1} - X_t$  lo llamaremos **ruido fraccionario** de parámetro  $H$ .

El ruido fraccionario es el primer modelo que vemos que tiene memoria larga, pero tendrá memoria larga solo cuando  $0.5 < H < 1$ . En tal caso su parámetro será  $d = H - 0.5$ .

**Teorema 5.**

Sea  $r_t$  el ruido fraccionario de parámetro  $H$  derivado de  $\{X_t\}$ , un  $H - SSSI$ . Entonces:

$$Cov(r_k, r_s) = \frac{\mathbb{E}(X_1^2)}{2} [(h+1)^{2H} + (h-1)^{2H} - 2h^{2H}] \quad (4.1)$$

$$\sim \mathbb{E}(X_1^2) 2H(2H-1)h^{2H-2}, \quad \text{cuando } h \rightarrow \infty \quad (4.2)$$

para  $k > s \geq 0$  y  $h = k - s$ .

**Demostración.**

$$\begin{aligned} Cov(r_k, r_s) &= \mathbb{E}((X_{k+1} - X_k)(X_{s+1} - X_s)) \\ &= \mathbb{E}(X_{k+1}X_{s+1}) - \mathbb{E}(X_{k+1}X_s) - \mathbb{E}(X_kX_{s+1}) + \mathbb{E}(X_kX_s) \\ &= \frac{\mathbb{E}(X_1^2)}{2} [(h+1)^{2H} + (h-1)^{2H} - 2h^{2H}] \end{aligned}$$

Ahora,

$$\begin{aligned}
\frac{\text{Cov}(r_k, r_s)}{\mathbb{E}(X_1^2)(4H^2 - 2H)h^{2H-2}} &= \frac{(h+1)^{2H} + (h-1)^{2H} - 2h^{2H}}{4H(2H-1)h^{2H-2}} \\
&= \frac{\sum_{j=0}^{\infty} \binom{2H}{j} h^{2H-j} + \sum_{j=0}^{\infty} \binom{2H}{j} (-h)^{2H-j} - 2h^{2H}}{4H(2H-1)h^{2H-2}} \\
&= \frac{\sum_{j=0}^{\infty} \binom{2H}{j} (h^{2H-j} + (-h)^{2H-j}) - 2h^{2H}}{4H(2H-1)h^{2H-2}} \\
&= \frac{\sum_{j=2}^{\infty} \binom{2H}{j} (h^{2H-j} + (-h)^{2H-j})}{4H(2H-1)h^{2H-2}} \\
&= \frac{2\binom{2H}{2}h^{2H-2} + \sum_{j=3}^{\infty} \binom{2H}{j} (h^{2H-j} + (-h)^{2H-j})}{4H(2H-1)h^{2H-2}} \\
&= 1 + \frac{\sum_{j=3}^{\infty} \binom{2H}{j} (h^{2H-j} + (-h)^{2H-j})}{4H(2H-1)h^{2H-2}}
\end{aligned}$$

Tomando el límite,

$$\lim_{h \rightarrow \infty} 1 + \frac{\sum_{j=3}^{\infty} \binom{2H}{j} (h^{2H-j} + (-h)^{2H-j})}{4H(2H-1)h^{2H-2}} = 1$$

□

Ahora,  $\mathbb{E}(X_1^2)2H(2H-1)$  es un constante respecto a  $h$ . Y si  $H \in (0.5, 1)$  y hacemos  $d = H - 0.5$  entonces:

$$\text{Cov}(r_k, r_s) \sim ch^{2d-1}, \quad \text{cuando } h \rightarrow \infty$$

para  $d \in (0, 0.5)$ . Por lo tanto el ruido fraccionario tiene memoria larga.

Si el proceso  $H - SSSI$ , con media cero, es un proceso gaussiano entonces, se conoce como **movimiento browniano fraccionario** y al ruido fraccionario generado por él se le llama **ruido gaussiano fraccionario**.

El movimiento browniano fraccionario con parámetro  $H \in (0.5, 1)$  fue usado por Mandelbrot y Van Ness (1968) y por Mandelbrot y Wallis (1968) para construir un modelo consistente con las características inusuales del nivel del agua en el río Nilo observadas por Hurst (1951). En este caso el ruido gaussiano fraccionario que generó el modelo, representaría los incrementos en el nivel del agua del río.

R tiene una biblioteca que nos permitirá hacer simulaciones de movimientos brownianos fraccionarios y, a partir de ellos, de los ruidos gaussianos fraccionarios.

La biblioteca es '*dvfBm*', una vez instalada, para simular un movimiento browniano fraccionario se usa la función *circFBM*.

Simularemos 4 movimientos brownianos fraccionarios con parámetros  $H = 0.2, 0.4, 0.6, 0.8$ . El código para simular esto es:

```
library("dvfBm")
x2<-circFBM(1000,0.2)
x4<-circFBM(1000,.4)
x6<-circFBM(1000,.6)
x8<-circFBM(1000,.8)
mBf2<-ts(x2)
mBf4<-ts(x4)
mBf6<-ts(x6)
mBf8<-ts(x8)
par(mfrow=c(2,2))
plot(mBf2,type="l",col=Red",main="mBf, H=0.2")
plot(mBf4,type="l",col="Blue",main="mBf, H=0.4")
plot(mBf6,type="l",col="green",main="mBf, H=0.6")
plot(mBf8,type="l",col=".orange",main="mBf, H=0.8")
```

La gráfica resultante se puede ver en la Figura 4.1.

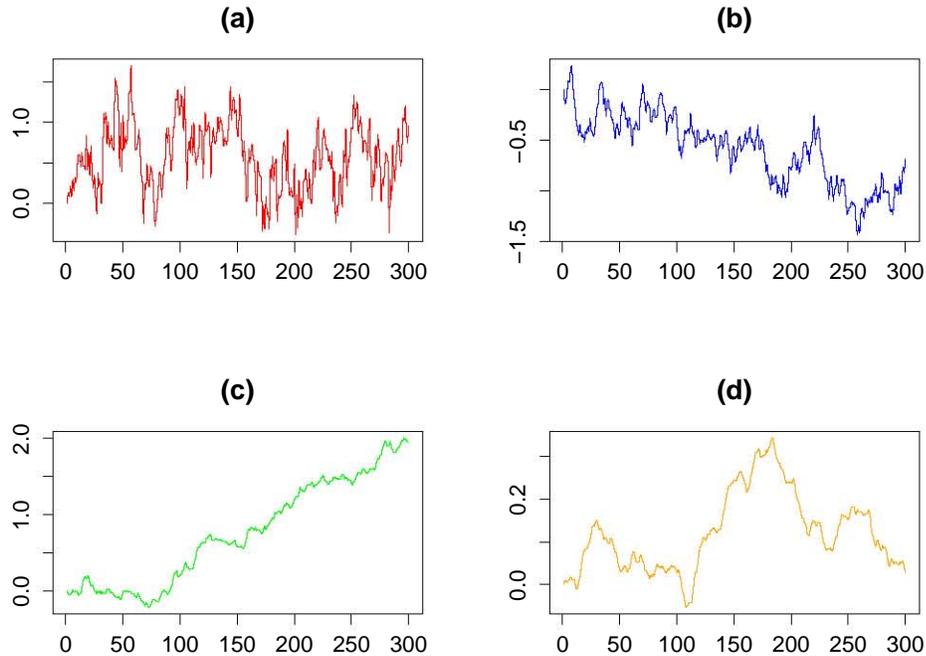


Figura 4.1: Simulación de movimientos brownianos fraccionarios con parámetro  $H =$  (a)0.2, (b)0.4, (c)0.6 y (d)0.8.

Una vez que tenemos esta simulación podemos generar un ruido gaussiano fraccionario de cada una de las simulaciones tomando los procesos de incrementos  $r_t = X_{t+1} - X_t$  por cada movimiento browniano fraccionario.

En R podemos obtener los procesos de incrementos y sus ACF con el siguiente código:

```
plot(diff(mBf2),type="l",col=Red",main=rGf, H=0.2")
```

```

plot(plot(mBf4),type="l",col="Blue",main=rGf, H=0.4")
plot(diff(mBf6),type="l",col="green",main=rGf, H=0.6")
plot(diff(mBf8),type="l",col="orange",main=rGf, H=0.8")

```

## 4.2. ARFIMA

Los procesos ARFIMA (Autoregressive fractionally integrated moving average), por su nombre, parecerían una generalización de los ARIMA pero no lo son. En realidad los procesos ARIMA y los ARFIMA son excluyentes.

Fueron introducidos por Granger y Joyeux (1980) y Hosking (1981).

### Definición 22.

Sea  $(\epsilon_t)_{t \in \mathbb{Z}}$  un  $WN(0, \sigma_\epsilon^2)$ . Decimos que  $X_t$  es un **proceso ARFIMA** $(p, d, q)$  si satisface:

$$\tilde{\phi}_p(B)(1 - B)^d X_t = \tilde{\theta}_q(B)\epsilon_t \quad (4.3)$$

para  $t \in \mathbb{Z}$ ,  $p, q \in \mathbb{N} \cup \{0\}$  y con  $d \in (0, 0.5)$ .

Para este modelo será muy fácil intuir cuál es el parámetro de memoria pues, como puede pensarse, es  $d$ .

Los procesos ARIMA y los ARFIMA son excluyentes ya que, a pesar de que la ecuación que los define es la misma, en el proceso ARIMA  $d$  es un entero y en el ARFIMA es un real entre 0 y 0.5. El hecho de que  $d \in (0, 0.5)$  causa problemas al definir al operador  $(1 - B)^d$  por lo que lo podemos definir con su expansión de

