



**LOS LIBERTADORES**  
FUNDACIÓN UNIVERSITARIA

**PROPUESTA DE UN MODELO DE PRONÓSTICO PARA LA  
VOLATILIDAD DE LOS RETORNOS DE LA ACCIÓN CEMARGOS  
DURANTE EL PERIODO 08/JUNIO/2012 AL 17/MARZO/2017**

CHRISTIAN CAMILO CORTES GARCIA

**FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES**  
DEPARTAMENTO DE CIENCIAS BÁSICAS  
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA  
BOGOTÁ D.C.  
2017

**PROPUESTA DE UN MODELO DE PRONÓSTICO PARA LA  
VOLATILIDAD DE LOS RETORNOS DE LA ACCIÓN CEMARGOS  
DURANTE EL PERIODO 08/ENERO/2012 AL 17/MARZO/2017**

CHRISTIAN CAMILO CORTES GARCIA

Trabajo de grado realizado para obtener el título de  
Especialista en Estadística Aplicada

Asesor  
Juan Camilo Santana  
Magister en Estadística

**FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES**  
DEPARTAMENTO DE CIENCIAS BÁSICAS  
ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA  
BOGOTÁ D.C.  
2017

Nota de aceptación

---

---

---

---

---

---

---

---

Firma del presidente del jurado

---

Firma del jurado

---

Firma del jurado

Bogotá D.C, 27 de Mayo de 2017

Las directivas de la Fundación Universitaria Los  
Libertadores, los jurados calificadores y el cuer-  
po docente no son responsables por los criterios e  
ideas expuestas en el presente documentos. Esto  
corresponde únicamente al autor.

# Agradecimientos

Doy gracias a mi familia por su apoyo incondicional en este proceso y a mi asesor por su buena disposición y el enriquecimiento a mi trabajo.

# Resumen

En este trabajo se presenta una breve introducción a los instrumentos estadísticos y modelos necesarios para explicar la volatilidad de los precios de activos, siguiendo la metodología de series temporales, involucrando el efecto de heterocedasticidad condicional. Con estos lineamientos definidos, se modela la volatilidad de los retornos de precios de cierre diarios de las acciones de la empresa colombiana de Cementos Argos S.A, tomando como referencia los modelos ARCH, GARCH, TGARCH, IGARCH, EGARCH y APARCH, e incorporar back-testing, realizando contraste con un modelo  $SVt$ -AR(1), para determinar la efectividad de los modelos por fuera de la muestra. El modelo que mejor explica la volatilidad condicional de los retornos es el EGARCH(1,1) y el modelo que mejor realiza pronósticos de volatilidad es el  $SVt$ -AR(1).

**Palabras Claves:** *Volatilidad condicional, volatilidad estocastica, efecto ARCH, retornos diarios, distribución GED, Criterios de información bayesiano.*

# Índice general

<b>1. Introducción</b>	<b>10</b>
1.1. Planteamiento y formulación del problema . . . . .	10
1.2. Objetivos . . . . .	12
1.2.1. Objetivo General . . . . .	12
1.2.2. Objetivos Específicos . . . . .	12
1.3. Justificación . . . . .	12
<b>2. Marco de Referencia</b>	<b>13</b>
<b>3. Marco Teórico</b>	<b>14</b>
3.1. Distribuciones de probabilidad básicas . . . . .	14
3.1.1. Distribución uniforme . . . . .	14
3.1.2. Distribución normal . . . . .	15
3.1.3. Distribución <i>t</i> -student . . . . .	16
3.1.4. Distribución chi-cuadrado . . . . .	16
3.1.5. Distribución de error generalizada . . . . .	17
3.1.6. Distribución beta . . . . .	18
3.2. Serie estacionaria . . . . .	18
3.3. Autocorrelaciones en series de tiempo . . . . .	19
3.4. Modelos de volatilidad . . . . .	20
3.4.1. Modelos de volatilidad condicional . . . . .	23
3.4.2. Modelo de volatilidad estocástica . . . . .	27
<b>4. Marco Metodológico</b>	<b>30</b>
<b>5. Análisis y resultados</b>	<b>34</b>
5.1. Modelo de volatilidad condicional . . . . .	34
5.2. Modelo de volatilidad estocástica . . . . .	39
5.3. Modelo de pronóstico . . . . .	42
<b>Conclusiones</b>	<b>45</b>
<b>Bibliografía</b>	<b>46</b>
<b>Anexos</b>	<b>47</b>

# Índice de figuras

3.1.	Distribuciones de probabilidad clásica. . . . .	18
3.2.	Ejemplo de series temporales. . . . .	18
3.3.	Ejemplo de autocorrelaciones en series temporales. . . . .	19
3.4.	Efecto de apalancamiento. . . . .	24
4.1.	Boxplot por año precio CEMARGOS. . . . .	30
4.2.	Serie de precios y retornos CEMARGOS. . . . .	31
5.1.	Histograma de acciones CEMARGOS. . . . .	35
5.2.	ACP y PACF retornos CEMARGOS. . . . .	35
5.3.	Serie de retornos CEMARGOS con dos desviaciones estándar a un nivel de confianza del 95 %. . . . .	38
5.4.	ACF residuales estandarizados EGARCH(1,1). . . . .	39
5.5.	Residuales estandarizados EGARCH(1,1). . . . .	39
5.6.	Modelo SV-AR(1). . . . .	40
5.7.	Modelo SV $t$ -AR(1). . . . .	41
5.8.	Bandas de volatilidad a un nivel de confianza del 95 %. Rojo: SV-AR(1), verde: SV $t$ -AR(1). . . . .	41
5.9.	Pronóstico serie de retornos recortada con los últimos 20 datos. . . . .	41
5.10.	Bandas de volatilidad para la serie $R_t$ con 15 datos posteriores a línea verde $t = 1165$ . . . . .	42
5.11.	Pronóstico volatilidad de la serie $R_t$ para 15 datos. . . . .	42

# Índice de tablas

1.1. Agentes Accionarios Cementos Argos S.A. . . . . .	11
4.1. Descriptivos acciones CEMARGOS. . . . .	30
5.1. BIC modelos de volatilidad condicional para los retornos. . . . .	36
5.2. Estimación de parámetros modelos de volatilidad condicional. . . . .	37
5.3. Probabilidades obtenidas por la prueba de Ljung-Box a residuales estandarizados. . . . .	37
5.4. Estimaciones modelo SV-AR(1). . . . .	40
5.5. Estimaciones modelo SV $t$ -AR(1). . . . .	41
5.6. Pronósticos precios de cierre CEMARGOS. . . . .	43
5.7. Intervalos de confianza para los pronósticos precios de cierre CEMAR- GOS dados por el modelo SV $t$ -AR(1). . . . .	44

# Capítulo 1

## Introducción

En este capítulo se explica los motivos y la importancia que tiene los modelos de volatilidad, condicional y estocástica, por los accionistas al momento de invertir en una de las acciones más cotizadas por la bolsa de valores de Colombia: CEMARGOS.

### 1.1. Planteamiento y formulación del problema

La compañía colombiana de Cementos Argos S.A. (Argos, 2012) nació en Medellín en el año 1934 e inició producción en el año 1936, dedicándose principalmente a la explotación, producción y comercialización de cemento, concreto y cualquier otro material a base de cemento, cal o arcilla. El nombre de la compañía viene del acrónimo del apellido Arango, el cual pertenecía a la familia que propulsó el inicio de la empresa. Después de diferentes asociaciones con inversionistas de varias regiones del país, surgieron las siguientes compañías: Cementos del Valle en 1938, Cementos del Caribe en 1944, Cementos El Cairo en 1946, Cementos de Caldas en 1955, Tolcemento en 1972, Colclinker en 1974, Cementos Rioclaro en 1982 y adquiere participación accionaria en Cementos Paz del Río a mediados de los 90.

En 1998 empieza su internalización mediante la compra de diferentes plantas en países como Venezuela y en el 2005 se fusionan todas las compañías productoras y distribuidoras de cementos en Colombia bajo el nombre de Cementos Argos S.A. Hoy en día, su mercado se divide principalmente en tres regiones: Colombia, Estados Unidos y el Caribe, siendo el país cafetero el de mayor participación en los ingresos totales de la compañía.

En la actualidad, Cementos Argos cuenta con 1151672310 acciones ordinarias en circulación, las cuales están adquiridas en su mayoría por Inversiones Argos S.A. y Protección Fondo de Pensiones Obligatorias como se observa en la Tabla 1.1.

Por otro lado, las series temporales de alta frecuencia que se observan en estos tipos de mercados financieros no siguen una distribución normal, caracterizándose principalmente por ser asimétricas, leptocúrticas, elevada persistencia en la volatilidad, correlaciones en los retornos al cuadrado, entre otros. De igual forma, una medida importante en las finanzas es el riesgo asociado con un activo y la volatilidad es una

Accionista	Acciones	Participación
Inversiones Argos S.A.	715067729	62.09 %
Protección Fondo de Pensiones Obligatorias	87511387	5.44 %
Amalfi S.A.	62593875	3.93 %
Fondo de Pensiones Horizonte	32690797	2.84 %
Fondo de Pensiones Obligatorias Colfondos	27321747	2.37 %
ING Fondo de Pensiones	26391597	2.29 %
Fondo de Pensiones Obligatorias Sandia S.A.	11067024	0.96 %
Fondo de Pensiones Obligatorias Porvenir	11044623	0.96 %
JMRV & CIA S.C.A.	7641391	0.66 %
Azurita S.A.	5683635	0.49 %
Conscar & Cia S.C.A.	5414455	0.47 %
Scarpetta de Piedrahita Gloria	4605365	0.40 %
LIPU & CIA S.C.A.	4590450	0.40 %
Fondo de Pensiones Protección	4396502	0.38 %
Fondo de Cesantías Protección	3223625	0.28 %
MultiFund SKANDIA	2483499	0.22 %

Tabla 1.1: Agentes Accionarios Cementos Argos S.A.

medida de incertidumbre, objeto de estudio por los inversionistas al momento de invertir.

Algunas investigaciones suponen que la varianza es determinista o constante a lo largo del tiempo, fijando como estimador puntual la varianza muestral de los retornos del activo. Sin embargo, los hechos estilizados de los activos financieros, tales como acciones, bonos o tasas de cambio, sugieren que la volatilidad no es constante. Por lo tanto, asumir homoscedasticidad en el tiempo no es un supuesto muy realista.

Aunque la volatilidad de activos está bien definida, no es directamente observable en la práctica, pues se observa son los precios de un activo y sus derivados. El hecho de que la volatilidad no es directamente observable tiene varias implicaciones importantes en estudiar y modelar la volatilidad. Por lo tanto, para modelar las variaciones de las series financieras y acercarse un poco más a la realidad, se han desarrollado modelos, tanto de volatilidad condicional como de volatilidad estocástica, como los ARCH, GARCH, IGARCH, EGARCH, TGARCH, APARCH, SV-AR(1) y SVt-AR(1), los cuales permiten ajustar la serie de retornos de cualquier activo, involucrando la teoría de series temporales, con el fin de identificar la estructura condicional respecto al tiempo de cotización que se encuentra detrás de los retornos. La efectividad de los modelos de volatilidad condicional pueden observarse en los trabajos realizados por (Franco et al., 2007), (Awartani and Corradi, 2005) y (Gutiérrez and Gutiérrez, 2013).

Para los inversionistas, el análisis de los retornos de un activo presenta un resumen completo de la oportunidad de invertir, tanto los precios a futuros como el grado de incertidumbre, por esta razón, la pregunta de investigación planteada es la siguiente:

*¿Cómo explicar y pronosticar la volatilidad de los precios diarios de acciones CEMARGOS, tomando como referencia modelos de volatilidad condicional y estocástica?*

## **1.2. Objetivos**

### **1.2.1. Objetivo General**

Proponer un modelo de volatilidad, condicional o estocástica, que logre pronosticar la varianza de los retornos de precios de cierre diarios de las acciones CEMARGOS posteriores al 17/marzo/2017.

### **1.2.2. Objetivos Específicos**

- Definir modelos de volatilidad, tanto condicional como estocástica, que logren capturar las varianzas de los retornos de precios de cierre diario de las acciones CEMARGOS, comprendidos entre 08/junio/2012 al 17/marzo/2017.
- Realizar back-testing para observar y determinar la efectividad de los modelos de volatilidad condicional y estocástica por fuera de la muestra.
- Pronosticar los precios de cierre de acciones CEMARGOS después del 17/marzo/2017 y comparar con los precios ya establecidos por la bolsa de valores de Colombia.

## **1.3. Justificación**

Dado que las decisiones por los inversionistas conllevan un alto grado de incertidumbre al momento de invertir en los mercados financieros, equivalente al riesgo que enfrentan los activos, es importante modelar los efectos de estos riesgos para tomar decisiones. Este riesgo se asocia a la volatilidad, la cual se define como una medida de dispersión o variabilidad de la rentabilidad de un activo financiero como los índices bursátiles, precios de acciones o moneda.

Por tal motivo, el objetivo de este trabajo es pronosticar la volatilidad de los retornos de precios de cierre diarios de una de las acciones mas cotizadas en la bolsa de valores de Colombia, Cementos Argos S.A, durante el periodo 08/julio/2007 - 17/marzo/2017. Para ello, se aplica los modelos de volatilidad condicional como el ARCH, GARCH, TGARCH, IGARCH, EGARCH y APARCH, y seleccionar el modelo que mejor se ajusten al conjunto de datos en estudio a través del criterio de información bayesiana (BIC). Por ultimo, se realiza un contraste con los modelos SV-AR(1) o SVt-AR(1) para determinar su efectividad y pronósticos por fuera de la muestra.

## Capítulo 2

# Marco de Referencia

Para la ejecución de este trabajo se toman como referencias (Franco et al., 2007), (Awartani and Corradi, 2005) y (Gutiérrez and Gutiérrez, 2013) con el fin de validar la efectividad de los modelos de volatilidad condicional tratados a lo largo de este trabajo.