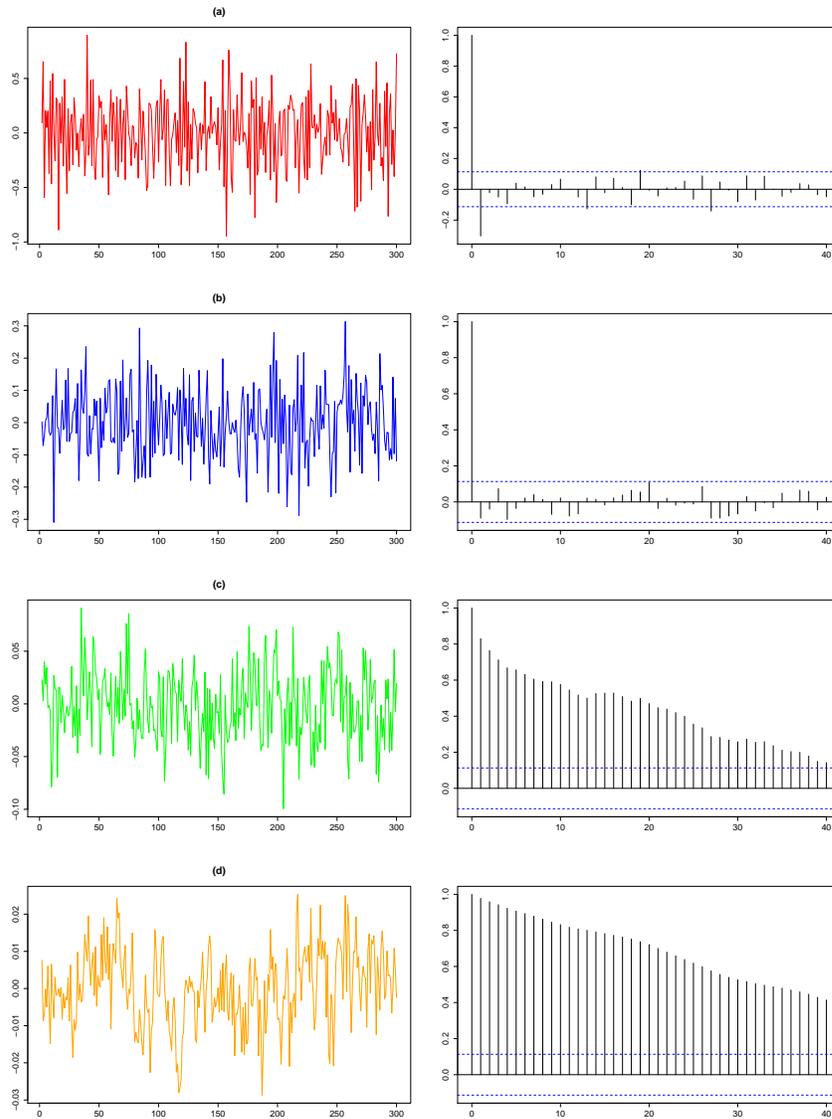


Figura 4.2: Ruidos Gaussianos fraccionarios con parámetro  $H =$  (a)0.2, (b)0.4, (c)0.6 y (d)0.8. Y sus ACF.

Taylor:

$$\begin{aligned} (1 - B)^d &= \sum_{n=0}^{\infty} \binom{d}{n} (-B)^n \\ &= \sum_{n=0}^{\infty} \frac{\Gamma(d+1)}{\Gamma(n+1)\Gamma(d-n+1)} (-B)^n \end{aligned}$$

Ahora, utilizando (3.4) construiremos la densidad espectral de este proceso.

#### 4.2.1. La densidad espectral de un proceso ARFIMA

Consideremos a  $X_t$ , un proceso ARFIMA( $p, d, q$ ), es decir,

$$\tilde{\phi}_p(B)(1 - B)^d X_t = \tilde{\theta}_q(B)\epsilon_t$$

tomaremos a los polinomios  $\tilde{\phi}_p$  y  $\tilde{\theta}_q$  tales que el proceso ARMA  $Z_t$ , con:

$$\tilde{\phi}_p(B)Z_t = \tilde{\theta}_q(B)\epsilon_t,$$

sea causal. Entonces tendremos la relación:

$$X_t = (1 - B)^{-d} Z_t$$

y nuevamente usando la expansión de Taylor podremos escribir:

$$\begin{aligned} (1 - B)^{-d} &= \sum_{n=0}^{\infty} \binom{-d}{n} (-B)^n \\ &= \sum_{n=0}^{\infty} \frac{\Gamma(d+n)}{\Gamma(n+1)\Gamma(d)} B^n \end{aligned}$$

Por lo tanto

$$\begin{aligned} X_t &= \sum_{n=0}^{\infty} \frac{\Gamma(d+n)}{\Gamma(n+1)\Gamma(d)} B^n Z_t \\ &= \sum_{n=0}^{\infty} \pi_n B^n Z_t \\ &= \sum_{n=0}^{\infty} \pi_n Z_{t-n} \end{aligned}$$

Usando la fórmula de Stirling

$$\Gamma(x) \approx \sqrt{2\pi} e^{-x} x^{x-0.5}, \quad \text{cuando } x \rightarrow \infty$$

tenemos:

$$\begin{aligned} \pi_n &= \frac{\Gamma(d+n)}{\Gamma(n+1)\Gamma(d)} \\ &\approx \frac{1}{\Gamma(d)} \frac{\sqrt{2\pi} e^{-n-d} (n+d)^{n+d-0.5}}{\sqrt{2\pi} e^{-n-1} (n+1)^{n+1-0.5}} \\ &\approx \frac{1}{\Gamma(d)n^{(1-d)}}, \quad \text{cuando } n \rightarrow \infty \end{aligned}$$

Por lo que la sucesión  $\{\pi_n\}$  será absolutamente sumable si y solo si  $2(1-d) > 1$ , es decir  $d < 0.5$ . Y como tomamos a  $Z_t$  causal entonces,  $X_t$  será causal si y solo si  $d < 0.5$ .

Con este supuesto y con (3.4), la densidad espectral de  $X_t$  es:

$$f(\nu) = \frac{\sigma^2}{2\pi} (1 - e^{-i\nu})^{-2d} \left( \frac{|\tilde{\theta}(e^{-i\nu})|}{|\tilde{\phi}(e^{-i\nu})|} \right)^2 \quad (4.4)$$

Con esto podemos calcular  $f(0)$  para ver que un proceso ARFIMA tiene memoria larga:

$$\begin{aligned} \lim_{\nu \rightarrow 0} \frac{\sigma^2}{2\pi} (1 - e^{-i\nu})^{-2d} \left( \frac{|\tilde{\theta}(e^{-i\nu})|}{|\tilde{\phi}(e^{-i\nu})|} \right)^2 &= \frac{\sigma^2}{2\pi} \left( \frac{|\tilde{\theta}(1)|}{|\tilde{\phi}(1)|} \right) \lim_{\nu \rightarrow 0} (1 - e^{-i\nu})^{-2d} \\ &= \frac{\sigma^2}{2\pi} \tilde{\psi}(1)^2 \lim_{\nu \rightarrow 0} (1 - e^{-i\nu})^{-2d} \\ &= \infty \end{aligned}$$

ya que  $-1 < -2d < 0$ .

Concluimos entonces que un proceso ARFIMA tiene memoria larga. Se puede demostrar, incluso, que un ARFIMA tiene parámetro de memoria  $d$  (el mismo que su parámetro de integración). Pero aquí solo haremos la demostración para el caso del ARFIMA(0,  $d$ , 0).

### 4.2.2. Ruidos fraccionales integrados

Ahora estudiaremos un caso particular: el ARFIMA(0,  $d$ , 0) a este proceso se le suele llamar ruido fraccionario integrado. Este proceso será tal que

$$(1 - B)^d X_t = \epsilon_t$$

Sabemos, por (4.4), que la densidad espectral de este proceso es:

$$\begin{aligned} f(\nu) &= \frac{\sigma^2}{2\pi} (1 - e^{-i\nu})^{-2d} \\ &= \frac{\sigma^2}{2\pi} \left[ 2 \sin\left(\frac{\nu}{2}\right) \right]^{-2d} \end{aligned}$$

Con lo que podemos calcular su autocorrelación:

$$\begin{aligned} \gamma(k) &= \int_{-\pi}^{\pi} e^{ik\nu} f(\nu) d\nu \\ &= \frac{\sigma^2}{2\pi} \int_0^{\pi} 2 \cos(k\nu) \left[ 2 \sin\left(\frac{\nu}{2}\right) \right]^{-2d} d\nu \\ &= (-1)^k \sigma^2 \frac{\Gamma(1 - 2d)}{\Gamma(k - d + 1)\Gamma(1 - d - k)} \\ &= \sigma^2 \frac{\Gamma(1 - 2d)}{\Gamma(d)\Gamma(1 - d)} \frac{\Gamma(k + d)}{\Gamma(k - d + 1)} \end{aligned}$$

Hacemos

$$c_1 = \sigma^2 \frac{\Gamma(1 - 2d)}{\Gamma(d)\Gamma(1 - d)}$$

y usando nuevamente la fórmula de Stirling tenemos, para  $k$  grandes:

$$\begin{aligned} \gamma(k) &\sim c_1 \frac{\sqrt{2\pi} e^{-k-d} (k+d)^{k+d-0.5}}{\sqrt{2\pi} e^{-k+d-1} (k-d+1)^{k-d+0.5}} \\ &\sim c_1 e^{1-2d} k^{2d-1} \end{aligned}$$

finalmente para  $c = c_1 e^{1-2d}$  tenemos:

$$\gamma(k) \sim c k^{2d-1}$$

Con esto podemos concluir que el proceso ARFIMA(0,  $d$ , 0) tiene memoria larga de parámetro  $d$ .

En R podemos obtener procesos ARFIMA y sus ACF con el siguiente código:

```
library("fArma")
farima1 <- farimaSim(n, model=list(ar=c(.5, -.5), d=0.1, ma=.1))
farima2 <- farimaSim(n, model=list(ar=c(.5, -.5), d=0.2, ma=.1))
farima3 <- farimaSim(n, model=list(ar=c(.5, -.5), d=0.3, ma=.1))
farima4 <- farimaSim(n, model=list(ar=c(.5, -.5), d=0.4, ma=.1))
plot(farima1, type="l", main=".^RFIMA(2,.1,1)")
plot(farima2, type="l", main=".^RFIMA(2,.2,1)")
plot(farima3, type="l", main=".^RFIMA(2,.3,1)")
plot(farima4, type="l", main=".^RFIMA(2,.4,1)")
acf(farima1, 40)
acf(farima2, 40)
acf(farima3, 40)
acf(farima2, 40)
```

### 4.3. Modelos de memoria larga con volatilidad estocástica

En esta sección estudiaremos dos modelos de memoria larga que tienen volatilidad estocástica, como los modelos ARCH y GARCH que estudiamos en el Capítulo 2. El primer modelo solo lo mencionaremos y veremos que fácilmente se encuentra su parámetro de memoria.

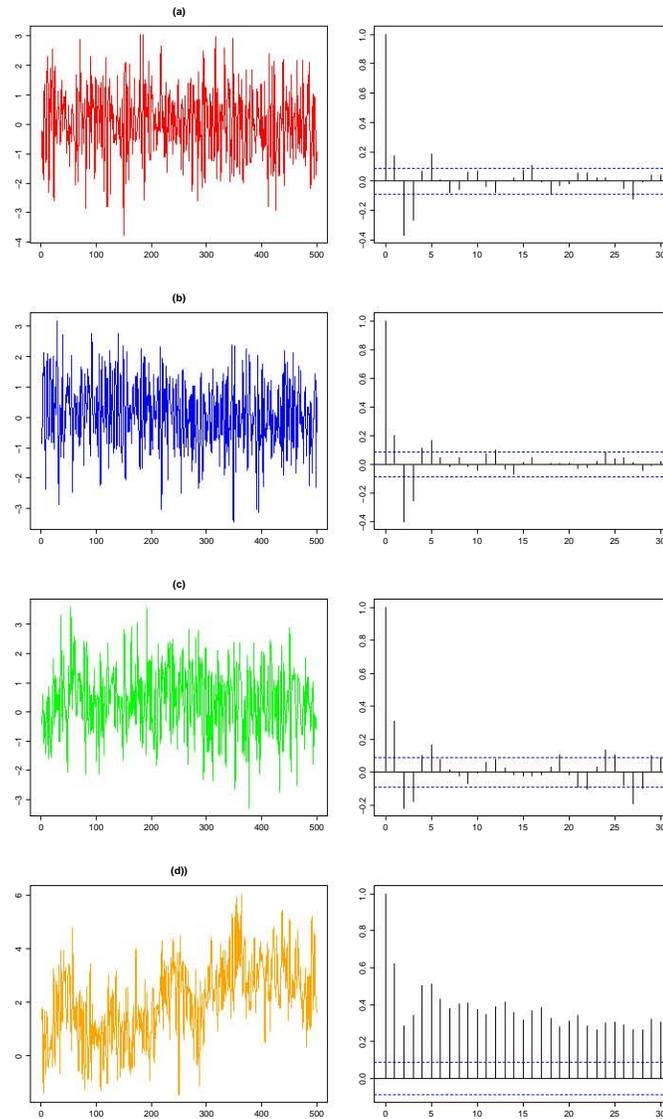


Figura 4.3: Simulación de procesos ARFIMA(2,  $d$ , 1) con parámetros  $\phi_1 = 0.5$ ,  $\phi_2 = -0.5$ ,  $\theta_1 = 0.1$  y (a)  $d = 0.1$ , (b)  $d = 0.2$ , (c)  $d = 0.3$  y (d)  $d = 0.4$ . Y sus ACF.

El segundo modelo es un ARFIMA con innovaciones GARCH, ya hemos trabajado bastante los aspectos teóricos de estos modelos, así que no hablaremos mucho de él, pero en el siguiente capítulo será el modelo que ajustaremos a los datos que analizaremos.

### 4.3.1. LMSV

Este modelo está especialmente construido para series de tiempo que representan los rendimientos de algún activo financiero.

El modelo LMSV, Long Memory Stochastic Volatility, para una serie débilmente estacionaria de rendimientos es:

$$r_t = \sigma_t \epsilon_t \quad (4.5)$$

$$\sigma_t = \sigma \exp\left(\frac{v_t}{2}\right) \quad (4.6)$$

donde  $\sigma > 0$ ,  $\{\epsilon_t\}$  es un  $SWN(0, 1)$  y  $\{v_t\}$  es un proceso de memoria larga independiente de  $\epsilon_t$ .

El proceso  $\{v_t\}$  puede ser cualquiera de los que estudiamos anteriormente, y definirá la forma en la que se mueve (estocásticamente) la volatilidad del modelo.

Como en el Capítulo 2 nos interesa más el comportamiento de los log-retornos de la serie, pero ahora nos interesaremos en los logaritmos de los cuadrados de los retornos. A partir de la definición de modelo LMSV tenemos:

$$\log(r_t^2) = \log(\sigma_t^2) + \log(\epsilon_t^2)$$

$$\log(\sigma_t^2) = \log(\sigma^2) + v_t$$

Y si hacemos:  $y_t = \log(r_t^2)$ ,  $\mu = \log(\sigma^2) + \mathbb{E}[\log(\epsilon_t^2)]$ ,  $\eta = \log(\epsilon_t^2) - \mathbb{E}[\log(\epsilon_t^2)]$ , tendremos:

$$y_t = \mu + v_t + \eta_t$$

A partir de esta relación podemos ver fácilmente que:

$$\gamma_{y(s,k)} = \gamma_{v(s,k)} + \sigma_{\eta}^2 \mathbb{I}_{\{k=s\}}$$

Con lo que podemos concluir directamente que el proceso  $\{y_t\}$  tiene memoria larga e incluso, si  $\{v_t\}$  tiene parámetro de memoria  $d$ , entonces  $\{y_t\}$  también tendrá parámetro de memoria  $d$ .

Este modelo se está empezado a usar para valuación de opciones, en los casos donde se observan señales de memoria larga, usando ruidos fraccionarios para el proceso  $\{v_t\}$ .

### 4.3.2. ARFIMA–GARCH

El modelo que estudiaremos con más interés es el ARFIMA–GARCH. Ahora veremos un modelo con las características principales de un ARFIMA (memoria larga) y con volatilidad estocástica definida por un proceso GARCH. La definición de este modelo se puede intuir a partir de la definición 13.

#### Definición 23.

Decimos que  $X_t$  es un **proceso ARFIMA**( $p_1, d, q_1$ )–**GARCH**( $p_2, q_2$ ) si es débilmente estacionario y cumple:

$$\begin{aligned} \tilde{\phi}_{p_1}(B)(1-B)^d X_t &= \tilde{\theta}_{q_1}(B)\epsilon_t \\ \epsilon_t &= \sigma_t Z_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p_2} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q_2} \beta_j \sigma_{t-j}^2 \end{aligned}$$

para  $t \in \mathbb{Z}$ ,  $\tilde{\phi}_{p_1}$  y  $\tilde{\theta}_{q_1}$  polinomios reales,  $0 < d < 0.5$ ,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $i = 1, 2, \dots, p_2$  y  $\beta_j \geq 0$ ,  $j = 1, 2, \dots, q_2$  y con la condición  $\sum_{i=1}^{p_2} \alpha_i + \sum_{j=1}^{q_2} \beta_j < 1$  y  $(Z_t)_{t \in \mathbb{Z}}$  un  $SWN(0, 1)$ .

### Medidas de riesgo

Supongamos que tenemos datos,  $x_1, x_2, \dots, x_n$ , a los que se les ajusta un modelo ARFIMA–GARCH, una vez que logremos estimar los parámetros, nos interesará calcular las medidas de riesgo de los datos para el siguiente tiempo, es decir para la variable  $X_{n+1}$ .

Con la hipótesis de que nuestra serie se comporta como un ARFIMA–GARCH tendremos:

$$\tilde{\phi}_{p_1}(B)(1-B)^d X_{n+1} = \tilde{\theta}_{q_1}(B)\epsilon_{n+1}$$

Por lo que la variable que nos interesa la podemos escribir como:

$$X_{n+1} = \sum_{i=0}^{\infty} \psi_i \epsilon_{n+1-i} - \sum_{j=1}^{\infty} \pi_j x_{n+1-j}$$

La realidad es que con la información finita que tenemos, con la estimación de los residuales y de los parámetros, la variable se verá de la forma:

$$X_{n+1} = \epsilon_{n+1} + \sum_{i=1}^{n-1} \hat{\psi}_i \hat{\epsilon}_{n+1-i} - \sum_{j=1}^n \hat{\pi}_j x_{n+1-j}$$

Finalmente, el único componente estocástico es  $\epsilon_{n+1}$ , la siguiente innovación. Entonces el VaR de la variable  $X_{n+1}$  será:

$$\text{VaR}_\alpha(X_{n+1}) = \text{VaR}_\alpha(\epsilon_{n+1}) + \sum_{i=1}^{n-1} \hat{\psi}_i \hat{\epsilon}_{n+1-i} - \sum_{j=1}^n \hat{\pi}_j x_{n+1-j} \quad (4.7)$$

De igual manera podemos calcular el expected-shortfall de la variable  $X_{n+1}$ , esto es:

$$\text{ES}_\alpha(X_{n+1}) = \text{ES}_\alpha(\epsilon_{n+1}) + \sum_{i=1}^{n-1} \hat{\psi}_i \hat{\epsilon}_{n+1-i} - \sum_{j=1}^n \hat{\pi}_j x_{n+1-j} \quad (4.8)$$

donde:

$$\sum_{j=1}^n \hat{\pi}_j x_{n+1-j} = \sum_{j=1}^n (-x_{n+1-j}) \frac{\Gamma(\hat{d} + 1)}{\Gamma(j + 1)\Gamma(\hat{d} - j + 1)}$$

El ARFIMA–GARCH hereda varias propiedades del ARFIMA y del GARCH, por ejemplo sabemos que tiene memoria larga y que su parámetro de memoria es  $d$ . También sabemos que, gracias a sus innovaciones, tendrá volatilidad estocástica. Y finalmente el estudio de los procesos ARFIMA y GARCH que hemos hecho nos permitió dar una forma de calcular el Var y el expected-shortfall de este modelo.

# Capítulo 5

## Aplicación. La inflación mexicana

Para finalizar este trabajo ajustaremos el modelo ARFIMA–GARCH a los datos de la inflación mexicana. Primero daremos una breve introducción histórica de la inflación en México y hablaremos de la metodología de su cálculo. Veremos que la serie tiene características que nos hacen pensar que tienen memoria larga, las características que estudiamos en el Capítulo 3.

Después veremos los resultados del ajuste del modelo de memoria larga y volatilidad estocástica, el parámetro de memoria estimado y las estimaciones de los demás parámetros del modelo. Usaremos estas estimaciones para calcular las medidas de riesgo de la inflación.

Por último, calibraremos el modelo y veremos si el modelo que ajustamos a la serie es bueno o no. Esto lo haremos con un Back-testing.

El parte del contenido teórico de este capítulo está basado en ORTEGA PÉREZ DE LEÓN, ARMANDO. ET. AL. [13]. La aplicación realizada en este capítulo está motivada por las aplicaciones realizadas en ZAMBRENO, J.A. [21] y OOMS, M., y DOORNIK, J. [12].

## 5.1. Inflación

La inflación es el aumento sostenido y generalizado de los precios de los bienes y servicios de un país (o zona económica) en una ventana de tiempo. La principal consecuencia de este fenómeno es la pérdida del poder de compra del dinero; las personas pueden comprar menos con la misma cantidad de dinero.

La inflación afecta directamente a todos los consumidores de un mercado pues su capacidad de consumir (poder adquisitivo) se ve disminuido. También tiene repercusiones en los acreedores de instrumentos financieros pues cuando se liquiden estos instrumentos la ganancia obtenida tendrá un poder de compra menor entonces, sus tasas de rendimiento serán menores de lo esperado.

Si las tasas de rendimiento bajan, entonces las inversiones también se ven reducidas y con esto se generan aumentos en las tasas de desempleo.

Es por ello que se puede considerar a la inflación como el problema que más incide negativamente en los esfuerzos de un país por lograr desarrollo y crecimiento económico.

Se considera que lo ideal es que los precios aumenten un poco cada año, pues una disminución de los precios, llamada deflación, tampoco es buena para la economía: si los bienes y servicios bajan de precio lo más normal es que se retracen las compras y las inversiones ("mañana será más barato") entonces, las empresas no obtienen ganancias, desaparecen y generan desempleo.

Entonces fácilmente podemos entender porqué es importante medir la inflación y administrar correctamente sus riesgos. Para hacer esa correcta administración de los riesgos existentes debemos conocer las causas de la inflación, sus consecuencias y su comportamiento matemático (como una serie de tiempo).