En (Franco et al., 2007) se estudió los efectos asimétricos y día de la semana en el índice de volatilidades VIX de la Chicago Board Option Exchange del 02/01/2003 al 30/03/2007, el cual utilizaron modelos GARCH asumiendo innovaciones gaussianas,  $t$ -student y de la distribución generalizada de los errores (GED). Se encontró que un modelo EGARCH(1,1) y un modelo TGARCH con regresores exógenos para el día lunes e innovaciones con distribución GED y  $t$ -Student, respectivamente, se ajustan bien a los datos. Ambos modelos indicaron que existen efectos asimétricos en la serie y que en el día lunes la volatilidad del mercado medida con el VIX es mayor que en los demás días.

En (Awartani and Corradi, 2005) se examinó la capacidad predictiva relativa de la muestra de diferentes modelos GARCH, con especial énfasis en el contenido predictivo del componente asimétrico. Como resultados, el modelo GARCH(1,1) es superado cuando se compara con la clase de GARCH asimétrica, mientras que no es superado en comparación con otros modelos GARCH que no permiten asimetrías.

Por ultimo, en (Gutiérrez and Gutiérrez, 2013) se evaluó la capacidad predictiva de una familia de modelos GARCH para predecir la volatilidad condicional de los rendimientos del petróleo Maya y Mezcla Mexicana de Exportación para el periodo 2 de enero de 1989 al 30 de diciembre de 2011. Los modelos de volatilidad se estimaron bajo tres distribuciones condicionales para las innovaciones: normal,  $t$ -student y GED. Los hallazgos empíricos señalan que el modelo EGARCH, con innovaciones  $t$ -student es más eficiente en el proceso de predicción de la volatilidad dentro de la muestra que los modelos basados en las distribuciones de errores generalizada y normal. Sin embargo, los resultados de la prueba de Diebold-Mariano muestran que los modelos GARCH, con innovaciones de errores generalizados y normales, superan el desempeño predictivo óptimo de los modelos GARCH con innovaciones  $t$ -student cuando se consideran las predicciones de la volatilidad fuera de la muestra.

# Capítulo 3

## Marco Teórico

En este capítulo se presenta una breve explicación de algunas distribuciones estadísticas, teoría básica de series temporales y algunos modelos de volatilidad condicional y estocástica utilizadas en el desarrollo de este trabajo. La teoría básica de series temporales y los modelos de volatilidad condicional fueron tomados de (Ghahlanos, 2013), (Tsay, 2014), (Wuertz et al., 2009) y los modelos de volatilidad estocástica de (Kastner, 2016a) y (Kastner, 2015).

### 3.1. Distribuciones de probabilidad básicas

Existen un conjunto de distribuciones que se han utilizado tradicionalmente en la modelación del comportamiento del retorno de un activo, dentro de las que se destacan las siguientes distribuciones, tomados de (Jiménez Moscoso, 2014).

#### 3.1.1. Distribución uniforme

Es una familia de distribuciones de probabilidad para variables aleatorias continuas, tales que para cada miembro de la familia, todos los intervalos de igual longitud en la distribución en su rango son igualmente probables. El dominio está definido por dos parámetros,  $a$  y  $b$ , que son sus valores mínimo y máximo. La distribución es escrita en forma abreviada como  $\mathcal{U}(a, b)$ .

La función de densidad de probabilidad de la distribución uniforme continua es:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{para } a \leq x \leq b \\ 0 & \text{para } x < a \text{ o } x > b. \end{cases}$$

Los valores en los dos extremos  $a$  y  $b$  no son por lo general importantes porque no afectan el valor de las integrales de  $f(x)dx$  sobre el intervalo, ni de expresiones

similares. La función de distribución de probabilidad es:

$$F(x) = \begin{cases} 0 & \text{para } x < a \\ \frac{x-a}{b-a} & \text{para } a \leq x < b \\ 1 & \text{para } x \geq b. \end{cases}$$

La función generadora de momentos es

$$M_x = E(e^{tx}) = \frac{e^{tb} - e^{ta}}{t(b-a)},$$

a partir de la cual se pueden calcular los momentos  $m_k$

$$m_1 = \frac{a+b}{2},$$

$$m_2 = \frac{a^2 + ab + b^2}{3},$$

y en general,

$$m_k = \frac{1}{k+1} \sum_{i=0}^k a^i b^{k-i}.$$

Para una variable aleatoria que satisface esta distribución, la esperanza matemática es entonces  $m_1 = \frac{a+b}{2}$  y la varianza es  $m_2 - m_1^2 = \frac{(b-a)^2}{12}$ .

### 3.1.2. Distribución normal

Si  $X$  sigue una distribución normal con parámetros  $\mu$  y  $\sigma^2$ , denotada por  $X \sim \mathcal{N}(\mu, \sigma^2)$ , la función de distribución de probabilidad esta dada por

$$\varphi(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2 \right\}, \quad x \in \mathbb{R}, \sigma > 0.$$

La función de densidad acumulada de una variable normal viene dada por

$$\Phi(x, \mu, \sigma) = \int_{-\infty}^x \varphi(u, \mu, \sigma) du.$$

En este caso, la función generadora de momentos es

$$M_X(t) = \exp \left( \mu t + \frac{1}{2} t^2 \sigma^2 \right),$$

y los momentos centrales de orden  $n$  se define como

$$\mu_n = \begin{cases} 0 & \text{si } n \text{ es impar} \\ \frac{n! \sigma^n}{2^{\frac{n}{2}} \left(\frac{n}{2}\right)!} & \text{si } n \text{ es par.} \end{cases}$$

Haciendo el cambio de variable

$$z = \frac{x-\mu}{\sigma},$$

se dice que  $Z$  tiene una distribución normal estándar, es decir,  $Z \sim \mathcal{N}(0, 1)$ .

### 3.1.3. Distribución *t*-student

Supongamos que  $X_1, \dots, X_n$  son variables aleatorias independientes distribuidas normalmente, con media  $\mu$  y varianza  $\sigma^2$ , denotado por  $t_\nu(\mu, \sigma^2)$ . Sea

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$$

la media muestral. Entonces

$$Z = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

sigue una distribución normal de media 0 y varianza 1.

Sin embargo, dado que la desviación estándar no siempre es conocida de antemano, se estudia un cociente relacionado,

$$T = \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}},$$

$$S^2(x) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

La función de densidad de  $T$  es

$$f(t) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\Gamma(\nu/2)} (1+t^2/\nu)^{-(\nu+1)/2},$$

donde  $\nu$  es igual a  $n-1$  y  $\Gamma$  es la función gamma.

La distribución de  $T$  se llama distribución *t*-student, donde  $\nu$  representa el número de grados de libertad. La distribución depende de  $\nu$ , pero no de  $\mu$  o  $\sigma$ , lo cual es muy importante en la práctica.

### 3.1.4. Distribución chi-cuadrado

La distribución de Pearson, o chi cuadrado, ( $\chi^2$ ), es una distribución de probabilidad continua con un parámetro  $k$  que representa los grados de libertad de la variable aleatoria

$$X = Z_1^2 + \dots + Z_k^2,$$

donde  $Z_i$  son variables aleatorias normales independientes de media cero y varianza uno. Si variable aleatoria  $X$  posee dicha distribución, se representa habitualmente como  $X \sim \chi_k^2$ .

Su función de densidad es:

$$f(x; k) = \begin{cases} \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k/2)-1} e^{-x/2} & \text{para } x > 0 \\ 0 & \text{para } x \leq 0, \end{cases}$$

donde  $\Gamma$  es la función gamma.

Su función de distribución es

$$F_k(x) = \frac{\gamma(k/2, x/2)}{\Gamma(k/2)},$$

donde  $\gamma(k, z)$  es la función gamma incompleta.

El valor esperado y la varianza de una variable aleatoria  $X$  con distribución  $\chi^2$  son, respectivamente,  $k$  y  $2k$ .

### 3.1.5. Distribución de error generalizada

Si  $X$  sigue una distribución de error generalizada con parámetros  $\mu$ ,  $\sigma$  y  $\kappa$ , denotada por  $X \sim \text{GED}_\kappa(\mu, \sigma)$ , entonces la función de distribución de probabilidad viene dada por

$$f_X(x; \mu, \sigma, \kappa) = \frac{\beta}{2\sigma\Gamma(\kappa + 1)} \exp \left\{ - \left( \beta \frac{|x - \mu|}{\sigma} \right)^{\frac{1}{\kappa}} \right\}, \quad x \in \mathbb{R},$$

donde  $0 \leq \kappa \leq 1$  es un parámetro que controla el grosor de la cola y

$$\beta = \sqrt{\frac{\Gamma(3\kappa)}{\Gamma(\kappa)}}.$$

Si  $\mu = 0$  y  $\sigma = 1$ , entonces  $X$  se dice que tiene una distribución de error generalizada estándar, es decir,  $X \sim \text{GED}_\kappa(0, 1)$ .

La función de densidad acumulada de una variable GED es dado por

$$F_X(x; \mu, \sigma, \kappa) = \frac{1}{2} \left[ 1 + \text{sign}(x - \mu) \gamma \left( \left( \beta \frac{|x - \mu|}{\sigma} \right)^{\frac{1}{\kappa}}, \kappa \right) \right],$$

donde  $\gamma(z, \kappa)$ ,  $\kappa > 0$  denota la función gamma incompleta y viene dada por:

$$\gamma(z, \kappa) = \frac{1}{\Gamma(\kappa)} \int_0^z t^{\kappa-1} e^{-t} dt.$$

Los momentos centrales de orden  $n$  están dados por

$$\mu_n = \begin{cases} 0 & \text{si } n \text{ es impar} \\ \frac{\Gamma((n+1)\kappa)}{\Gamma(\kappa)} \left[ \frac{\sigma}{\beta} \right]^n & \text{si } n \text{ es par,} \end{cases}$$

donde

$$\frac{\Gamma((n+1)\kappa)}{\Gamma(\kappa)} = (2\pi)^{-\frac{n}{2}} (n+1)^{(n+1)\kappa-1/2} \prod_{j=1}^n \Gamma \left( \kappa + \frac{j}{n+1} \right).$$

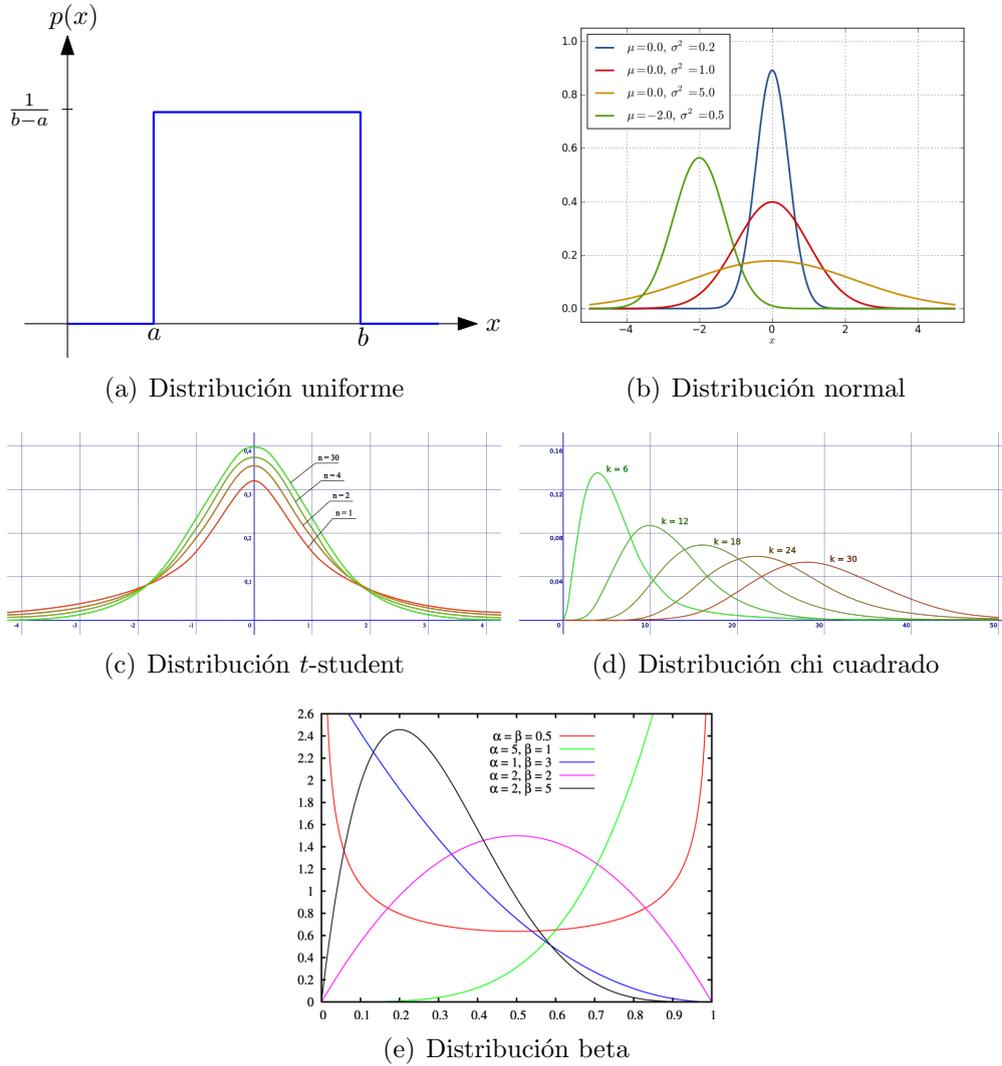


Figura 3.1: Distribuciones de probabilidad clásica.

### 3.1.6. Distribución beta

La distribución beta es una distribución de probabilidad continua sujeta a dos parámetros  $a$  y  $b$ , denotada por  $X \sim \mathcal{B}(a, b)$ , cuya función de densidad para valores  $0 \leq x \leq 1$  es

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1},$$

donde  $\Gamma$  es la función gamma.