Para poder administrar los riesgos de la inflación, hay que medirla. La forma más común de hacerlo es crear un conjunto de bienes y servicios cuyo uso se considere necesario para la población del país en cuestión, una canasta básica. Una vez establecida esta canasta se obtienen los precios de los productos en el mercado y a partir de ellos se construye un índice llamado (generalmente) Índice de Precios al Consumidor. Finalmente, con la metodología que el banco central de dicho país determine, se calcula la inflación del índice obtenido y el resultado se conoce como tasa de inflación.

En México se denomina al índice: Índice Nacional de Precios al Consumidor (INPC). En la siguiente sección hablaremos de cómo se calcula y un poco de su historia.

Después de ver como se mide la inflación nos concentraremos en estudiar el comportamiento de las series de tiempo de las tasas de inflación.

### **5.1.1. La inflación en México**

En México se utiliza el Índice Nacional de Precios al Consumidor (INPC) para medir la inflación. El INPC comenzó a publicarse en 1969, siendo facultad del Banco de México su cálculo y publicación. A partir de la primera quincena de julio de 2011 (por el presidencial publicado en el Diario Oficial de la Federación el 16 de abril de 2008) se le confiere al Instituto Nacional de Estadística y Geografía (INEGI) la facultad de calcular y publicar el INPC.

La metodología que se usa para el cálculo del índice es la siguiente.

El INPC se calcula considerando los precios de una canasta de bienes y ser-

vicios representativa del consumo de los mexicanos. Para construir la canasta de consumo se usa la información de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) que hace el INEGI. A esta información se incorporan ponderadores de las variaciones de precios.

Por ejemplo, la ENIGH realizada en 2010 recabó información de una canasta de 283 bienes y servicios genéricos. Después, se seleccionó una muestra representativa de productos específicos correspondientes a cada producto genérico y a cuyos precios se da seguimiento en los establecimientos donde compran los consumidores. Esta muestra es cercana a 83,500 bienes y servicios específicos.

Actualmente se consideran 46 ciudades, distribuidas a lo largo y ancho del territorio nacional, para extraer la muestra de precios.

Los precios no cambian simultáneamente, ni avanzan todo el tiempo a la misma velocidad. Esta es una de las razones por las que se ponderan los productos de la canasta de bienes y servicios para el cálculo del INPC. Las ponderaciones actuales son:

Vivienda :	28.18 %
Alimentos, bebidas y tabaco :	23.29 %
Transporte :	14.64 %
Educación y esparcimiento :	9.16 %
Salud y cuidado personal :	7.82 %
Ropa, calzado y accesorios :	5.04 %
Muebles, aparatos y accesorios domésticos :	4.1 %
Otros :	7.76 %

Para hacer las comparaciones del cambio de los precios se considera el año

base, que es el punto de referencia en el tiempo a partir del cual se efectúan las comparaciones. La estructura de ponderadores para el año base se calcula de acuerdo con el comportamiento de los gastos de los ciudadanos.

La fórmula básica del índice de precios en el periodo  $[t_1, t_2]$  se puede definir como:

$$I_{[t_1, t_2]} = \sum_j W_{j, t_1} \left( \frac{P_{j, t_2} - P_{j, t_1}}{P_{j, t_1}} \right)$$

donde  $W_{j, t_1}$  es la ponderación de producto  $j$  en el tiempo  $t_1$  y  $P_{j, t_2}$  y  $P_{j, t_1}$  son los precios del producto  $j$  en  $t_2$  y  $t_1$  respectivamente.

El histórico de la inflación mexicana se presenta en la Figura 5.1.

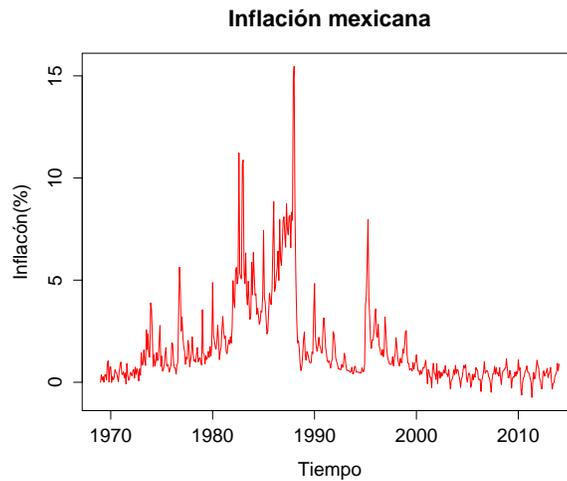


Figura 5.1: La inflación mexicana desde 1969 a la fecha.

Como puede verse en la Figura 5.1 la inflación mexicana tiene momentos de gran volatilidad en los años 80's. Esto se debe a la crisis de 1982. Todo inicia

con una dramática caída de los precios del petróleo en 1981 que se conjugó con un fuerte aumento en las tasas de interés. En 1982, el presidente López Portillo suspende los pagos de la deuda externa, devalúa el peso y nacionaliza el sistema bancario.

La confianza en México se había perdido y pocas instituciones financieras internacionales estaban dispuestas a dar préstamos a México. Al inicio de 1982 el presidente Miguel de la Madrid implementa reformas importantes como continuas devaluaciones, lo cual produjo altas tasas de inflación que llegaron hasta el 15.97 % anual en 1987.

Contrario a lo que se piensa actualmente, los incrementos en el precio de los productos no se comparan con las subidas de los años 80's. Por ejemplo: si el precio del kilo de tortilla en 1982 era de 3 pesos (\$3.00), a finales del sexenio del presidente De la Madrid (1982-1988) llegó a costar \$89.00 pesos (incremento de 2961 %).

Para hacer el estudio de la inflación mexicana tomaremos en cuenta la serie desde 1990, después de los efectos de los problemas que ya mencionamos.

Lo interesante de esta serie de tiempo es el comportamiento de su ACF. Veremos la ACF para la serie completa y para la serie truncada de 3 formas: desde 1980, desde 1990 y desde 2000:

Es bastante notorio, a pesar de truncar la serie, que no podemos pensar este proceso como uno de memoria corta. Además, los clusters que aparecen en los 90's nos hacen pensar en una serie de volatilidad estocástica.

Un análisis exploratorio de los datos nos arroja los resultados de la tabla 5.1 y de las figuras 5.3 y 5.4.

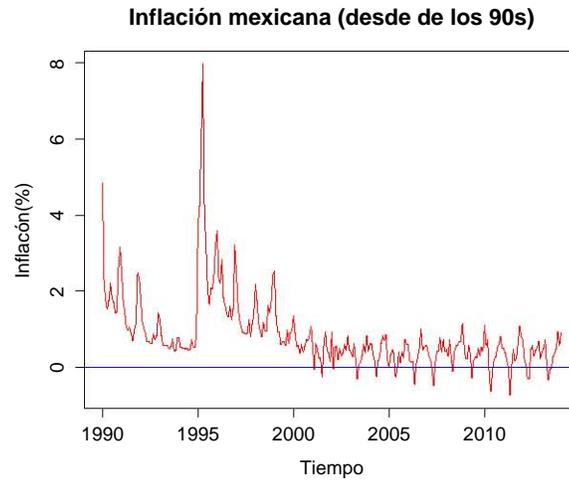


Figura 5.2: La inflación mexicana desde 1990 a la fecha.

Ahora con las evidencias de memoria larga, como muestra la figura 5.3, y volatilidad estocástica que muestra el histórico de la serie, seguiremos con la estimación de un modelo que cumpla estas características.

Cuadro 5.1: Medidas descriptivas

Media	Kurtosis	Varianza
0.8498962	16.49699	0.9315649

También podemos ver cómo se ve la comparación de los datos contra la distribución normal en la Figura 5.5.

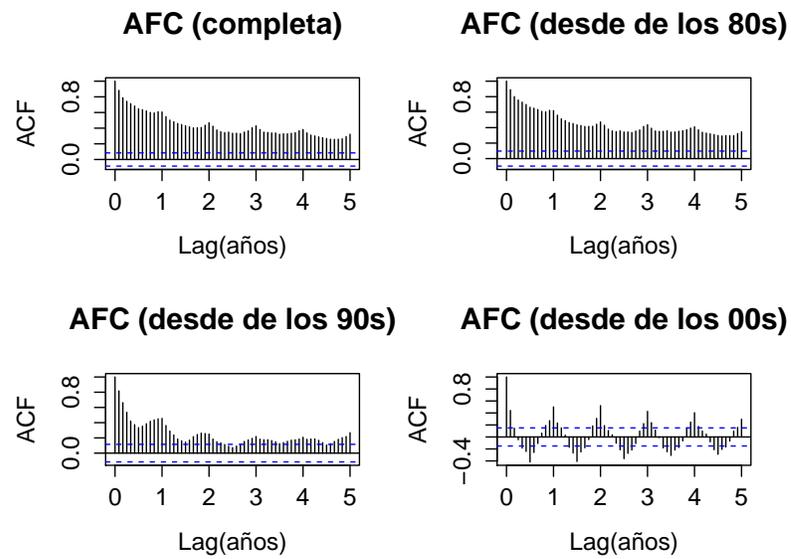


Figura 5.3: ACF de la serie de la inflación mexicana en varias ventanas de tiempo.

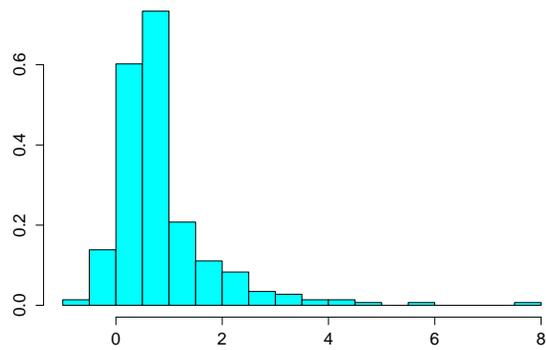


Figura 5.4: Histograma de la inflación mexicana desde 1990.

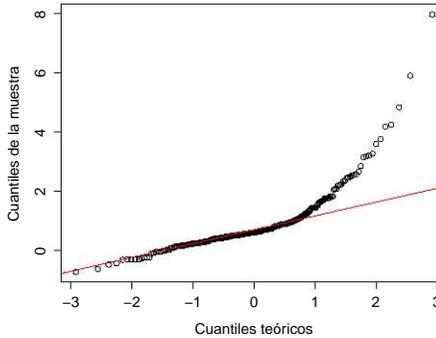


Figura 5.5: Q-Qplot de la inflación mexicana desde 1990 contra una normal estándar.

## 5.2. Ajuste del modelo ARFIMA-GARCH

La inflación mexicana parece cumplir los hechos estilizados, excepto por una diferencia: no se comporta como un ruido blanco, el ACF muestra fuerte correlación serial (Figura 5.3). Este hecho es el que nos hace pensar que será mejor ajustar un modelo de memoria larga en esta ocasión. Además, el comportamiento de la volatilidad de la serie nos hace pensar que no es constante, por lo que ajustaremos un modelo que también tenga volatilidad estocástica. Trataremos a esta serie como un ruido fraccionario integrado (para simplificar lo más posible el modelo) con innovaciones GARCH.

Primero veamos que el periodograma de la serie muestra cierta evidencia de divergencia en 0 (Figura 5.6).

Ahora debemos hacer es un prueba para saber si la serie tiene o no memoria larga y de ser así entonces estimar el parámetro de memoria de la serie. Esto lo haremos con la prueba de la estadística R/S y con los estimadores que vimos en el capítulo anterior.

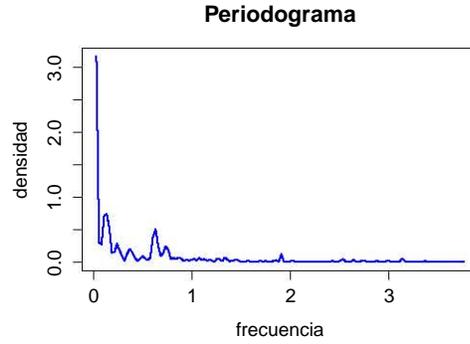


Figura 5.6: Periodograma de la inflación mexicana desde 1990.

$${}_n L_{R/S} = 0.279$$

Por lo que concluimos que la serie sí tiene memoria larga. Ahora hacemos la estimación del parámetro de memoria:

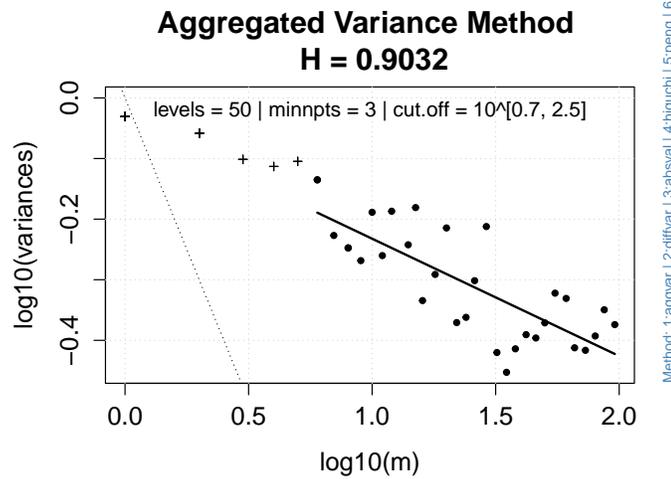
$$\begin{aligned} m \hat{d}_{GPH} &= 0.3845 \\ k \hat{d}_V &= 0.4032 \end{aligned}$$

Usando  $k = 50$  y  $m = n^{0.6} = 30$ , que son parámetros que se suelen usar en la práctica.

Con ayuda del Teorema 4 podemos construir un intervalo del 99 % de confianza para el parámetro de memoria, este es:

$$(0.3168, 0.4522)$$

El variance plot es el siguiente:



El parámetro que usaremos es  $d = 0.3845$ , por nuestra hipótesis de que la serie es un ruido fraccionario integrado y sabemos por lo visto en el capítulo 3 que el estimador GPH se comporta bastante bien en este tipo de series.

Ahora que tenemos el parámetro de memoria, transformaremos la serie aplicando la transformación:  $(1 - B)^d X_t$  que suponemos se comporta como un GARCH; después le ajustaremos un GARCH al resultado. La transformación la hacemos con el siguiente código:

```
difffrac1<-function(a,d,m){
  n<-length(a)
  y <- x[m:n]
  for(i in 1:(m-1)){
    y <- y+((-1)i*(((g(d+1))/(g(i+ 1)*g(d-i +1))))*x[(m-i):(n-i)])}
  return(y)}
xd<-difffrac1(x,0.3845,90)
plot(xd,type="l")
```

Al aplicar la transformación a la serie obtenemos la Figura 5.7.

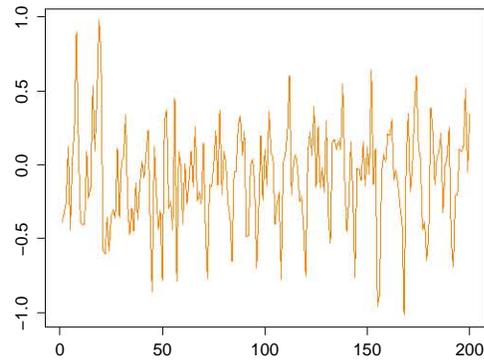


Figura 5.7: Inflación mexicana desde 1990 al aplicarle la diferenciación fraccional.

Si nuestra hipótesis es correcta, la nueva serie debe comportarse como un ruido blanco. Para probarlo podemos ver el ACF de la serie transformada (Figura 5.8).

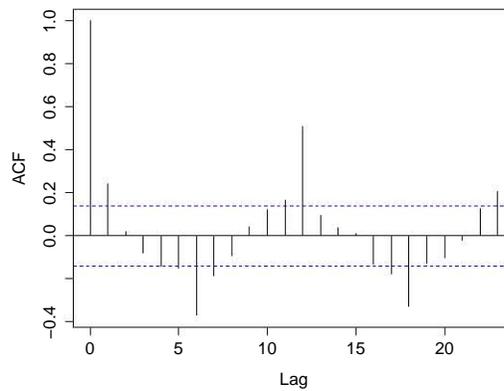


Figura 5.8: ACF de la inflación mexicana desde 1990 al aplicarle la diferenciación fraccional.

Ahora ajustaremos un GARCH a la serie trasformada, de la misma forma que en el Capítulo 2 pero ahora las innovaciones serán gaussianas. El resultado se observa en el cuadro 5.2 y las características de los residuales en la Figura 5.9.

Cuadro 5.2: Parámetros estimados

	Estimación	Error estándar	t value
$\mu$	-0.0847	0.02870	0.666
$\alpha_0$	0.1031	0.1068	0.965
$\alpha_1$	0.1214	0.2221	0.546
$\beta_1$	0.00000001	1.079	0.0000

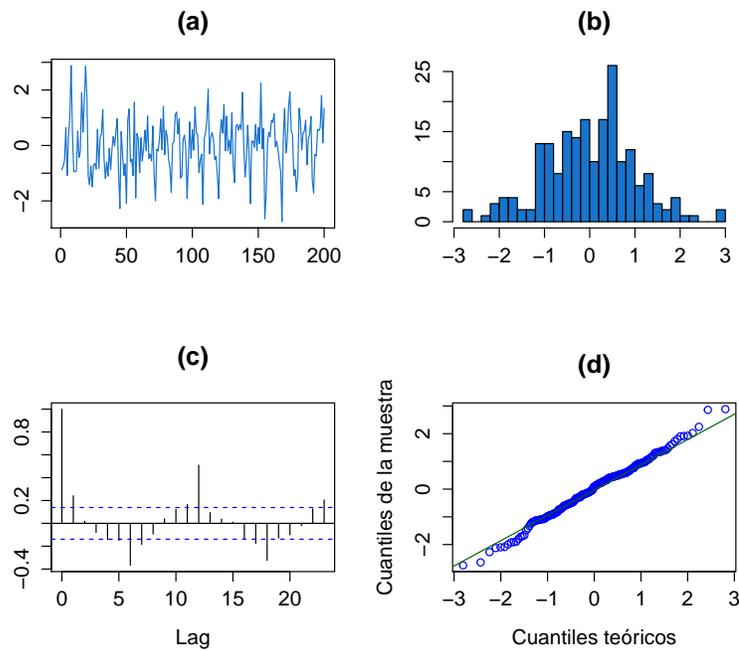


Figura 5.9: (a)Residuales del ajuste (b)Histograma de los residuales del ajuste (c)ACF de los residuales del ajuste (d) QQplot de los residuales del ajuste contra una normal estándar.

Incluso los residuales parecen tener un buen ajuste con la distribución normal.

Concluimos que un ARFIMA(0,  $d$ , 0)–GARCH(1, 1) es el modelo indicado para la serie y los parámetros son:

$$\begin{aligned}(1 - B)^{0.3845}(X_t - 0.8499) &= \epsilon_t \\ \epsilon_t &= \sigma_t Z_t \\ \sigma_t^2 &= 0.1031 + (0.1214)\epsilon_{t-1}^2 + (0.00000001)\sigma_{t-1}^2\end{aligned}$$

para  $(Z_t)_{t \in \mathbb{Z}}$  un  $SWN(0, 1)$  con distribución normal.

### 5.2.1. Medidas de riesgo

La primera cuestión importante que debemos tratar es aclarar qué significa el VaR para la inflación. El  $\text{VaR}_\alpha$  es el nivel que la tasa del aumento de los precios de los bienes y servicios no excederá con  $(100 \times \alpha)\%$  de confianza.

Ahora vamos a calcularlo. Como las innovaciones son un proceso GARCH podemos estimar la volatilidad del proceso  $\epsilon_t$  en el tiempo.

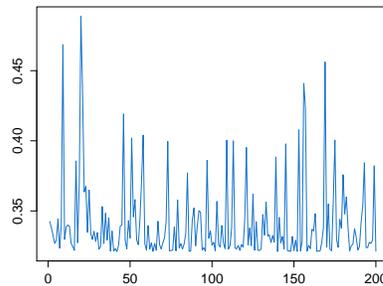


Figura 5.10: Volatilidad estimada de la serie fraccionalmente diferenciada.

Y la estimación de la volatilidad para el siguiente tiempo (el siguiente mes):

$$\hat{\sigma}^2 = 0.1173289$$

Ahora, usando (4.7) y (4.8) podremos calcular las medidas de riesgo para la inflación mexicana con niveles de 90 % y 95 % de confianza:

Cuadro 5.3: Riesgo de la inflación

VaR <sub>90%</sub>	=	1.00956
ES <sub>90%</sub>	=	1.171726
VaR <sub>95%</sub>	=	1.134003
ES <sub>95%</sub>	=	1.277134

### 5.2.2. Back-testing para el VaR

Después de calcular las medidas de riesgo de la serie debemos ver que tan bueno es el modelo que tenemos. Para probar el modelo se realiza una prueba de Back-testing (ver A.2) para el VaR. Haremos el Back-testing sobre los últimos 100 datos de la serie. Para realizar el Back-testing utilizamos el siguiente código:

```

backtest_var <- function(x, niv) {
  m <- mean(x)
  c <- (x - m)
  db <- 0
  var <- c(0)
  q <- qnorm(niv)
  for(i in 1:w) {
    xb <- c[(288 - 280 + i):(288 - 100 + i)]
    db <- fdGPH(xb, 0.6)$d
  }
}

```

```

xbd<-difffrac(xb,db,30)
xbd<-ts(xbd)
n<-length(xb)
n1<-length(xbd)
fit=garchFit(formula= garch(1, 1),data=xbd)
m2 <-fit@fit$par[1]
a_0<-fit@fit$par[2]
a_1<-fit@fit$par[3]
b_1<-fit@fit$par[4]
v<-c(0)
v[1]<-0
for(h in 2:(n1+1)){
v[h]<- (a_0)+(a_1*(xbd[h-1]^2))+(b_1*(v[h-1]))}
k<-0
for(j in 1:20){
k <- k+((-1)^(j)*(((g(db+1))/(g(j+1)*g(db-j+1)))*xb[n+1-j]))}
var[i] <- m2 + m + (sqrt(v[n1+1])*q) - k }
return(var)}
for(i in 1:10){b90<-backtest_ var(x,0.89+i)
plot(xx-b90,type="h")
abline(h=0,col="red")}

```

Este código nos arroja los excedentes que resultaron del Back-testing, el el p-valor del binomial test (ver A.2) y la forma gráfica del Back-testing.