El valor esperado y la varianza de una variable aleatoria  $X$  con distribución beta son

$$E(X) = \frac{a}{a+b},$$

$$\text{Var}(X) = \frac{ab}{(a+b+1)(a+b)^2}.$$

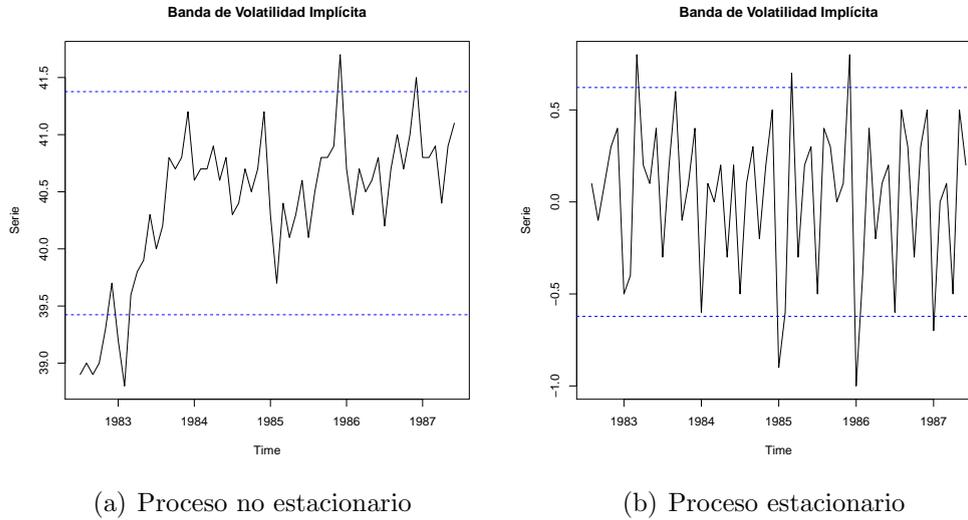


Figura 3.2: Ejemplo de series temporales.

## 3.2. Serie estacionaria

Un proceso estocástico es una sucesión de variables aleatorias  $X_t$  ordenadas en el tiempo  $t$ . Las series temporales se definen como un caso particular de los procesos estocásticos. De todos los tipos de procesos estocásticos posibles, son de especial interés los siguientes:

- *Proceso estocástico estacionario*: Un proceso estocástico se dice que es estacionario si su media y su varianza son constantes en el tiempo y si el valor de la covarianza entre dos periodos depende solamente de la distancia o rezago entre estos dos periodos de tiempo y no del tiempo en el cual se ha calculado la covarianza, como lo observado en la Figura 3.2(b).
- *Proceso gaussiano de ruido blanco*: Es una sucesión de variables aleatorias independientes e idénticamente distribuidas a lo largo del tiempo, con media igual a cero y varianza constante  $\sigma^2$ .
- *Caminata aleatoria*: Es un proceso estocástico no estacionario representado por

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

donde  $\epsilon_t \sim \mathcal{N}(0, 1)$  es el termino de error.

Para determinar si una serie  $X_t$  sigue un proceso estacionario, es decir, no sigue una caminata aleatoria, se emplea el modelo

$$X_t = \phi_1 X_{t-1} + \epsilon_t \tag{3.1}$$

y se usa la prueba de Dickey - Fuller, el cual contrasta

$$\begin{aligned} \text{Ho: } & \phi_1 = 1, \\ \text{Ha: } & \phi_1 < 1. \end{aligned}$$

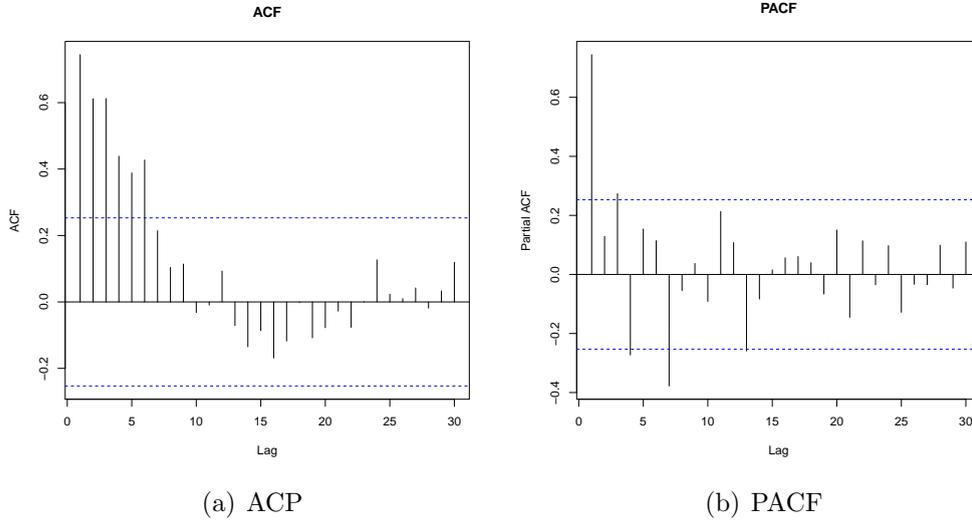


Figura 3.3: Ejemplo de autocorrelaciones en series temporales.

Una estadístico de prueba conveniente es la relación  $t$  de la estimación DF de  $\phi_1$  bajo la hipótesis nula. Para la ecuación (3.1), el método  $LS$  da

$$\hat{\phi}_1 = \frac{\sum_{t=1}^T X_{t-1} X_t}{\sum_{t=1}^T X_{t-1}^2},$$

$$\hat{\sigma}_\epsilon^2 = \frac{\sum_{t=1}^T (X_t - \hat{\phi}_1 X_{t-1})^2}{T - 1},$$

donde  $X_0 = 0$  y  $T$  es el tamaño de la muestra. La relación  $t$  es:

$$\text{DF} \equiv t - \text{ratio} = \frac{\hat{\phi}_1 - 1}{\text{std}(\hat{\phi}_1)} = \frac{\sum_{t=1}^T X_{t-1} \epsilon_t}{\hat{\sigma}_\epsilon \sqrt{\sum_{t=1}^T X_{t-1}^2}}.$$

### 3.3. Autocorrelaciones en series de tiempo

En una serie de tiempo puede presentar que los valores que toma una variable en el tiempo no son independientes entre si, sino que un valor determinado depende de los valores anteriores, existen dos formas de medir esta dependencia de las variables: la función de autocorrelación simple (ACF) y la función de autocorrelación parcial (PACF).

La autocorrelación simple mide la correlación entre dos variables separadas por  $k$  periodos, definidas por la ecuación

$$\rho_j = \text{corr}(X_j, X_{j-k}) = \frac{\text{cov}(X_j, X_{j-k})}{\sqrt{\text{Var}(X_j)} \sqrt{\text{Var}(X_{j-k})}}.$$

La función de autocorrelación simple tiene las siguientes propiedades:

$$\begin{aligned}\rho_0 &= 1, \\ -1 &< \rho_j < 1, \\ \rho_j &= \rho_{-j}.\end{aligned}$$

Por otro lado, la autocorrelación parcial mide la correlación entre dos variables separadas por  $k$  periodos cuando no se considera la dependencia creada por los retardos intermedios existentes entre ambas, definida por

$$\begin{aligned}\pi_j &= \text{corr}(X_j, X_{j-k} | X_{j-1} X_{j-2} \cdots X_{j-k-1}) \\ &= \frac{\text{cov}(X_j - \widehat{X}_j, X_{j-k} - \widehat{X}_{j-k})}{\sqrt{\text{Var}(X_j - \widehat{X}_j)} \sqrt{\text{Var}(X_{j-k} - \widehat{X}_{j-k})}}.\end{aligned}$$

La prueba Ljung-Box prueba en forma conjunta de que todos los coeficientes de autocorrelación son simultáneamente iguales a cero, esto es, son independientes. Para ello, se contrasta el siguiente planteamiento de hipótesis:

$$\begin{aligned}\text{Ho} &: \rho_1 = \rho_2 = \cdots = \rho_k = 0, \\ \text{Ha} &: \rho_i \neq 0 \text{ para algún } i = 1, 2, \cdots, k\end{aligned}$$

donde  $k$  es el número de rezagos a probar, y estadístico igual a

$$Q(k) = T(T+2) \sum_{i=1}^k \left( \frac{\widehat{\rho}_i^2}{n-i} \right) \sim \chi_k^2,$$

con  $T$  el tamaño de la muestra.

### 3.4. Modelos de volatilidad

Dado el comportamiento de las series financieras, una de las principales preocupaciones de los inversionistas y de los llamados agentes que participan en el mercado financiero es tener un conocimiento aproximado del riesgo que presentan las inversiones que realizan. A partir de los precios diarios  $X_t$  se puede calcular y modelar la rentabilidad como una variable aleatoria a la volatilidad de sus rentabilidades. Para medir el riesgo de mercado, se investiga el comportamiento de todas esas variables que influyen en la determinación de su rendimiento, o se considera la serie de retornos de los precios de mercado, que para valores pequeños de la rentabilidad precio, resulta ser una buena aproximación de la rentabilidad real, y permite la suma de las rentabilidades. Dichos retornos define la tasa de variación del activo, el cual se calcula dependiendo del tiempo de sus cotizaciones.

Si los activos se cotizan para un cierto tiempo  $t$  de forma continua, los retornos son de la forma

$$R_t = \frac{X_t - X_{t-1}}{X_{t-1}},$$

en cambio, si se cotizan para un tiempo  $t$  de forma discreta, entonces

$$R_t = \text{Ln} \left( \frac{X_t}{X_{t-1}} \right),$$

donde, para ambos casos,  $X_{t-1}$  es el precio del activo en el tiempo  $t - 1$  y  $X_t$  es el precio en el tiempo  $t$ .

Note que las dos formas anteriores son aproximadamente las mismas, ya que:

$$\text{Ln} \left( \frac{X_t}{X_{t-1}} \right) = \text{Ln} \left( 1 + \frac{X_t - X_{t-1}}{X_{t-1}} \right) \approx \frac{X_t - X_{t-1}}{X_{t-1}}.$$

Sea  $R_t$  el retorno de un activo en un tiempo  $t$  que se comporta, generalmente, como una distribución leptocúrtica, asimétrica y un proceso estacionario. La idea central detrás del estudio de la volatilidad es probar la heterocedasticidad condicional de la serie  $R_t^2$ .

Según (Tsay, 2014), para realizar estos modelos, es importante considerar la esperanza y varianza condicional de  $R_t$  dado  $F_{t-1}$ , son:

$$\begin{aligned} \mu_t &= \text{E}(R_t|F_{t-1}), \\ \sigma_t^2 &= \text{Var}(R_t|F_{t-1}) = \text{E}[(R_t - \mu_t)^2|F_{t-1}], \end{aligned}$$

donde  $F_{t-1}$  es el conjunto de información disponible en el tiempo  $t - 1$  que, por lo general, expresa los retornos pasados.

Se asume que  $\mu_t$  y  $R_t$  siguen un modelo ARMA( $p, q$ ), donde

$$\begin{aligned} R_t &= \mu_t + a_t, \quad a_t \sim \text{IID}(0, 1) \\ \mu_t &= \phi_0 + \sum_{i=1}^p \phi_i R_{t-i} + \sum_{j=1}^q \theta_j a_{t-j}, \end{aligned} \tag{3.2}$$

con  $|\phi_i| < 1$  y  $|\theta_j| < 1$ .

De la ecuación anterior se puede deducir que

$$\sigma_t^2 = \text{Var}(R_t|F_{t-1}) = \text{Var}(a_t|F_{t-1}),$$

donde  $\sigma_t$  es la volatilidad. Los modelos condicionales heteroscedásticos dados en la siguiente sección se ocupan de la evolución de  $\sigma_t^2$ . La forma bajo que  $\sigma_t^2$  evoluciona con el tiempo distingue un modelo de volatilidad de otro. Los modelos heteroscedásticos condicionales se pueden clasificar en dos categorías generales. Los de la primera categoría utilizan una función exacta para gobernar la evolución de  $\sigma_t^2$ , mientras que los de la segunda categoría usan una ecuación estocástica para describir  $\sigma_t^2$ .

El modelo GARCH pertenece a la primera categoría, mientras que el modelo SV se encuentra en la segunda categoría. La expresión  $a_t$  hace referencia a los shocks o las innovaciones de un retorno en el tiempo  $t$ . El modelo de  $\mu_t$  en la ecuación (3.2)

se denomina la ecuación media para  $R_t$  y el modelo  $\sigma_t^2$  la ecuación de volatilidad para  $R_t$ .

Para la construcción de modelos  $\sigma_t^2$  se verifica la existencia de correlaciones significativa en la serie  $R_t^2$  mediante la prueba Ljung-Box o si los residuales para la media,  $a_t = \mu_t - R_t$ , tienen efectos ARCH estadísticamente significativos. Para ello, se realiza o bien una prueba Ljung-Box con estadístico  $Q(k)$  a la serie  $a_t^2$  o una prueba multiplicadora de Lagrange. Esta prueba es equivalente a usar una distribución de Fisher para probar que  $\alpha_i = 0$  para todo  $i = 1, \dots, k$  en la regresión lineal

$$a_t^2 = \alpha_0 + \sum_{i=1}^k \alpha_i a_{t-i}^2 + \epsilon_t, \quad t = k+1, \dots, T$$

donde  $\epsilon_t$  es el término de error,  $k$  es un entero positivo pre-especificado y  $T$  es el tamaño de la muestra. Si  $SSR_0 = \sum_{t=k+1}^T (a_t^2 - \bar{w})^2$ , con  $\bar{w} = \left(\frac{1}{T}\right) \sum_{t=1}^T a_t^2$  y,  $SSR_1 = \sum_{t=k+1}^T \hat{\epsilon}_t^2$ , donde  $\hat{\epsilon}_t$  es el residuo de mínimos cuadrados de la regresión lineal anterior, entonces  $F$  es dado por:

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

con distribución  $F$  para  $k$  y  $T - 2k - 1$  grados de libertad.

El propósito de la prueba es determinar si existen efectos ARCH en la serie de retornos, esto es,  $\alpha_i \neq 0$  para algún  $i = 1, \dots, k$ .

Luego de proponer un modelo en media condicional  $\mu_t$  y un modelo en varianza condicional o estocástica  $\sigma_t^2$ , se verifica los supuestos en sus residuales estandarizados:

$$\tilde{a}_t = \frac{a_t}{\sigma_t}.$$

Las pruebas para comprobar los supuestos son las siguientes:

- Los estadísticos de Ljung-Box de  $\tilde{a}_t$  para verificar la no existencia de autocorrelaciones significativas y validar la ecuación en media.
- Los estadísticos de Ljung-Box de  $(\tilde{a}_t)^2$  para probar la validez de la ecuación de la volatilidad condicional o estocástica.
- La adecuada distribución de  $\tilde{a}_t$ , según sea el caso, planteada en los términos de error  $\epsilon_t$  de los modelo de volatilidad condicional o estocástica.

Lo siguiente será describir varios modelos de volatilidad condicional y modelos de volatilidad estocástica, tomados de (Tsay, 2014) y (Kastner, 2016a), respectivamente.

### 3.4.1. Modelos de volatilidad condicional

#### Modelo ARCH

Los modelos de heteroscedasticidad condicional autorregresivo, ver (Engle, 1982), conocidos como modelos ARCH, recoge los episodios de agrupamiento temporal de volatilidad que suele observarse en las series de rentabilidad de casi todo mercado financiero.

La idea básica para los modelos ARCH, en su aplicación a las series temporales financieras, es que un shock en el retorno de un activo financiero no está autocorrelacionado, pero sí que es dependiente.

La varianza de un proceso puede ser descrita a través del modelo ARCH( $m$ ) dado por:

$$\begin{aligned}R_t &= \mu_t + a_t, \\a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{IID}(0, 1) \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2,\end{aligned}$$

donde  $\epsilon_t$  son variables aleatorias independientes e idénticamente distribuidas con media cero y varianza uno,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$  para todo  $i > 0$ .

Para cumplir la condición de estacionariedad en varianza, la suma de todos los parámetros debe ser menor que la unidad, esto es,

$$\sum_{i=1}^m \alpha_i < 1,$$

el cual debe tomarse como un contraste de validez del modelo, ya que este no sería aceptable si este estima varianzas negativas.

Se podría suponer que  $\epsilon_t$  sigue o bien una distribución normal, una distribución  $t$ -student estandarizada o una distribución de errores generalizada (GED) y en algunos casos, se utiliza una distribución sesgada.

La varianza no condicional de dicho proceso es dada por la siguiente ecuación:

$$\text{Var}(R_t) = \frac{\alpha_0}{1 - \sum_{i=1}^m \alpha_i}.$$

De este modo, los valores de  $R_t$  no son independientes, a pesar de no estar correlacionados, dada la relación existente entre sus segundos momentos.

A pesar de su utilidad, estos modelos ARCH también tienen una serie de limitaciones que dejan su validez en entredicho. Al emplear una función cuadrática para calcular la varianza, se asume que los shocks positivos afectan a la volatilidad de precios y cotizaciones que los shocks negativos, el cual impide estimar apropiadamente el efecto apalancamiento, pues los shocks negativos afectan más intensamente a la volatilidad

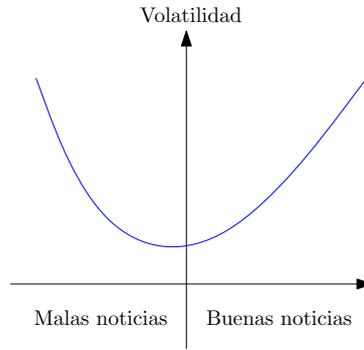


Figura 3.4: Efecto de apalancamiento.

de los activos financieros que los shocks positivos como lo observado en la Figura 3.4.

De esa forma, los modelos ARCH no son capaces de capturar este efecto, ya que asumen que la intensidad es la misma tanto para shocks positivos como negativos.

Tras analizar esta limitación que tiene este modelo a la hora de ser aplicado, se toma como punto de partida la familia de modelos GARCH( $m, s$ ).

### Modelo GARCH

En muchos casos, la especificación ARCH que recoge la estructura de autocorrelación en varianza necesita de un elevado número de retardos. Para evitar que el alto número de coeficientes en términos autoregresivos, generalmente bastante relacionados, produzca una importante pérdida de precisión en su estimación, se ha propuesto una parametrización alternativa, restringida, dependiente de un número reducido de parámetros, ver (Bollerslev, 1986).

El modelo de heteroscedasticidad condicional autoregresivo generalizado, conocido como GARCH( $m, s$ ), es dado por:

$$\begin{aligned} R_t &= \mu_t + a_t, \\ a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{IID}(0, 1) \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-i}^2, \end{aligned}$$

donde  $\epsilon_t$  es una sucesión de variables idénticamente distribuido con media cero y varianza uno,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$  y  $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$  para garantizar la condición de estacionariedad en media. En este caso,  $\alpha_i = 0$  si  $i > m$  y  $\beta_j = 0$  si  $j > s$ . La última restricción en  $\alpha_i + \beta_j$  implica que la varianza condicional de  $a_t$  es finita, mientras que la varianza condicional  $\sigma_t^2$  evoluciona con el tiempo.

Como en el modelo anterior,  $\epsilon_t$  pueden seguir o una distribución normal, una distribución  $t$ -student estandarizada o una distribución GED. Por otro lado, la ecuación anterior se reduce a un modelo ARCH( $m$ ) si  $s = 0$ .

La esperanza y varianza no condicional de dicho proceso es:

$$\begin{aligned} E(R_t) &= 0, \\ \text{Var}(R_t) &= \frac{\alpha_0}{1 - \sum_{i=1}^m \alpha_i}. \end{aligned}$$

### Modelo IGARCH

Las siglas IGARCH hacen referencia a Integrated Generalized Autoregressive Conditional Heteroskedastic o modelo generalizado autorregresivo integrado de heterocedasticidad condicionada.

El propósito de esta variante de modelo GARCH( $m, s$ ) es estimar de una forma conveniente el efecto de apalancamiento. Para lograrlo, el modelo supone que la varianza incondicional es finita:

$$\sum_{i=1}^m \alpha_i + \sum_{j=1}^s \beta_j < 1.$$

Por tanto, al suponer que  $\beta_j = (1 - \alpha_j)$ , el modelo IGARCH es dado por:

$$\begin{aligned} R_t &= \mu_t + a_t, \\ a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{IID}(0, 1) \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 + \sum_{j=1}^s (1 - \alpha_j) \sigma_{t-j}^2. \end{aligned}$$

Este nuevo modelo hace que el efecto de la varianza condicional no desaparezca nunca, a diferencia del modelos GARCH( $m, s$ ), transcurriendo múltiples periodos de tiempo hasta que la varianza alcance de nuevo el valor medio.

Este modelo ha recibido diversas críticas, el cual basa su análisis en tres puntos clave:

1. La dificultad para lograr que los parámetros no sean negativos.
2. La incapacidad de estimar el efecto de apalancamiento.
3. La dificultad que entraña implementar estos modelos IGARCH, dado lo confuso que es el término de persistencia en varianza.

Es así como surge el modelo EGARCH, cuya especificación es más general y acaba con los problemas a la hora de replicar apropiadamente la asimetría en los shocks.

### Modelo EGARCH

El modelo EGARCH, o modelo exponencial generalizado autorregresivo y condicionalmente heterocedástico, acaba con las asimetrías en la estimación del efecto de los shocks, al implementar una función  $g(\epsilon_t)$  de las innovaciones  $\epsilon_t$  y en las que

por tanto el valor de las innovaciones queda recogido por medio de la expresión  $|\epsilon_t| - E(|\epsilon_t|)$ . Matemáticamente:

$$g(\epsilon_t) = \theta\epsilon_t + \gamma[|\epsilon_t| - E(|\epsilon_t|)],$$

donde  $\theta$  y  $\gamma$  son constantes reales. Tanto  $\epsilon_t$  como  $|\epsilon_t| - E(|\epsilon_t|)$  son variables igual e idénticamente distribuidas de media cero y distribuciones continuas. Por lo tanto,  $E[g(\epsilon_t)] = 0$ . La asimetría de  $g(\epsilon_t)$  se puede ver reescribiéndolo como

$$g(\epsilon_t) = \begin{cases} (\theta + \gamma)\epsilon_t - \gamma E(|\epsilon_t|) & \text{si } \epsilon_t \geq 0 \\ (\theta - \gamma)\epsilon_t - \gamma E(|\epsilon_t|) & \text{si } \epsilon_t < 0. \end{cases}$$

Un modelo EGARCH( $m, s$ ) puede escribirse como

$$\begin{aligned} R_t &= \mu_t + a_t, \\ a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{IID}(0, 1) \\ \text{Ln}(\sigma_t^2) &= \alpha_0 + \frac{1 + \beta_1 B + \dots + \beta_{s-1} B^{s-1}}{1 - \alpha_1 B + \dots - \alpha_m B^m} g(\epsilon_{t-1}), \end{aligned}$$

donde  $\alpha_0$  es una constante,  $B$  es el operador retraso tal que  $Bg(\epsilon_t) = g(\epsilon_{t-1})$  y  $1 + \beta_1 B + \dots + \beta_{s-1} B^{s-1}$  y  $1 - \alpha_1 B + \dots - \alpha_m B^m$  son polinomios con ceros fuera del círculo unitario y no tienen factores comunes, esto es, que los valores absolutos de los ceros son mayores que 1. La ecuación anterior utiliza la parametrización ARMA para describir la evolución de la varianza condicional de  $a_t$ .

Sobre la base de esta representación, algunas propiedades del modelo EGARCH pueden obtenerse de manera similar a las del modelo GARCH. Por ejemplo, la media incondicional de  $\text{Ln}(\sigma_t^2)$  es  $\alpha_0$ . Sin embargo, el modelo difiere del modelo GARCH de varias maneras. En primer lugar, utiliza la varianza condicional registrada para relajar la restricción de positividad de los coeficientes del modelo. En segundo lugar, el uso de  $g(\epsilon_t)$  permite que el modelo responda asimétricamente a valores rezagados positivos y negativos de  $a_t$ .

Para facilitar la estimación, se puede utilizar una forma alternativa para el modelo EGARCH( $m, s$ ):

$$\text{Ln}(\sigma_t^2) = \alpha_0 + \sum_{i=1}^m \left( \alpha_i \frac{|a_{t-i}|}{\sigma_{t-i}} + \gamma_i \frac{a_{t-i}}{\sigma_{t-i}} \right) + \sum_{j=1}^s \beta_j \text{Ln}(\sigma_{t-j}^2),$$

donde el signo positivo de  $a_{t-i}$  contribuye a  $\alpha_i(1 + \gamma_i)|\epsilon_t|$  de la log volatilidad, mientras que el signo negativo de  $a_{t-i}$  contribuye a  $\alpha_i(1 - \gamma_i)|\epsilon_t|$ , donde  $\epsilon_t = \frac{a_{t-i}}{\sigma_{t-i}}$  y, el parámetro  $\gamma_i$  significa el efecto de apalancamiento de  $a_{t-i}$ .

## Modelo TGARCH

La idea del modelo TGARCH es dividir la distribución de los shocks en intervalos disjuntos, para luego aproximar una función lineal por tramos para la desviación estándar condicional o para la volatilidad condicional. Si sólo hay dos intervalos,

la división es normalmente en un umbral identificado con el número cero, donde la influencia de los shocks positivos se identifica con valores por encima de cero y negativos por debajo de este valor. En este caso, el modelo TGARCH( $m, s$ ) se puede escribir como:

$$\begin{aligned} R_t &= \mu_t + a_t, \\ a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{IID}(0, 1) \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^m (\alpha_i + \gamma_i N_{t-i} a_{t-i}^2) + \sum_{j=1}^s \beta_j \sigma_{t-i}^2, \end{aligned}$$

donde  $N_{t-i}$  es un indicador para  $a_{t-i}$  definido por:

$$N_{t-i} = \begin{cases} 1, & \text{si } a_{t-i} < 0 \\ 0, & \text{si } a_{t-i} \geq 0, \end{cases}$$

y  $\alpha_i, \gamma_i, \beta_j$  son parámetros no negativos que satisfacen condiciones similares a los modelos GARCH. A partir del modelo, un número positivo  $a_{t-i}$  contribuye  $\alpha_i a_{t-i}^2$  a  $\sigma_t^2$ , mientras que un número negativo  $a_{t-i}$  tiene un mayor impacto  $(\alpha_i + \gamma_i) a_{t-i}^2$  con  $\gamma_i > 0$ .

### Modelo APARCH

El modelo TGARCH pertenece a la clase de modelos de poder asimétrico autoregresivo condicional heteroscedástico, denominado APARCH( $m, s$ ), el cual se describe como

$$\begin{aligned} R_t &= \mu_t + a_t, \\ a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{IID}(0, 1) \\ \sigma_t^\delta &= \alpha_0 + \sum_{i=1}^m \alpha_i (|a_{t-i}| + \gamma_i a_{t-i})^\delta + \sum_{j=1}^s \beta_j \sigma_{t-i}^\delta, \end{aligned}$$

donde  $\delta > 0$  y los coeficientes  $\alpha_0, \alpha_i, \gamma_i$  y  $\beta_j$  satisfacen ciertas condiciones de regularidad para que la volatilidad sea positiva. Tres casos especiales de los modelos APARCH son de interés. Cuando  $\delta = 2$ , el modelo APARCH se reduce a un modelo TGARCH. Cuando  $\delta = 1$ , el modelo utiliza la volatilidad directamente en la ecuación de volatilidad. El caso de  $\delta = 0$  se toma  $\delta \rightarrow 0$  y en este caso el modelo se convierte en un modelo EGARCH.

### 3.4.2. Modelo de volatilidad estocástica

Los modelos de heteroscedasticidad condicional como los ARCH o GARCH asumen que la volatilidad puede ser observada un paso adelante con la información que proporciona las variables aleatorias  $\epsilon_t$  o  $a_t$ . En estos modelos, las noticias o choques se incorporan únicamente a través de las innovaciones  $\epsilon_t$ , es decir que sólo existen perturbaciones sobre la ecuación de la media. De esta forma, la volatilidad se estima a través de la información pasada, tornándose predecible porque para cada momento  $t$  ya se conoce la información en  $t - 1$ .

Partiendo de esto, un modelo más realista para la volatilidad se puede basar en una

ecuación de comportamiento que tiene un componente predecible que depende de la información pasada, como en los GARCH, y un componente de ruido inesperado. En este caso, la volatilidad es una variable no observada latente. Una interpretación de la volatilidad latente es que representa la llegada de nueva información en el mercado. Este tipo de modelación corresponde a los denominados modelos de volatilidad estocástica: SV-AR(1) o SVt-AR(1).

### Modelo SV-AR(1)

El modelo de volatilidad estocástica autorregresivo o SV-AR(1) es un modelo de estado-espacio donde la variable de estado es la log-volatilidad, dada por

$$R_t = \sigma_t \epsilon_t,$$

$$\sigma_t^2 = \exp \left\{ \frac{h_t}{2} \right\},$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t,$$

donde  $\mu \in \mathbb{R}$ ,  $h_t$  es el logaritmo de la volatilidad que sigue un proceso estacionario AR(1) con parámetro de persistencia  $|\phi| < 1$ .