Los resultados para la serie que estamos estudiando y con el modelo dado son los siguientes:

El cuadro 5.4 nos muestra un resumen del Back-testing realizado y nos también nos da el p-value del binomial test los diferentes niveles de confianza.

Cuadro 5.4: Back-testing para el VaR de 90 % a 99 %

	Excedentes	Excedentes esperados	p-valor (binomial test)
VaR <sub>0.90</sub>	10	10	0.9999
VaR <sub>0.91</sub>	8	9	0.8619
VaR <sub>0.92</sub>	7	8	0.8545
VaR <sub>0.93</sub>	7	7	0.9999
VaR <sub>0.94</sub>	7	6	0.6704
VaR <sub>0.95</sub>	6	5	0.6418
VaR <sub>0.96</sub>	5	4	0.6033
VaR <sub>0.97</sub>	3	3	0.9999
VaR <sub>0.98</sub>	3	2	0.4559
VaR <sub>0.99</sub>	1	1	0.9999

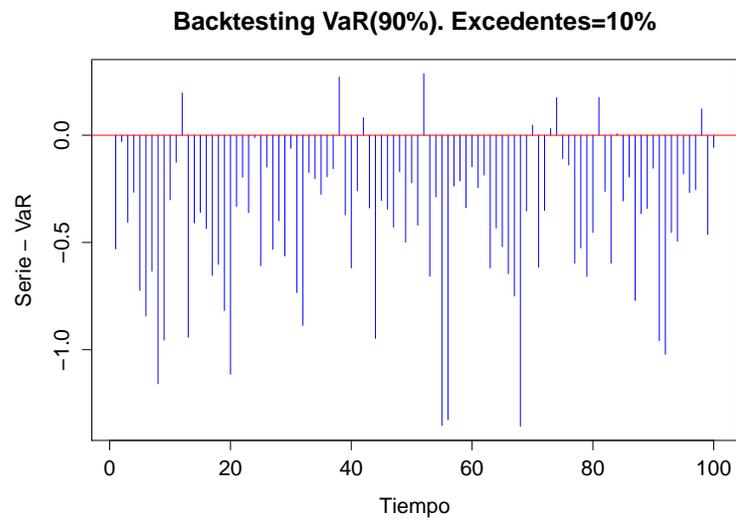


Figura 5.11: Resultado gráfico del back-testing del VaR a 90 %.

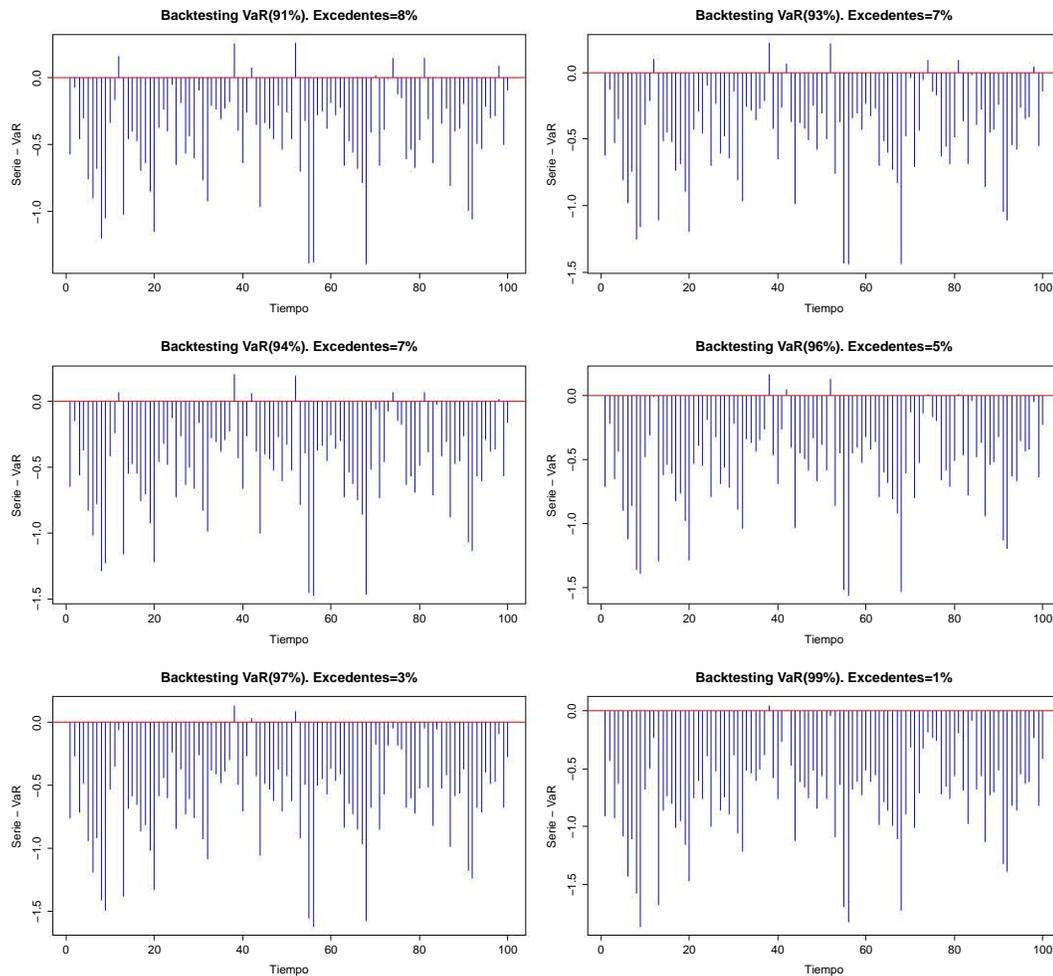


Figura 5.12: Resultado gráfico del back-testing del VaR a 95 %.

Las Figuras 5.11 y 5.12 nos muestran gráficamente el resultado del back-testing del modelo. Las gráficas muestran a la serie de tiempo menos el VaR calculado, los puntos que exceden de cero son los puntos en los que la serie excedió en VaR. La idea es que la proporción de excedentes sea  $1 - \alpha$ , donde  $\alpha$  es el nivel de confianza a que se está calculando en VaR.

### 5.3. Conclusiones

Sabemos que el  $\text{VaR}_\alpha$  de la inflación es la tasa de inflación más alta que podemos observar con  $(100 \times \alpha) \%$  de confianza. Por lo que podemos decir que: con una confianza del 95 % la tasa de inflación del siguiente mes (febrero de 2014) no será mayor a 1.134003 %.

El Expected-shortfall, al nivel  $\alpha$  de confianza, es el valor esperado de la inflación si llegase a ocurrir que esta fuera mayor al  $\text{VaR}_\alpha$ . Es decir, si la inflación de febrero de 2014 llega a ser mayor que 1.134003 %, entonces tendrá un valor esperado de 1.277134 % con 95 % de confianza.

Para finalizar este trabajo mencionaremos algunas conclusiones y comentarios:

1. El análisis exploratorio de los datos es la llave para la determinación de modelo que debemos ajustar. En el capítulo 2 las evidencias de no correlación de la serie observadas en el ACF, las pruebas de leptokurtosis que arrojaron el histograma y la estimación de la kurtosis, y la observación que hicimos de que la varianza de la serie no podría ser constante, fueron los indicadores de que un modelo GARCH sería el ideal para realizar el ajuste.

En el capítulo 5 el análisis exploratorio de los datos mostró que la correlación serial de los datos es bastante fuerte así como pruebas de una distribución leptokurtica y de volatilidad no constante en el tiempo; por lo que un modelo que podría servir para el ajuste sería uno de memoria larga con volatilidad estocástica.

2. Las evidencias de memoria larga de una serie serían principalmente dos: una fuerte correlación serial y que la media muestral parece no ser constante en el tiempo (como muestra el corolario del Teorema de Féjèr en (3.11)).
3. Una vez establecido el modelo, debe buscarse no hacer una sobreparametrización. Es decir, cuando sabemos qué modelo vamos a ajustar, debemos

buscar su forma más sencilla posible para hacer el ajuste.

4. Los resultados del Back-testing verifican que el modelo que construimos para la estimación del VaR de la inflación es adecuado. Aunque es claro que este modelo debe supervisarse para darle mantenimiento ante posibles cambios en la estructura de serie.

# Apéndice A

## Medidas de riesgo

Primero hablaremos de manera no matemática de las diversas cuestiones que forman los antecedentes de los desarrollos que se han hecho en la gestión de riesgos.

Hablaremos del riesgo, su definición y características; también, un poco de las regulaciones mundiales actuales en cuanto a gestión de riesgos.

### A.1. Gestión de riesgos

Se define el riesgo como "peligro, una oportunidad de malas consecuencias, la exposición a la mala suerte", esto para el caso general, es decir, un riesgo en cualquier perspectiva.

Ahora para el caso de riesgos financieros podemos dar la siguiente definición: "la posibilidad de ocurrencia de cualquier evento o acción que pueda afectar negativamente la capacidad de una organización para alcanzar sus objetivos y ejecutar sus estrategias" o "la cuantificable probabilidad de pérdida o de tener menos de lo esperado".

Estas definiciones pueden captar algunos elementos de riesgo, pero con una sola frase no tendremos una definición satisfactoria en todos los contextos y perspectivas.

El riesgo se relaciona fuertemente con la incertidumbre y por tanto con la probabilidad. Es aquí donde entra la gestión de riesgos, el estudio y aplicación de modelos matemáticos que midan matemáticamente el riesgo (lo que sea que definamos como tal).

La medición del riesgo de algo es -esencialmente- determinar su distribución de probabilidades. Para lograr esto, necesitamos un modelo calibrado correctamente para estimar dicha distribución.