Se asume que  $\epsilon_t \sim \mathcal{N}(0, 1)$  y  $\eta_s \sim \mathcal{N}(0, \sigma_\eta^2)$  son las perturbaciones o ruidos no esperados para los retornos y para la volatilidad, respectivamente, mutuamente independiente para todo  $t$  y  $s$ .

Para completar la configuración del modelo, una distribución a priori para el vector de parámetros  $\theta = (\mu, \phi, \sigma_\eta)$  debe especificarse con el fin de tomar la mediana, como estimador robusto, de cada distribución como parámetro puntual del modelo. Para ello, se eligen componentes independientes para cada parámetro, es decir,  $p(\theta) = p(\mu)p(\phi)p(\sigma_\eta)$ .

El parámetro  $\mu$  sigue una distribución normal a priori  $\mathcal{N}(b, B)$ . Según (Kastner, 2016a), se puede tomar  $b = 0$  y  $B \geq 100$ , pero dada varia experiencia de diversos autores con datos empíricos, su elección exacta no suele ser muy importante.

Para el parámetro  $\phi \in (-1, 1)$ , se elige  $(\phi + 1) \sim 2\mathcal{B}(a_0, b_0)$ , lo que implica que esta distribución se encuentra en el intervalo  $(-1, 1)$  y garantiza la estacionariedad del proceso de volatilidad autorregresiva. Su valor esperado y varianza se dan a través de las expresiones

$$E(\phi) = \frac{2a_0}{a_0 + b_0} - 1,$$

$$\text{Var}(\phi) = \frac{4a_0b_0}{(a_0 + b_0)^2(a_0 + b_0 + 1)}.$$

La distribución a priori de  $\phi$  depende sólo de la relación entre  $a_0$  y  $b_0$ . Este parámetro es mayor que cero si y sólo si  $a_0 > b_0$  y menor que cero si y sólo si  $a_0 < b_0$ . La distribución uniforme en  $(-1, 1)$  surge como un caso particular cuando  $a_0 = b_0 = 1$ .

Para un conjunto de datos financieros con no demasiadas observaciones, esto es  $T \leq 1000$ , la elección de  $a_0$  y  $b_0$  puede ser bastante influenciada sobre la forma de la distribución a priori de  $\phi$ . De hecho, cuando el proceso de generación de datos subyacente es (casi) homocedástico, la volatilidad de la log-varianza  $\sigma_\eta$  es (muy cerca de) cero y por lo tanto la probabilidad contiene poca o ninguna información sobre  $\phi$ . En consecuencia, la distribución a priori de  $\phi$  es (casi) igual a su anterior, no importa cuántos puntos de datos se observan. Según (Kastner, 2016a), se puede tomar  $a_0 = 5$  y  $b_0 = 1.5$ , lo que implica una media anterior de 0.86 con una desviación estándar anterior de 0.11 y, por tanto, muy poca masa para valores no positivos de  $\phi$ .

Para la volatilidad de la log varianza  $\sigma_\eta \in \mathbb{R}^+$ , se elige  $\sigma_\eta^2 \sim B \times \chi_1^2$ , donde la elección del hiperparámetro  $B$  resulta ser de menor importancia en aplicaciones empíricas, siempre y cuando no se fije demasiado pequeño. En algunas investigaciones, se puede tomar  $B = 1$ .

### Modelo SVt-AR(1)

El modelo SV-AR(1) con  $\epsilon_t$  siguiendo una distribución  $t$ -student, denominado SVt-AR(1), se da a través de

$$\begin{aligned} R_t &= \sigma_t \epsilon_t, \\ \sigma_t^2 &= \exp \left\{ \frac{h_t}{2} \right\}, \\ h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \end{aligned}$$

donde  $\epsilon_t \sim t_\nu(0, 1)$  y  $\eta_s \sim \mathcal{N}(0, \sigma_\eta^2)$ .

Se asume que la distribución a priori del parámetro de grados de libertad  $\nu$  sigue una distribución uniforme sobre el intervalo  $(a, b)$ , esto es,  $\nu \sim \mathcal{U}(a, b)$  y las distribuciones a priori de los demás parámetros son como los dados en el modelo SV-AR(1). Según (Kastner, 2015), se puede tomar  $\nu \sim \mathcal{U}(2, 100)$ .

# Capítulo 4

## Marco Metodológico

La serie temporal a utilizar son los retornos diarios de las acciones CEMARGOS de la empresa colombiana de Cementos Argos S.A. Para el cálculo de esta serie de retornos, se utiliza los precios de cierre diarios de acciones en los que existe mercado en el periodo comprendido entre 08/julio/2012 y el 17/marzo/2017, con un total de 1165 datos, obtenidos por la bolsa de valores de Colombia, ver (de Colombia, 2011).

Medida	Estadístico	
	Precio de cierre	Retornos
Media	9861.631	0.0007299
Mediana	9670	0
Varianza	1878935	0.0002966789
Des. estándar	1370.742	0.01722437
IQR	2170	0.01840002
Asimétrica	-0.1736806	-0.1922218
Curtosis	-0.700694	2.148553

Tabla 4.1: Descriptivos acciones CEMARGOS.

Como lo contrastado en la Figura 4.1, se deduce que entre los años 2012, 2013 y 2014 hubo un incremento significativo en los precios de sus acciones y una fuerte caída en lo transcurrido del año 2015, conllevando a obtener precios por debajo de sus rangos intercuartílicos. En el año 2016 y lo transcurrido del año 2017, se ha aumentado dichos precios pero con caídas aun significativas. Por otro lado, en la Tabla 4.1 muestra que los datos de la serie no siguen una distribución normal, estando ligeramente sesgada a la izquierda y platocúrtica, con un coeficiente de asimetría de  $-0.1736806$  y curtosis de  $-0.700694 < 0$ .

En la modelación de series temporales es importante identificar si el fenómeno presenta un proceso estacionario a lo largo de tiempo  $1 \leq t \leq 1165$ . Para ello, aplicando la prueba de Dickey-Fuller, que contrasta

- Ho : La serie sigue un proceso no estacionario,
- Ha : La serie sigue un proceso estacionario,

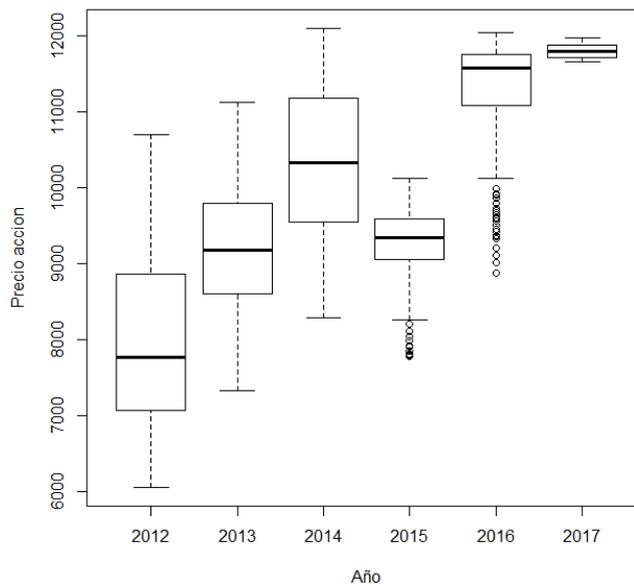


Figura 4.1: Boxplot por año precio CEMARGOS.

se concluye que, con un nivel de significancia del 5 %, la serie de los precios no se comporta estadísticamente como un proceso estacionario, con una probabilidad de 4.543 %. Por tanto, es de total interés analizar los retornos de la serie dada por la ecuación

$$R_t = \frac{X_t - X_{t-1}}{X_{t-1}},$$

donde  $X_t$  es el precio de cierre en el tiempo  $t$  y  $X_{t-1}$  en el tiempo  $t - 1$ .

En la Figura 4.2 se observa la diferencia entre la serie de los precios de cierre  $\{X_t\}_{t=1}^{T=1165}$  y la serie de retornos  $\{R_t\}_{t=1}^{T=1164}$ . Al usar la prueba de Dickey-Fuller, el cual contrasta,

- Ho : La serie  $R_t$  sigue un proceso no estacionario,  
 Ha : La serie  $R_t$  sigue un proceso estacionario,

se concluye que, con un nivel de significancia del 5 %, la serie se comporta estadísticamente como un proceso estacionario, con una probabilidad del 1 %.

En la Tabla 4.1 se observa que la serie  $R_t$  no sigue una distribución normal, puesto que se encuentra ligeramente sesgada a la izquierda y leptocúrtica, con un coeficiente de asimetría de -0.1922218 y curtosis 2.148553 > 0.

Para aplicar los modelos de volatilidad condicional se sugieren los siguientes pasos:

1. Aplicar una prueba de Ljung-Box para determinar si existen autocorrelaciones simples y autocorrelaciones parciales significativas en la serie de retornos  $R_t$

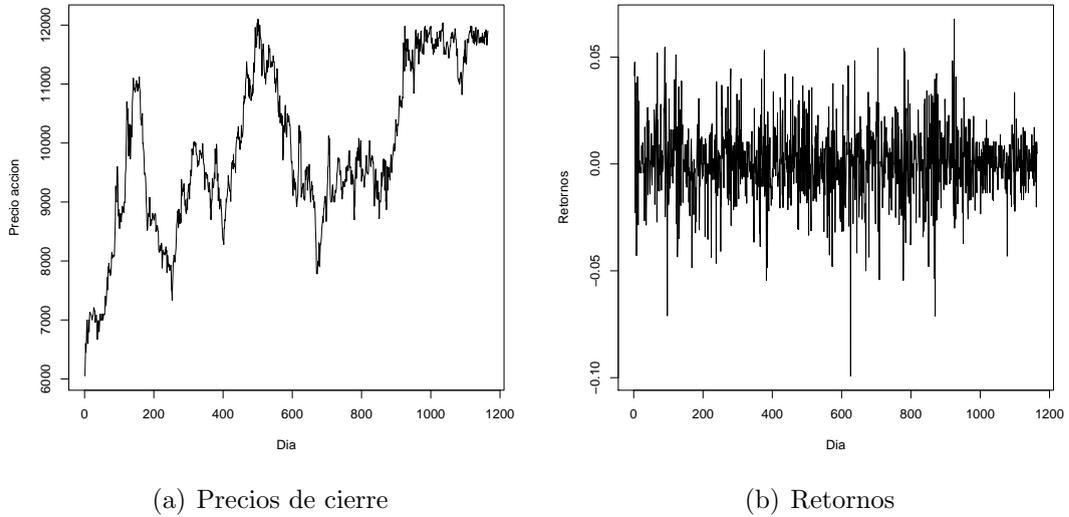


Figura 4.2: Serie de precios y retornos CEMARGOS.

y  $R_t^2$  para determinar si es necesario establecer un modelo  $ARMA(p, q)$  para media  $\mu_t$  y/o un modelo de volatilidad condicional o estocástica  $\sigma_t^2$ , respectivamente.

2. Si existen autocorrelaciones simples y parciales significativas en la serie  $R_t$  y  $R_t^2$ , se realiza las gráficas de función de autocorrelación (ACF) y la función de autocorrelación parcial (PACF) para cada una de las series mencionadas, con el fin de identificar los parámetros  $(p, q)$  y  $(m, s)$  que logren estimar el mejor modelo  $ARMA(p, q)$  junto con los modelos  $ARCH(m)$  y pertenecientes de la familia  $GARCH(m, s)$  para explicar la media  $\mu_t$  y volatilidad de la serie  $R_t$ .
3. Si existen autocorrelaciones simples y autocorrelaciones parciales significativas en la serie  $R_t^2$  mas no en la serie  $R_t$ , se encoje los valores de  $m$  y  $s$ , dependiendo de los primeros rezagos significativos proporcionados por las gráficas ACP y PACF de la serie  $R_t^2$ , para estimar los modelos  $ARCH(m)$  y pertenecientes de la familia  $GARCH(m, s)$  sin modelos para la media.
4. Por medio de los estadísticos descriptivos de la serie  $R_t$ , se realiza una primera inspección de aquellas distribuciones que podrían seguir los términos de errores  $\epsilon_t$  expresados por los modelos de volatilidad condicional.
5. Por medio del paquete *rugarch* del software estadístico R, se calcula los criterios de información bayesiana BIC para cada uno de los modelos de volatilidad condicional:
  - $ARMA(p, q)+ARCH(m)$ ,
  - $ARMA(p, q)+GARCH(m, s)$ ,
  - $ARMA(p, q)+IGARCH(m, s)$ ,
  - $ARMA(p, q)+EGARCH(m, s)$ ,

- $\text{ARMA}(p, q) + \text{TGARCH}(m, s)$ ,
- $\text{ARMA}(p, q) + \text{APARCH}(m, s)$ ,

con  $\epsilon_t \sim \text{IID}(0, 1)$  planteadas por el paso anterior.

6. Se verifica la validación de los parámetros de uno de cada tipo de modelo de volatilidad condicional que poseen menor BIC, esto es, el valor absoluto de cada uno de los parámetros puntuales de los modelo  $\text{ARMA}(p, q)$  sea menor que uno, supuestos de estacionariedad de los modelos ARCH y familia pertenecientes a modelos GARCH.
7. Se escoge el mejor modelo de volatilidad condicional, por medio del menor BIC y validación de los supuestos en sus residuales estandarizados, que logre explicar la volatilidad de la serie de retornos.

Por otro lado, para la construcción de los modelos  $\text{SV-AR}(1)$  y  $\text{SVt-AR}(1)$ , se usa el paquete *stochvol* siguiendo ditribuciones a priori, en cada uno de sus parámetros, establecidos por defecto y validar los supuestos en sus residuales.

Para finalizar este trabajo, se incorpora back-testing con el fin de determinar la efectividad de los modelos fuera de la muestra. Para ello,

1. Se calculan los primeros quince pronósticos de volatilidades para cada uno de los mejores modelos de volatilidad condicional y estocástica.
2. Con los primeros quince datos reales de precios de acciones CEMARGOS por fuera de la muestra, comprendidos entre 21/marzo/2017 al 10/abril/2017, se realiza una nueva gráfica de retornos, comprendidas entre 08/junio/2012 al 10/abril/2017, de tal forma que se realice el conteo de los datos que logren salirse de las bandas de volatilidad estimadas y pronosticadas por cada uno de los modelos de volatilidad, a un nivel de confianza del 95 %.
3. Si existe un modelo de tal forma que posea el menor numero de retornos por fuera de las bandas de volatilidad, dicho modelo será el mejor que pronostique los precios de los activos. En cambio, si existen dos o mas modelos que tienen el mismo numero de retornos por fuera de las bandas de volatilidad, se estiman los precios de los activos y compararlos con los precios reales por medio de estimación de errores MAE, RMSE, MAPE y AMAPE y tomar el mejor modelo de pronóstico que posea el menor numero de estas estimaciones.

# Capítulo 5

## Análisis y resultados

La serie de retornos  $R_t$  en el periodo evaluado cuenta con 1164 datos. En la Figura 4.2 se evidencia las fluctuaciones en la serie de precios y retornos.

El histograma de la Figura 5.1 muestra que la densidad incondicional de los retornos es simétrica con un centramiento cercano a cero y se caracteriza por una curtosis mayor a la presentada en la distribución normal. Al aplicar una prueba  $t$  para la serie de retornos, el cual contrasta

$$\begin{aligned} H_0 & : \text{La media es igual a cero,} \\ H_a & : \text{La media no es igual a cero,} \end{aligned}$$

evidentemente, con un nivel de significancia del 5 %, la media puede ser considerada estadísticamente como cero, con una probabilidad del 14.85 %.

En la Figura 5.2 se observa que la función de autocorrelación simple de la serie  $R_t$  poseen pocos rezagos significativos que se encuentran por fuera de las bandas de confianza al 95 %, sin embargo se aplica la prueba de Ljung-Box para comprobar la significancia de la autocorrelaciones hasta el rezago 20.

La probabilidad asociada al estadístico de prueba es del 7.259 % el cual permite concluir, con un nivel de significancia del 5 %, que los retornos no presentan rezagos significativos, por tanto no hay presencia de autocorrelación serial. Se observan autocorrelaciones significativas en los retornos al cuadrado, con un probabilidad menor que 1 %, indicando que no existe independencia serial en dicha serie.

A continuación se plantea los modelos de volatilidad condicional y estocásticas para la serie de retornos.

### 5.1. Modelo de volatilidad condicional

Con el fin de identificar un modelo apropiado, se inició con los modelos ARCH( $m$ ), GARCH( $m, s$ ), TGARCH( $m, s$ ), EGARCH( $m, s$ ), IGARCH( $m, s$ ) y APARCH( $m, s$ ), con  $\mu_t = 0$  conforme al resultado de la prueba de Ljung-Box para la serie  $R_t$ , para valores de  $m, s = 1, 2$  que son los primeros rezagos significativos observados en las

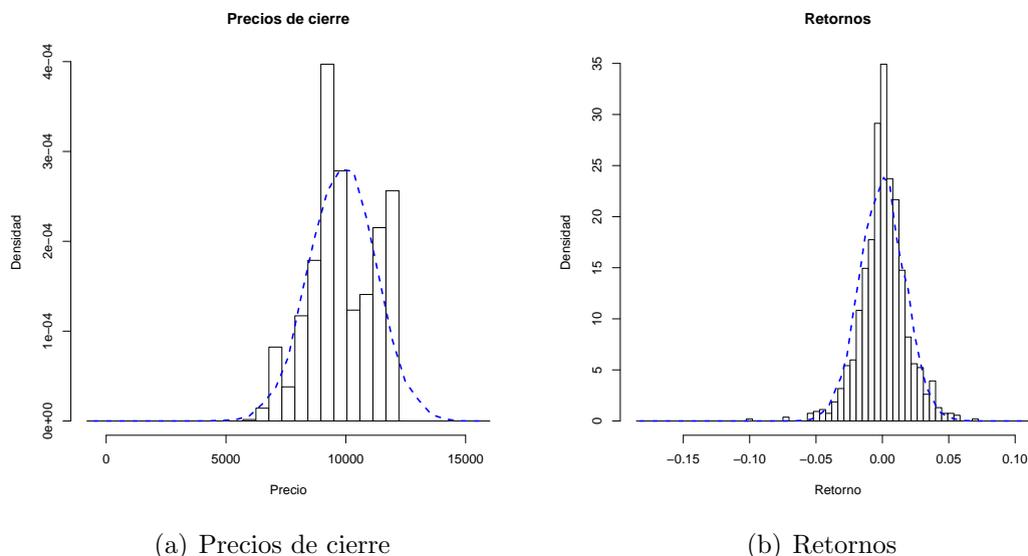


Figura 5.1: Histograma de acciones CEMARGOS.

gráficas ACP y PACP de la serie  $R_t^2$ . Como las series de retornos financieros poseen unas regularidades empíricas llamadas hechos estilizados, la cual se encuentra las colas pesadas de la distribución de los retornos donde esta distribución suele ser leptocúrtica, con colas mas pesadas que la distribución normal, se usa la distribución  $t$ -student estandarizada y la distribución GED.

En la Tabla 5.1 se muestran los valores dados por los criterios de información bayesiana BIC para cada uno de los modelos mencionados, respectivamente, por medio del paquete *rugarch* del programa estadístico *R version 3.3.3*, ver (Ghalanos, 2013).

Tomando los menores valores dados por el criterio de informacion bayesiano (BIC) para cada modelo, se selecciona el modelo ARCH(1) con  $\epsilon \sim t_{4.4867}^*$  y los modelos GARCH(1, 1), TGARCH(1, 1), EGARCH(1, 1), IGARCH(1, 1) y APARCH(1, 1) con  $\epsilon_t \sim \text{GED}_\kappa(0, 1)$ . En las Tablas 5.2 y 5.3 se muestran los parámetros estimados, incluyendo su nivel de significancia, y los estadísticos de prueba de Ljung-Box a los residuales estandarizados para cada uno de los modelos, de los cuales se concluye:

- El modelo ARCH(1) de la forma

$$\begin{aligned}
 R_t &= a_t, \\
 a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim t_{4.4867}^* \\
 \sigma_t^2 &= 2.16 \times 10^{-3} + 0.3858 a_{t-1}^2,
 \end{aligned} \tag{5.1}$$

posee estadísticamente, para un nivel de significancia del 5%, rezagos significativos en los residuales estandarizados al cuadrado como lo observado en la Tabla 5.3, por tanto no es un buen modelo que explique la volatilidad de la serie de retornos.

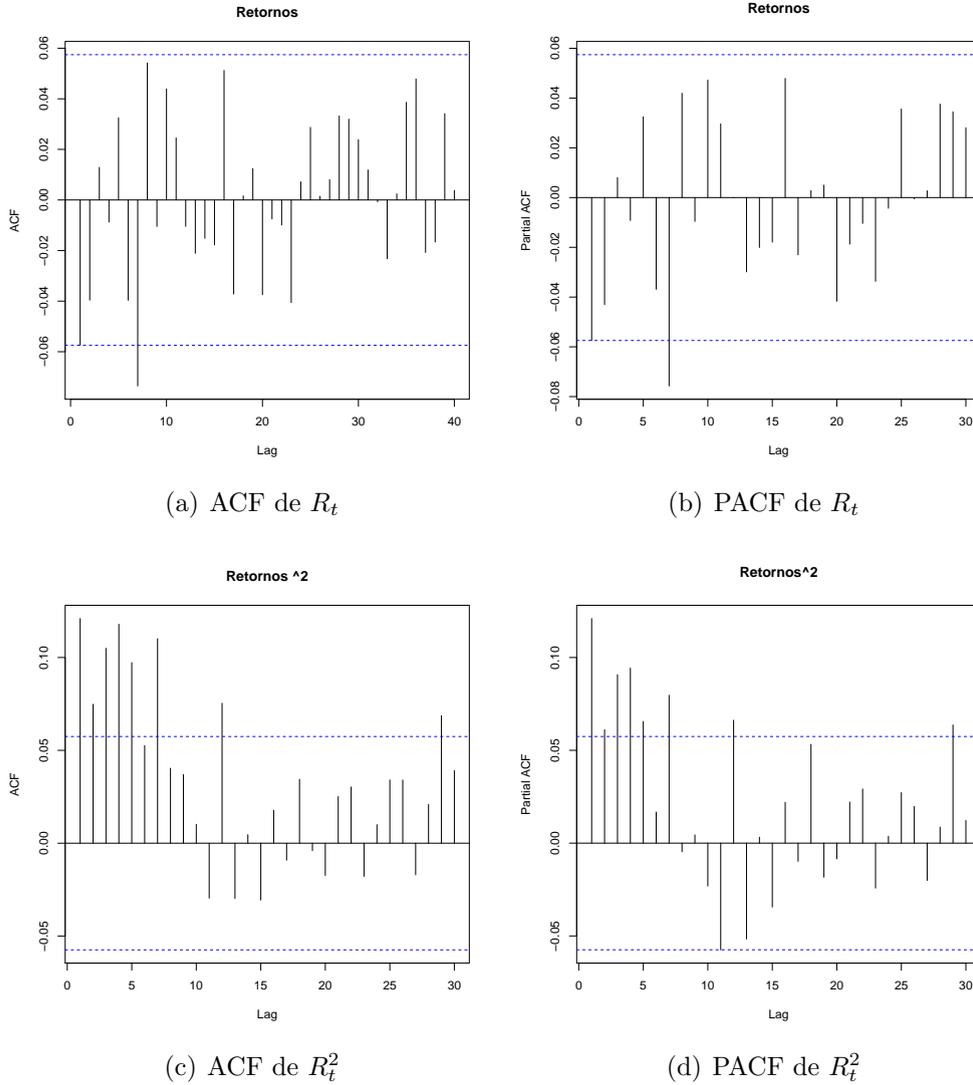


Figura 5.2: ACP y PACF retornos CEMARGOS.

- El modelo IGARCH(1,1) de la forma

$$\begin{aligned}
 R_t &= a_t, \\
 a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{GED}_{1.109}(0, 1) \\
 \sigma_t^2 &= 9 \times 10^{-7} + 0.161308 a_{t-1}^2 + 0.838692 \sigma_{t-1}^2,
 \end{aligned} \tag{5.2}$$

los parámetros estimados son significativos, como lo observado en la Tabla 5.2 para un nivel de significancia del 5%.

- El modelo EGARCH(1, 1) dado por

$$\begin{aligned}
 R_t &= a_t, \\
 a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{GED}_{1.176}(0, 1) \\
 \text{Ln}(\sigma_t^2) &= -0.84274 + 0.018381 \frac{|a_{t-1}|}{\sigma_{t-1}} + 0.28686 \frac{a_{t-1}}{\sigma_{t-1}} + \\
 &\quad + 0.89754 \text{Ln}(\sigma_{t-1}^2),
 \end{aligned} \tag{5.3}$$

	Modelo	BIC		
		$\epsilon_t \sim \mathcal{N}(0, 1)$	$\epsilon_t \sim t_{\nu}^*(0, 1)$	$\epsilon_t \sim \text{GED}_{\kappa}(0, 1)$
ARCH	ARCH(1)	-5.3190	-5.3889	5.4003
	ARCH(2)	-5.3106	-5.3815	3.7880
	ARCH(3)	-5.3086	-5.3807	5.0021
	ARCH(4)	-5.3028	-5.3752	-5.3865
	ARCH(5)	-5.3046	-5.3707	-5.3827
	ARCH(6)	-5.3104	-5.3681	-5.3777
	ARCH(7)	-5.3167	-5.3676	-5.3769
	ARCH(8)	-5.3111	-5.3619	-5.3709
	ARCH(9)	-5.3058	-5.3506	-5.3659
GARCH	GARCH(1,1)	-5.3522	-5.4056	-5.4167
	GARCH(1,2)	-5.3444	-5.3989	-5.4102
	GARCH(2,1)	-5.3438	-5.3975	-5.4089
	GARCH(2,2)	-5.3383	-5.3928	-5.4041
IGARCH	IGARCH(1,1)	-5.3438	-5.4063	-5.4159
	IGARCH(1,2)	-5.3370	-5.4000	-5.4102
	IGARCH(2,1)	-5.3358	-5.3981	-5.4084
	IGARCH(2,2)	-5.3309	-5.3939	-5.4041
EGARCH	EGARCH(1,1)	-5.3511	-5.4020	-5.4133
	EGARCH(1,2)	-5.3435	-5.3969	-5.4072
	EGARCH(2,1)	-5.3408	-5.3910	-5.4027
	EGARCH(2,2)	-5.3361	-5.3856	-5.3975
TGARCH	TGARCH(1,1)	-5.3435	-5.3983	-5.4093
	TGARCH(1,2)	-5.3319	-5.3902	-5.4011
	TGARCH(2,1)	-5.3277	-5.3826	-5.3936
	TGARCH(2,2)	-5.3198	-5.3781	-5.3890
APARCH	APARCH(1,1)	-5.3414	-5.4208	-5.4055
	APARCH(1,2)	-5.3311	-5.3874	-5.3984
	APARCH(2,1)	-5.3268	-5.3785	-5.3912
	APARCH(2,2)	-5.3246	-5.3769	-5.3888

Tabla 5.1: BIC modelos de volatilidad condicional para los retornos.

el parámetro estimado  $\alpha_1$  no es significativo, para un nivel de significancia del 5%, con una probabilidad de 56.37%. Al considerar  $\alpha_1 = 0$ , el nuevo modelo es expresado por:

$$\begin{aligned}
R_t &= a_t, \\
a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{GED}_{1.17598}(0, 1) \\
\text{Ln}(\sigma_t^2) &= -0.83994 + 0.28937 \frac{a_{t-1}}{\sigma_{t-1}} + \\
&\quad + 0.89781 \text{Ln}(\sigma_{t-1}^2),
\end{aligned} \tag{5.4}$$

con  $\text{BIC} = -5.4190$ .