Según Kloman (1990) la gestión de riesgos se trata de:

*Para muchos analistas, políticos y académicos es la gestión de riesgos ambientales y nucleares, los riesgos macroeconómicos generados por la tecnología que parece poner en peligro nuestra existencia. Para los banqueros y financieros oficiales es el sofisticado uso de técnicas tales como la cobertura de divisas y swaps de tasas de interés. Para los compradores o vendedores de seguros es la coordinación de riesgos asegurables y la reducción de los costos del seguro.*

***En resumen, la gestión de riesgos es una disciplina para vivir con la posibilidad de que los eventos futuros pueden causar efectos adversos.***

## La regulación

El gestionar el riesgo es importante para evitar pérdidas monetarias entonces, al hacer gestión de riesgo se tiene un cierto valor monetario en juego. Los dueños de estos valores querrán que quien administre los riesgos de sus bienes sea un profesional capacitado y que siga cierta normativa, aceptada y dada por expertos, sobre cómo se debe gestionar el riesgo. Esta normativa es la regulación de la gestión de riesgos.

La regulación se remonta, al menos, a la época de los bancos venecianos y las empresas de seguros de Londres en el siglo XVIII. En esos días se dependía en gran medida en la autorregulación o regulación local, pero las reglas estaban allí. Sin embargo, las principales novedades que han dado lugar a la presente reglamentación, marco de gestión, del riesgo son una historia del siglo XX.

El primer Acuerdo de Basilea fue la primera regulación global (o que pretendía ser global) que aparece. En 1988 la Supervisión Bancaria dio un paso importante hacia un estándar mínimo internacional de capitales. El énfasis principal del acuerdo estaba en riesgo de crédito, entonces era claro que era la fuente más importante de riesgo en el sector bancario. En retrospectiva, sin embargo, el primer Acuerdo de Basilea adoptó un enfoque que era bastante grueso y el riesgo fue medido de manera insuficiente. También el tratamiento de los derivados se consideró insatisfactorio.

Este acuerdo trajo consigo muchas discusiones sobre la regulación internacional y también la generación de otros acuerdos que intentan cubrir las deficiencias de los anteriores. Entre ellos surgieron Basilea 2 y 3 y los acuerdos de Solvencia.

## El nacimiento del VaR

Los eventos de gran impacto (negativo) nos llevan a desarrollar medidas de protección. En este caso fueron varios eventos, iniciando con el colapso del mercado de valores de 1987 y las pequeñas crisis que le siguieron en poco tiempo. El golpe llegó a una gran variedad de mercados, incluyendo algunos poco correlacionados; esto dio la idea de que se trataba de un evento extremo.

Entonces nació Value-at-Risk (VaR) como medida del riesgo de mercado que estableció un estándar en toda la industria. Ahora discutiremos sobre los principales conceptos que se utilizan en la administración de riesgos, pondremos principal atención en los modelos modernos de medición de riesgo. Hablaremos de sus debilidades y fortalezas; particularmente analizaremos los casos del "Value-at-Risk" y del "expected shortfall".

### A.2. Algunas definiciones

Consideraremos a un portafolios de instrumentos financieros, cualesquiera, cuyo valor en el tiempo es una variable aleatoria. Denotaremos el valor de este portafolios a tiempo,  $t$ , como  $V(t)$ .

#### **Definición 24.**

*Supongamos que tenemos un portafolios entonces, para un horizonte de tiempo dado, digamos  $\Delta$ , y para un tiempo  $s$  definimos a la **pérdida del portafolios** como:*

$$L(s, s + \Delta) = -(V(s + \Delta) - V(s))$$

$L(s, s + \Delta)$  es una variable aleatoria cuya distribución suele llamarse distribución de pérdidas.

### A.2.1. Medidas de riesgo

Para lo siguiente supondremos que la pérdida  $L$  es una variable aleatoria definida sobre cierto espacio de probabilidad,  $(\Omega, \mathcal{F}, \mathbb{P})$ , para cierto horizonte de tiempo  $\Delta$ .

Denotaremos por  $L^0(\Omega, \mathcal{F}, \mathbb{P})$  al espacio de todas las variables definidas en  $(\Omega, \mathcal{F})$  que son casi seguramente finitas, en riesgo financiero se suele usar el conjunto  $\mathcal{M} \subset L^0(\Omega, \mathcal{F}, \mathbb{P})$  a las variables que interpretamos como la pérdida de un portafolios en cierto horizonte de tiempo.

Asumiremos que  $\mathcal{M}$  es un cono convexo, es decir, que se cumple que:

1. Si  $L_1 \in \mathcal{M}$  y  $L_2 \in \mathcal{M}$  entonces:  $L_1 + L_2 \in \mathcal{M}$ .
2. Si  $L_1 \in \mathcal{M}$  y  $\lambda \in \mathbb{R}$  entonces:  $\lambda L_1 \in \mathcal{M}$ .

Con estas ideas, daremos los axiomas de coherencia para las medidas de riesgo. Una medida de riesgo es una función  $\varrho : \mathcal{M} \rightarrow \mathbb{R}$ .

Decimos que  $\varrho$  es una medida de riesgo coherente, si cumple los siguientes axiomas.

**Axioma 1.** *Invarianza bajo traslación.*

*Para todas  $L \in \mathcal{M}$  y  $\lambda \in \mathbb{R}$  se tiene que:  $\varrho(L + \lambda) = \varrho(L) + \lambda$ .*

**Axioma 2.** *Subaditividad.*

*Para todas  $L_1, L_2 \in \mathcal{M}$  se tiene que:  $\varrho(L_1 + L_2) \leq \varrho(L_1) + \varrho(L_2)$ .*

**Axioma 3.** *Homogeneidad positiva.*

*Para todas  $L \in \mathcal{M}$  y  $0 < \lambda \in \mathbb{R}$  se tiene que:  $\varrho(\lambda L) = \lambda \varrho(L)$ .*

**Axioma 4.** *Monotonicidad.*

*Para todas  $L_1, L_2 \in \mathcal{M}$  con  $L_1 \leq L_2$  se tiene que:  $\varrho(L_1) \leq \varrho(L_2)$ .*

### A.2.2. VaR

La medida de riesgo probablemente más usada en administración de riesgos es el VaR (Valor en Riesgo). Incluso en la legislación de Basilea II se menciona para el cálculo del marco de capital adecuado.

Consideremos un cierto portafolios de instrumentos financieros sujeto al riesgo. Denotaremos a su respectiva distribución de pérdidas como:

$$F_L(l) = P[L \leq l]$$

Queremos definir una estadística, que sea función de  $F_L$ , que mida la gravedad del riesgo de la celebración de nuestra cartera. Un candidato obvio es la pérdida máxima posible, dada por:

$$\inf\{l \in \mathbb{R} : F_L(l) = 1\}$$

Esta es una medida de riesgo importante en el reaseguro. Sin embargo, en la mayoría modelos de interés el soporte de  $F_L$  no está acotado entonces, la pérdida máxima sería infinita.

El Valor en Riesgo es una extensión directa de la pérdida máxima, tomando en cuenta sus problemas.

La idea es simplemente reemplazar "pérdida máxima" por "la pérdida máxima que no sea excedida con una probabilidad dada", esta probabilidad es llamada nivel de confianza.

Existen varias definiciones de VaR, según la forma de plantear el problema. La definición usual de VaR es la siguiente.

**Definición 25.**

Dado un cierto nivel de confianza  $\alpha \in (0, 1)$ , definimos el **Valor en Riesgo**,  $\text{VaR}_\alpha$ , de una pérdida  $L$  como:

$$\text{VaR}_\alpha(L) = \inf\{l \in \mathbb{R} : P[L > l] \leq (1 - \alpha)\}$$

El VaR de la cartera con nivel de confianza  $\alpha$  es el número más pequeño de tal manera que la probabilidad de que la pérdida lo exceda no sea mayor que  $(1 - \alpha)$ .

En términos probabilísticos, el VaR es simplemente un cuantil de la distribución de pérdidas con la propiedad:

$$F_L(\text{VaR}_\alpha) = \alpha$$

Los valores típicos de  $\alpha$  son 0.90, 0.95 ó 0.99, y en gestión de riesgo de mercado el horizonte de tiempo es generalmente 1 ó 10 días. En la gestión del riesgo de crédito y operacional es por lo general un año.

Tengamos en cuenta que por su propia definición el VaR no nos da ninguna información sobre la gravedad de las pérdidas que pueden ocurrir con una probabilidad menor a  $(1 - \alpha)$ . Esto es, claramente, una desventaja de VaR como medida del riesgo.

Lamentablemente esta medida de riesgo no cumple los axiomas de coherencia pues no cumple el Axioma 2, nos es subaditivo.

A pesar de este problema se utiliza con gran frecuencia el VaR, en ocasiones por la común suposición de normalidad en la distribución de los activos que se quiere modela, esto pasa porque bajo una distribución normal el VaR si cumple los axiomas de coherencia.

### A.2.3. Expected shortfall

La otra medida que veremos está dada en términos del VaR, pero esta vez la medida no tomará sus desventajas, sino solo sus ventajas para hacer una excelente medida de riesgo.

**Definición 26.**

Dado un cierto nivel de confianza  $\alpha \in (0, 1)$ , definimos el **Expected shortfall**,  $ES_\alpha$ , de una pérdida  $L$  como:

$$\begin{aligned} ES_\alpha(L) &= \mathbb{E}[L | L > \text{VaR}_\alpha(L)] \\ &= \frac{1}{1 - \alpha} \int_\alpha^1 \text{VaR}_u(L) du \end{aligned}$$

donde  $F_L$  es la distribución de la pérdida.

Puede demostrarse que el Expected shortfall es una medida de riesgo coherente.

# Apéndice B

## Conceptos de estadística

En este apéndice revisaremos algunos de los conceptos estadísticos que fueron usados para el desarrollo y conclusiones del trabajo.

### B.1. t-value

El t-value es una herramienta que se usó como test de los ajustes de los parámetros de los modelos. Es realmente muy sencillo de calcular. Supongamos que tenemos una muestra de datos  $x_1, \dots, x_n$  y que queremos estimar un parámetro de la distribución de la muestra, digamos  $\beta$ , con la estadística  $\hat{\beta}$ .

El Student's t-test consiste en calcular el valor:

$$t_{\hat{\beta}} = \frac{\hat{\beta} - \beta_0}{s.e.(\hat{\beta})}$$

donde  $s.e.(\hat{\beta})$  es el error estándar de  $\hat{\beta}$  y  $\beta_0$  es una constante. Suele usarse (y nosotros usamos)  $\beta_0 = 0$ .

Cuando  $\hat{\beta}$  es un estimador obtenido de un modelo de regresión, entonces  $t_{\hat{\beta}}$  tiene una distribución t de Student con  $(n - k)$  grados de libertad, donde  $k$  es el

número de regresores. En el caso de nuestras estimaciones  $k = 1$ .

El valor  $P[> |t|]$  es entonces,  $2(1 - P[X < t_{\hat{\beta}}])$ , para  $X$  una variable aleatoria con distribución t de Student y  $(n - 1)$  grados de libertad. Es decir  $P[> |t|]$  es la probabilidad de caer en la región de rechazo.