Modelo	Parámetro	Estimador	Error Estándar	Probabilidad
ARCH(1)	$\alpha_0$	0.000216	0.0002	< 0.01 %
	$\alpha_1$	0.385851	0.08196	< 0.01 %
	$\nu$	4.486756	0.64196	<0.01 %
GARCH(1,1)	$\alpha_0$	0.000028	0.000012	2.1047 %
	$\alpha_1$	0.157013	0.040552	<0.01 %
	$\beta_1$	0.757110	0.067094	<0.01 %
	$\kappa$	1.168351	0.072883	<0.01 %
IGARCH(1,1)	$\alpha_0$	0.000009	0.000004	0.9011 %
	$\alpha_1$	0.161308	0.035461	<0.01 %
	$\kappa$	1.109824	0.067343	<0.01 %
EGARCH(1,1)	$\alpha_0$	-0.84274	0.301836	0.5238 %
	$\alpha_1$	0.01838	0.031855	56.3961 %
	$\beta_1$	0.89754	0.036735	<0.01 %
	$\gamma_1$	0.28686	0.055842	<0.01 %
	$\kappa$	1.17608	0.073525	<0.01 %
TGARCH(1,1)	$\alpha_0$	0.002134	0.000794	0.7233 %
	$\alpha_1$	0.170026	0.035817	<0.01 %
	$\beta_1$	0.748807	0.065046	<0.01 %
	$\gamma_1$	-0.054921	0.123345	65.6132 %
	$\kappa$	1.165914	0.072557	<0.01 %
APARCH(1,1)	$\alpha_0$	0.000168	0.000334	61.3974 %
	$\alpha_1$	0.167804	0.042064	<0.01 %
	$\gamma_1$	-0.055773	0.100267	57.8043 %
	$\beta_1$	0.755577	0.065200	<0.01 %
	$\delta$	1.585104	0.450267	0.0431 %
	$\kappa$	1.169885	0.072919	<0.01 %

Tabla 5.2: Estimación de parámetros modelos de volatilidad condicional.

Modelo	$\tilde{a}_t$			$\tilde{a}_t^2$		
	Q(10)	Q(15)	Q(20)	Q(10)	Q(15)	Q(20)
ARCH(1)	16.14 %	36.17 %	24.54 %	0.04367 %	0.02827 %	0.01275 %
GARCH(1,1)	16.65 %	37.15 %	29.44 %	95.81 %	35.66 %	42.9 %
IGARCH(1,1)	15.76 %	37.39 %	28.54 %	98.22 %	50.41 %	45.69 %
EGARCH(1,1)	15.5 %	33.63 %	27.57 %	98.75 %	29.45 %	37.51 %
TGARCH(1,1)	17.82 %	36.55 %	31.33 %	96.12 %	20.3 %	29.97 %
APARCH(1,1)	17.32 %	36.95 %	52.86 %	30.17 %	26.5 %	34.74 %

Tabla 5.3: Probabilidades obtenidas por la prueba de Ljung-Box a residuales estandarizados.

- El modelo TGARCH(1,1) de la forma,

$$\begin{aligned}
R_t &= a_t, \\
a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{GED}_{1.165}(0, 1) \\
\sigma_t^2 &= 2.134 \times 10^{-3} + (0.170026 - 0.054921 N_{t-1}) a_{t-1}^2 + \\
&\quad + 0.748807 \sigma_{t-1}^2,
\end{aligned} \tag{5.5}$$

posee el parámetro  $\gamma_1$  no significativo por ser un termino negativo, luego podría ser reducido a un modelo GARCH(1, 1) dado por:

$$\begin{aligned} R_t &= a_t, \\ a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{GED}_{1.1683}(0, 1) \\ \sigma_t^2 &= 2.848 \times 10^{-5} + 0.157013a_{t-1}^2 + 0.75711\sigma_{t-1}^2. \end{aligned} \quad (5.6)$$

- El modelo APARCH(1,1) formulado por:

$$\begin{aligned} R_t &= a_t, \\ a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim \text{GED}_{1.169}(0, 1) \\ \sigma_t^{1.23} &= 7.291 \times 10^{-4} + 0.1685(|a_{t-1}| - 0.06344a_{t-1})^{1.23} + 0.75\sigma_{t-1}^{1.23}, \end{aligned} \quad (5.7)$$

posee un parámetro negativo, no acorde a las condiciones básicas del modelo.

En la Figura 5.3 se observa la serie de retornos con dos desviaciones estándar, a un nivel de confianza del 95 %, estimadas por cada uno de los modelos tomados como referencia. Como lo estimado en la Tabla 5.3, el modelo ARCH(1) posee rezagos significativos, por lo tanto no se ajustan adecuadamente las bandas de volatilidad a la serie de retornos.

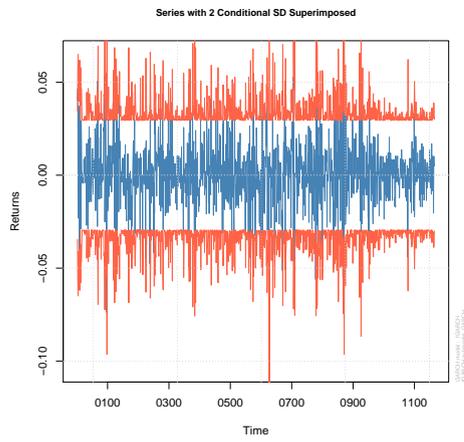
En base a los supuestos del modelo, la significancia de los parámetros, el menor BIC y las correlaciones nulas en los residuales estandarizados de los modelos mediante la prueba de Ljung-Box, el modelo que mejor se ajusta para explicar la volatilidad condicional de los retornos es un EGARCH(1,1) con  $\alpha_1 = 0$ , especificado en la ecuación (5.4).

Como lo observado en la Figura 5.5, la función de autocorrelación parcial de los residuales estandarizados al cuadrado poseen pocos rezagos significativos, por lo tanto el modelo de heterocedasticidad condicional está bien especificado; sin embargo se aplicó la prueba de Ljung-Box para verificar este resultado.

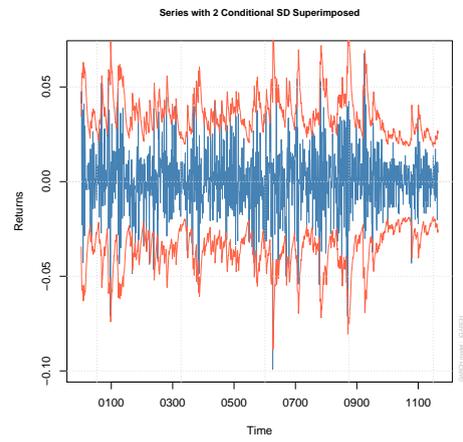
La probabilidad resultante es del 39.98 % el cual indica, a un nivel de significancia del 5 %, que las autocorrelaciones conjuntas hasta el rezago 20 de los residuales estandarizados al cuadrado no son significativas. Esto indica que el modelo en varianza se encuentra adecuadamente especificado.

Así mismo, no se considera un modelo en media condicional  $\mu_t$  ya que no existen correlaciones significativas con respecto a los residuales estandarizados, sin embargo se procedió aplicar la prueba Ljung-Box para evaluar formalmente la significancia de dichas correlaciones hasta el rezago 20.

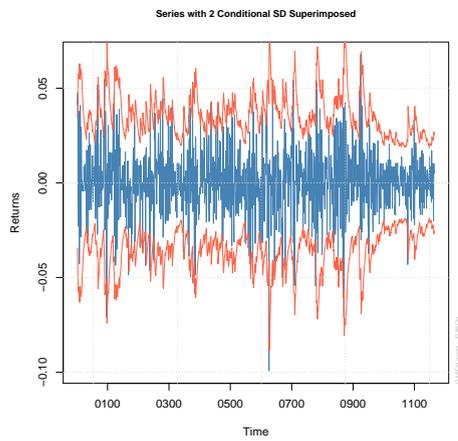
La probabilidad resultante es del 19.66 % el cual indica que no existe suficiente evidencia para rechazar la hipótesis nula y por lo tanto, las autocorrelaciones conjuntas hasta el rezago 20 de los residuales estandarizados no son significativas indicando que el modelo en media condicional dado por  $\mu_t = 0$  se encuentra correctamente especificado.



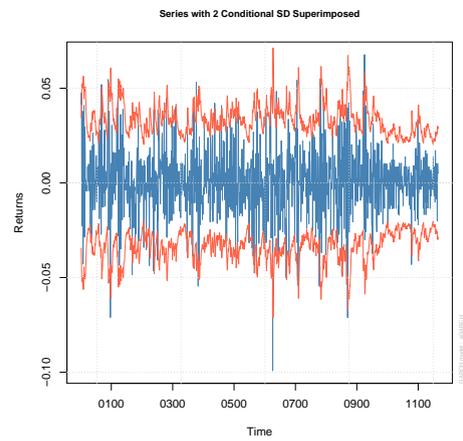
(a) ARCH(1)



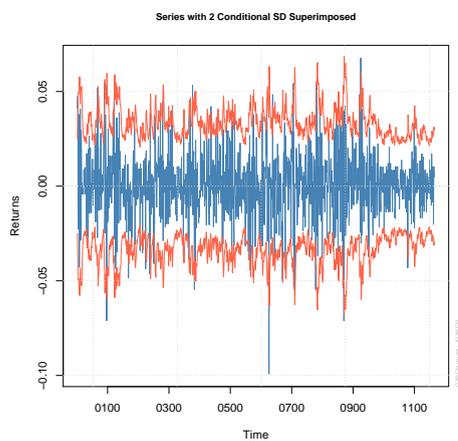
(b) GARCH(1,1)



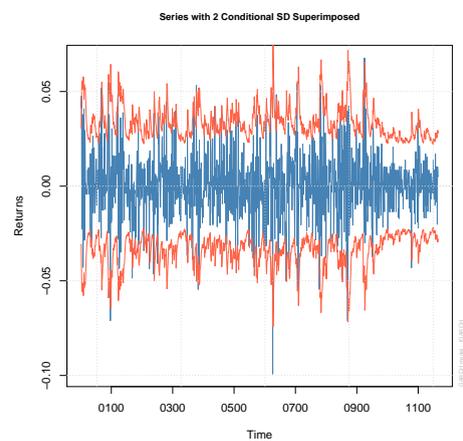
(c) IGARCH(1,1)



(d) EGARCH(1,1)

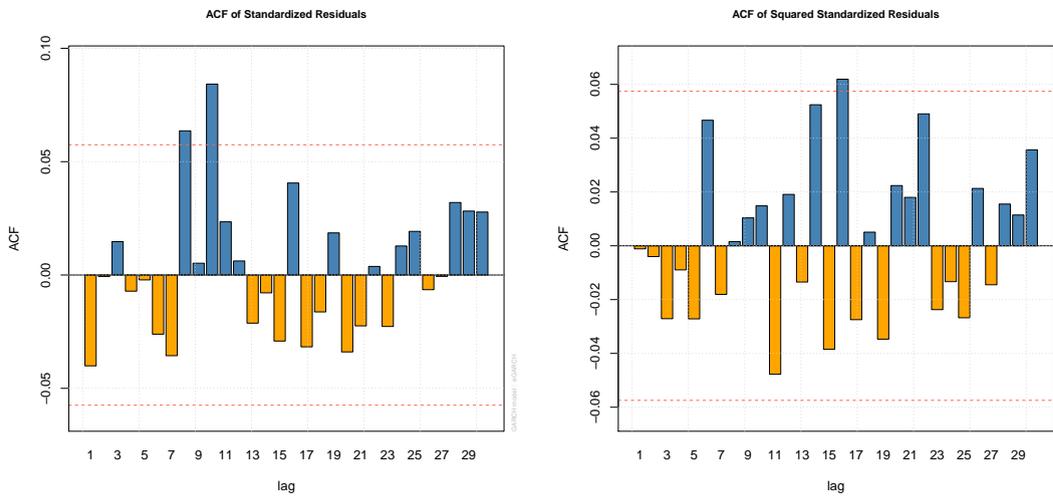


(e) TGARCH(1,1)



(f) APARCH(1,1)

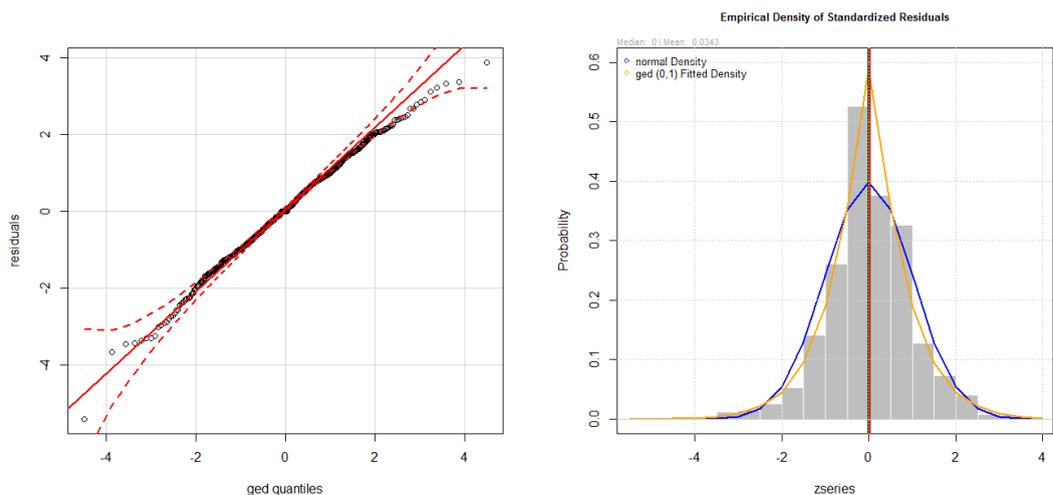
Figura 5.3: Serie de retornos CEMARGOS con dos desviaciones estándar a un nivel de confianza del 95 %.



(a) ACF de  $\tilde{a}_t$

(b) ACF de  $\tilde{a}_t^2$

Figura 5.4: ACF residuales estandarizados EGARCH(1,1).