## B.2. p-value

El p-value es otra herramienta estadística que nos ayuda a calificar una estimación. Pero en este caso el p-value sí nos dará directamente una probabilidad.

Este valor deberá ser interpretado de manera contraria al t-value, es decir, será la probabilidad caer en la región de no rechazo.

Supongamos, nuevamente, que tenemos una muestra de datos  $x_1, \dots, x_n$  y tenemos la estimación de un parámetro de la distribución de la muestra, digamos  $\beta$ , con la estadística  $\hat{\beta}$  pero ahora queremos hacer una prueba de hipótesis, digamos,  $H_0 : \beta = \hat{\beta}$ .

La prueba se realizará con una estadística de la muestra, digamos  $T = T(x_1, \dots, x_n)$  que se supone tiene la misma distribución que cierta variable  $X$ . Entonces el p-value es:

$$P[X > T]$$

Entonces el p-value es el nivel de significancia de la prueba, el nivel de confianza con el que se no se rechaza la hipótesis.

### B.3. QQplot

Cuando se hacen ajustes de modelos teórico a datos se pueden realizar diversas pruebas para evaluar el ajuste. Una de las pruebas más simples, ilustrativas y usadas es la comparación de cuantiles, la gráfica QQplot.

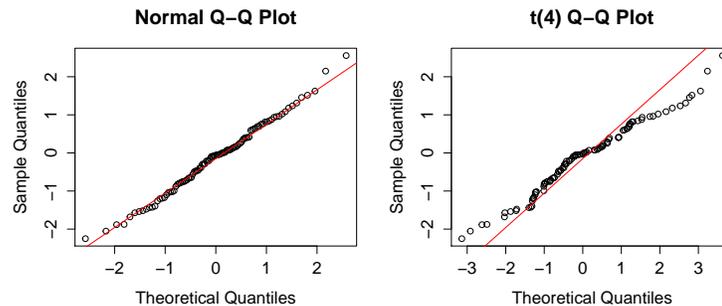
Dada una muestra (datos)  $x_1, \dots, x_n$  y una distribución teórica,  $F$ , que se cree es la distribución de los datos, se construye el QQplot de la siguiente forma:

1. Se ordena la muestra de menor a mayor:  $x_{(1)}, \dots, x_{(n)}$ . Entonces,  $x_{(i)}$  debe ser el cuantil  $\frac{i}{n}$ . Habitualmente se considera al cuantil  $\frac{i-0.5}{n}$ .
2. Se obtienen los cuantiles teóricos de la distribución que se quiere ajustar a los datos. Los cuantiles a obtener serían:  $q_1, \dots, q_n$  tales que:

$$F(q_i) = \frac{i - 0.5}{n}$$

3. Se grafican los puntos  $(q_i, x_{(i)})$ . Si el ajuste es correcto, entonces la gráfica se parecerá a la identidad.

La siguiente gráfica muestra un ejemplo, el QQplot de una muestra aleatoria normal contra una normal y contra un  $t$  con 4 grados de libertad.



## B.4. Back-testing para el VaR

Supongamos que tenemos una muestra de datos  $x_1, \dots, x_n$  y que queremos estimar alguna función de la distribución de la variable  $X_{n+1}$ . Cualquier modelo que estime este valor debe ser calificado, es decir, debe probarse qué tan bueno es el modelo. Para esto se puede hacer una prueba de Back-testing.

Esta prueba consiste en suponer que no se conocen los últimos  $k$  datos de la serie ( $k < n$ ) y con los  $(n - k)$  datos restantes hacer la estimación de la función buscada, y después comparar los resultados con lo que pasó en la realidad (los  $k$  datos).

Para el caso en que dicha función es el VaR, el Back-testing será el siguiente:

1. Se estiman los  $k$  valores del VaR (uno por cada dato que estamos omitiendo) con el modelo establecido, obteniendo el vector  $(\text{VaR}_\alpha(n - k + 1), \dots, \text{VaR}_\alpha(n))$ .
2. Se calculan los valores  $(\text{VaR}_\alpha(n - k + 1) - x_{n-k+1}), \dots, (\text{VaR}_\alpha(n) - x_n)$ . Estos valores producen la gráfica de excedentes del modelo.
3. Se calcula la estadística

$$T = \sum_{i=1}^k \mathbb{I}_{\{x_{n-k+i} > \text{VaR}_\alpha(n-k+i)\}}$$

Este valor debe ser la cantidad de datos de la serie que fueron mayores que el VaR que se estimó. De estar bien construido el modelo el porcentaje de datos que exceden su VaR estimado debe ser  $1 - \alpha$ , es decir,  $1 - \alpha \approx \frac{T}{k}$ .

4. Si el modelo es correcto, entonces la variable  $T$  debe comportarse como una Binomial con parámetros  $(k, 1 - \alpha)$ . Entonces, se realiza una prueba para corroborar que  $T \sim \text{Bin}(k, 1 - \alpha)$ . La prueba realizada es un Binomial test.

**Binomial test**

Sea  $X_1, \dots, X_n$  una muestra aleatoria de una población Bernoulli( $p$ ). Consideremos probar la hipótesis  $H_0 : p < p_0$  versus  $H_1 : p > p_0$  donde  $0 < p_0 < 1$  es un valor dado. El estimador máximo verosímil de  $p$  en una muestra de tamaño  $n$  viene dado por  $\hat{p}_n = \sum_{i=1}^n X_i/n$ . Dado que  $\hat{p}_n$  es la media de la muestra, tenemos que por el *Teorema del Límite Central* que para cualquier  $p$ ,  $0 < p < 1$ ,  $(\hat{p}_n - p)/\sigma_n$  converge a una variable aleatoria Normal estándar. Aquí tomamos  $\sigma_n = \sqrt{p(1-p)/n}$ , la cual depende claramente del parámetro desconocido  $p$ .

Un estimador razonable para  $\sigma_n$  es  $S_n = \sqrt{\hat{p}_n(1-\hat{p}_n)/n}$ . Entonces para cualquier  $p$  tal que  $0 < p < 1$  tenemos lo siguiente:

$$\frac{\hat{p}_n - p}{\sqrt{\hat{p}_n(1-\hat{p}_n)/n}} \rightarrow N(0, 1)$$

El estadístico para el test de Wald,  $Z_n$ , es definido reemplazando  $p$  por  $p_0$  y rechazaremos  $H_0$  si  $Z_n > z_\alpha$ .

Si tuviéramos interés en hacer una prueba de dos colas, es decir,  $H_0 : p = p_0$  versus  $H_1 : p \neq p_0$  donde  $0 < p < 1$ , podríamos volver a aplicar la misma estrategia anterior. Por el *Teorema del Límite Central* para cualquier  $0 < p < 1$ ,

$$\frac{\hat{p}_n - p}{\sqrt{p(1-p)/n}} \rightarrow N(0, 1)$$

Por lo tanto, si la hipótesis nula es verdadera, el estadístico

$$Z'_n = \frac{\hat{p}_n - p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0, 1) \text{ (aproximadamente)} \tag{B.1}$$

Entonces a un nivel  $\alpha$ , rechazamos  $H_0$  si  $|Z'_n| > z_{\alpha/2}$

# Bibliografía

- [1] BAILLIE, R.T. (1996), *Long memory processes and fractional integration in econometrics*, Journal of Time Series Analysis, 73, 5-59.
- [2] BAILLIE, R.T., BOLLERSLEV, T., y MIKKELSEN, H.O. (1996), *Fractionally integrated generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics, 74, 3-30.
- [3] BOLLERSLEV, T.P. (1986), *Generalized Autoregressive Conditional Heteroskedasticity*, Journal of Econometrics, 31, 307-327.
- [4] BREIDT, F.J., CRATO, N. y DE LIMA P.(1998), *The Detection and Estimation of Long Memory in Stochastic Volatility*, Journal of Time Series Analysis, 73, 325-48.
- [5] BROCKWELL P.J. y DAVIS R.A. (2002), *Introduction to Time Series and Forecasting*, segunda edición, Colorado, USA. Springer.
- [6] CHRONOPOULOU, A. y VIENS, F.G., *Estimation and Pricing under Long-Memory Stochastic Volatility*, USA, Department of Statistics, University St. West Lafayette.
- [7] ENGLE, R. (1982), *Autoregressive Conditional Heterosedasticity with Estimates of the Variance of United Kingdom Inflation*, Econometrica, vol. 50, issue 4.

- [8] GEWEKE, J. y PORTER-HUDAK S.(1983), *The Estimation and Application of Long-memory Time Series Models*, Journal of Time Series Analysis, 4, 221-38.
- [9] GRANGER C.W.J. y JOYEUX R. (1980), *An Introduction to Long-memory Time Series Model and Fractional Differencing*, Journal of Time Series Analysis, 1, 15-29.
- [10] GUEGAN, D. (2007), *How can we define the concept of long memory? an econometric survey*, Francia, Département d'Economie et Gestion.
- [11] MCNEIL A.J., FREY R. y EMBRECHTS P.(2005), *Quantitative Risk Management: Concepts, Techniques and Tools*, primera edición, USA, Princeton University Press.
- [12] OOMS, M., y DOORNIK, J. (1999), *Inference and Forecasting for Fractional Autoregressive Integrated Moving Average Models, with an application to US and UK inflation*, Econometric Institute Report, Erasmus University Rotterdam.
- [13] ORTEGA PÉREZ DE LEÓN, ARMANDO. ET. AL. (1982), *Inflación: Estudio Económico, Financiero y Contable*, Instituto Mexicano de Ejecutivos de Finanzas, A.C. México.
- [14] PALMA, W. (2007), *Long-Memory Time Series*, primera edición, Chile, Pontificia Universidad Católica de Chile.
- [15] QIAN, B. y RASHEED, K., *HURST EXPONENT AND FINANCIAL MARKET PREDICTABILITY*, USA, Department of Computer Science, University of Georgia.
- [16] ROBINSON, P. (2003), *Time series with long memory*, primera edición, UK, Oxford University Press.
- [17] SAMORODNITSKY, G., *Long Memory and Self-Similar Processes*.

- [18] WEI, W.W.S., (2006), *Time Series Analysis: Univariate and multivariate methods*, segunda edición, USA.
- [19] WILKINS, N., (2007), *Indirect Estimation of Long Memory Volatility Models*, Australia, School of Finance and Economics, University of Technology, Sydney.
- [20] XIE, L. (2006), *Long-Memory Stochastic Volatility Models: A New Method for Inference and Applications in Option Pricing*.
- [21] ZAMBRENO, J.A. (2012), *Análisis Comparativo de los modelos ARFIMA y ARMA: Inflación y Actividad Económica en Venezuela*, Universidad Autónoma de Barcelona.
- [22] ZONGWU, C., (2007), *Advanced Topics in Time Series Econometrics Using R*, USA, Department of Mathematics & Statistics and Department of Economics, University of North Carolina, Charlotte.