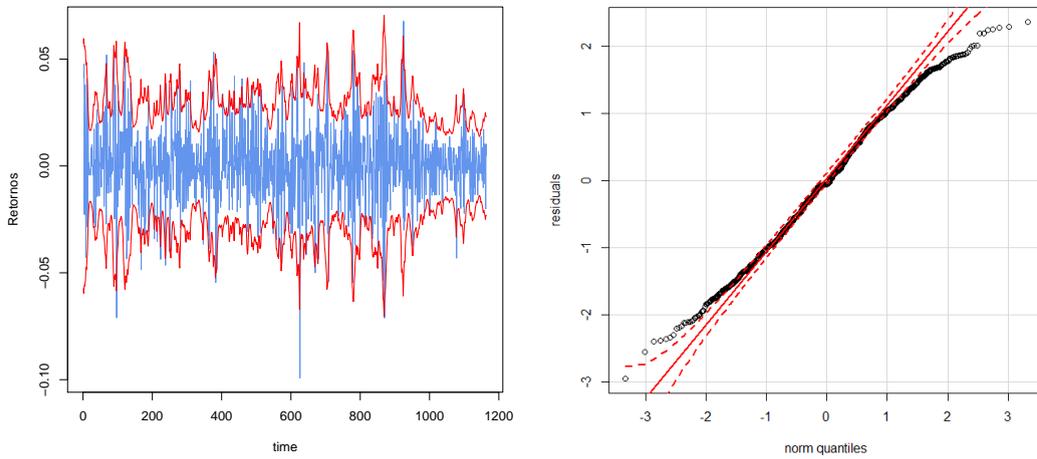
(a) QQ-Plot Distribución GED

(b) Histograma

Figura 5.5: Residuales estandarizados EGARCH(1,1).

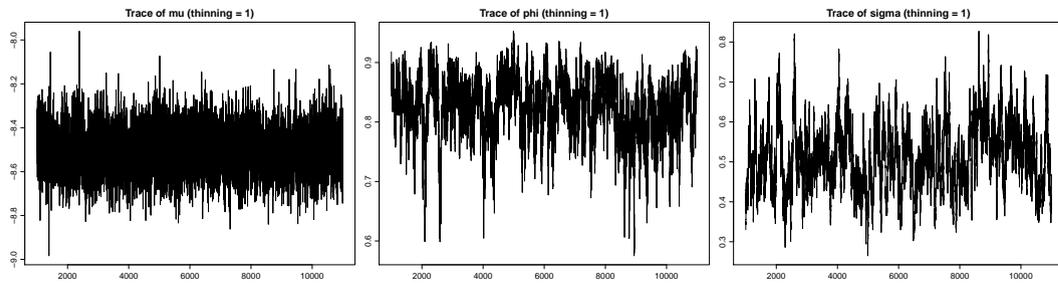
La Figura 5.3(d) muestra que el modelo en heterocedasticidad condicional logra capturar la volatilidad de los retornos a dos desviaciones estándar, con un nivel de confianza del 95 %, confirmando el buen ajuste del modelo.

Por otro lado, la verificación de la distribución de los errores (GED) con parámetro  $\kappa = 1.17$  se realizó de manera gráfica y se presentan en la Figura 5.5 observando que los residuales estandarizados se ajustan correctamente a dicha distribución.

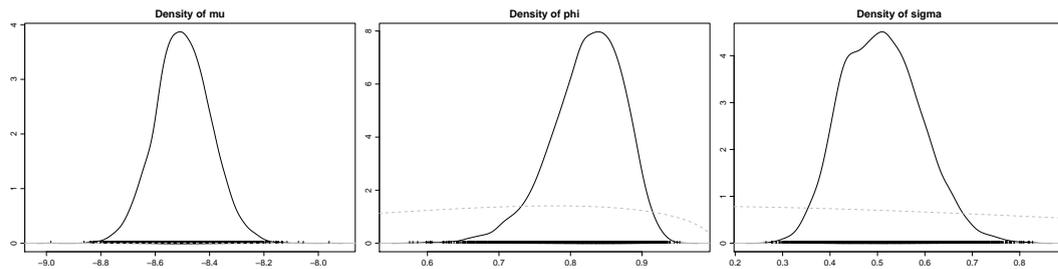


(a) Bandas de volatilidad a un nivel de confianza del 95 %

(b) Residuales estandarizados



(c) Traza distribuciones a priori



(d) Densidades distribuciones a priori

Figura 5.6: Modelo SV-AR(1).

## 5.2. Modelo de volatilidad estocástica

Mediante el paquete *stochvol* del software estadístico R, se estiman los parámetros de los modelos SV-AR(1) y SVt-AR(1) de la serie  $R_t$ , mediante distribuciones a priori, ver (Kastner, 2016b).

Parámetro	media	desviación estándar	5 %	50 %	95 %
$\mu$	-8.452	0.1086	-8.627	-8.453	-8.272
$\phi$	0.854	0.0436	0.774	0.858	0.918
$\sigma_\eta$	0.446	0.0759	0.324	0.442	0.577
$\exp(\mu/2)$	0.015	0.0008	0.013	0.015	0.016
$\sigma_\eta^2$	0.205	0.0695	0.105	0.195	0.333

Tabla 5.4: Estimaciones modelo SV-AR(1).

Parámetro	media	desviación estándar	5 %	50 %	95 %
$\mu$	-8.527	0.11	-8.717	-8.526	-8.341
$\phi$	0.854	0.046	0.769	0.860	0.920
$\sigma_\eta$	0.447	0.082	0.320	0.441	0.589
$\nu$	46.675	25	13.434	41.799	91.654
$\exp(\mu/2)$	0.014	0.00081	0.013	0.014	0.015
$\sigma_\eta^2$	0.207	0.076	0.103	0.194	0.347

Tabla 5.5: Estimaciones modelo SVt-AR(1).

El modelo SV-AR(1) es dado por:

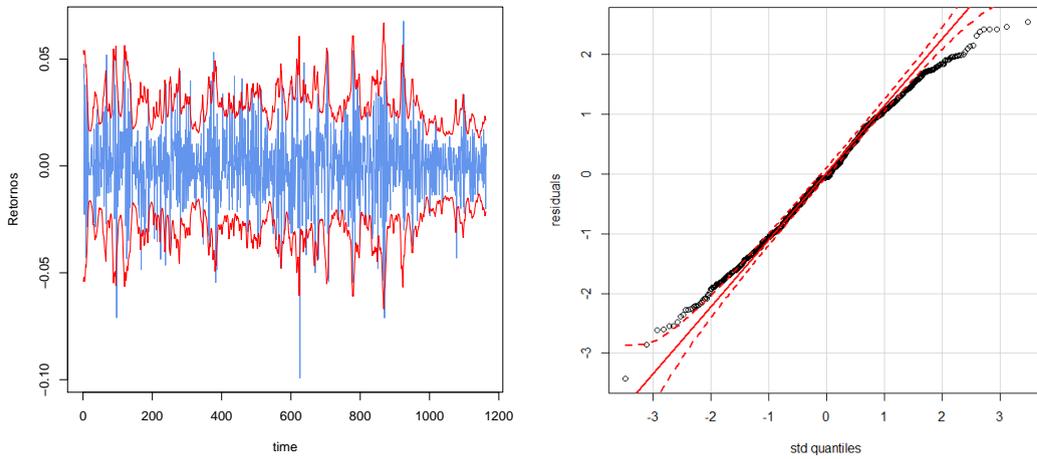
$$R_t = \sigma_t \epsilon_t,$$

$$\sigma_t^2 = \exp \left\{ \frac{h_t}{2} \right\}, \quad (5.8)$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t,$$

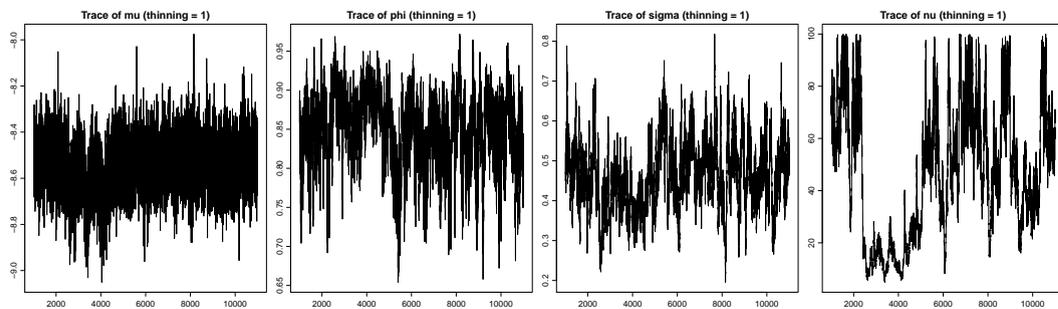
donde  $\mu \sim \mathcal{N}(0, 100)$ ,  $\frac{\phi+1}{2} \sim \mathcal{B}(5, 1.5)$  y  $\sigma^2 \sim \chi_1^2$ . Las estimaciones puntuales de dichos parámetros se muestra en la Tabla 5.4.

La traza y densidades de los parámetros, los residuales y las bandas de volatilidad para la serie de retornos a un nivel de confianza del 95 % son dadas en la Figura 5.6, cuyos parámetros fueron estimados a partir de la mediana para cada distribución a priori. Se observan que los residuales del modelo no sigue una distribución normal, pues no todos los puntos caen en las bandas de confianza del 95 %, por tanto no es un buen modelo que logre estimar apropiadamente la volatilidad de los retornos con relación al modelo EGARCH(1,1).

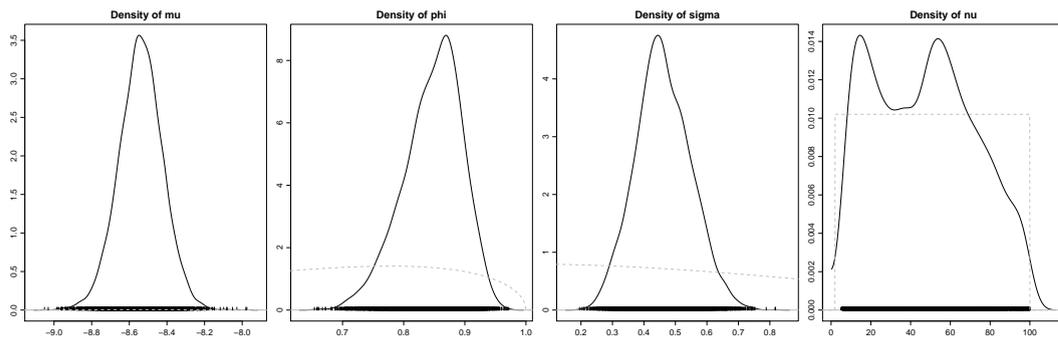


(a) Bandas de volatilidad a un nivel de confianza del 95 %

(b) Residuales estandarizados



(c) Traza distribuciones a priori



(d) Densidades distribuciones a priori

Figura 5.7: Modelo  $SVt$ -AR(1).

Análogamente, en la Tabla 5.5 se observan las estimaciones de parámetros para el modelo  $SVt$ -AR(1), con  $\nu \sim \mathcal{U}(2, 100)$ , donde la traza y densidades de cada distribución a priori, las bandas de volatilidad para la serie de retornos a un nivel de confianza del 95 % y, la distribución de sus residuales se observan en la Figura 5.7. Nuevamente, los residuales del modelo no sigue una distribución  $t$ -student, pues no todos los puntos caen en las bandas de confianza del 95 %, por tanto no es un buen modelo que logre estimar apropiadamente la volatilidad de los retornos.

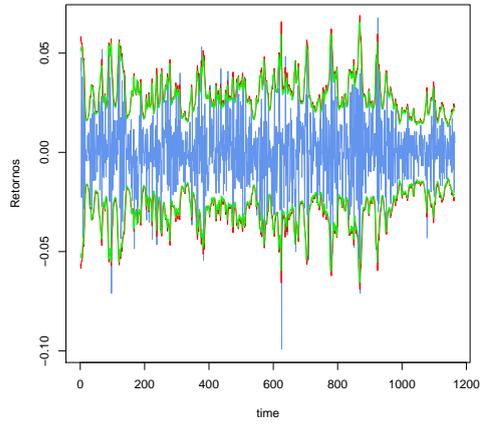
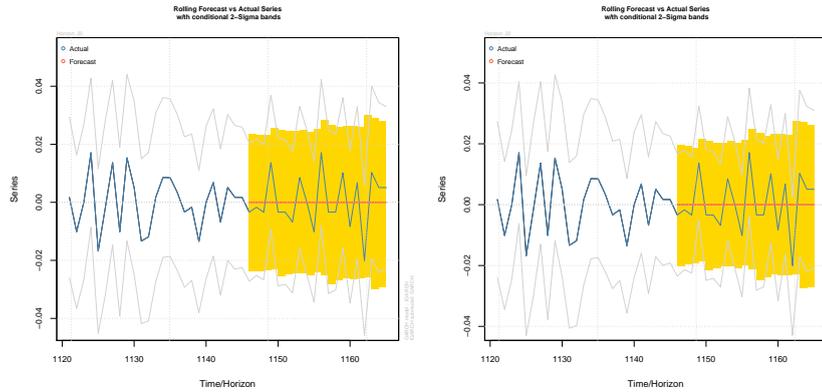
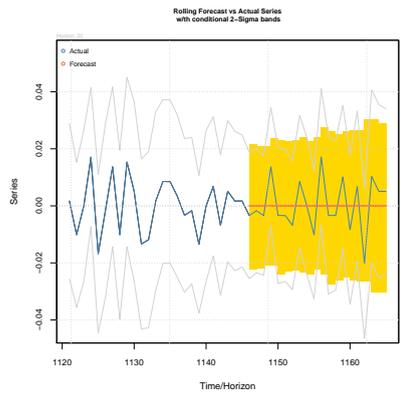


Figura 5.8: Bandas de volatilidad a un nivel de confianza del 95 %. Rojo: SV-AR(1), verde: SVt-AR(1).



(a) GARCH(1,1)

(b) IGARCH(1,1)



(c) EGARCH(1,1)

Figura 5.9: Pronóstico serie de retornos recortada con los últimos 20 datos.

Por ultimo, conforme a lo observado en la Figura 5.8, no existe diferencia significativa entre las estimaciones de volatilidad en ambos modelos, sin embargo se toma como referencia el modelo  $SVt-AR(1)$  a pesar que no cumple con los supuestos en sus residuales estandarizados.

### 5.3. Modelo de pronóstico

Tomando como referencia los modelos  $GARCH(1,1)$ ,  $IGARCH(1,1)$  y  $EGARCH(1,1)$  dados por las ecuaciones (5.6), (5.2) y (5.4), respectivamente, se analiza la efectividad de los modelos al eliminar los últimos veinte datos de la serie de retornos  $R_t$  y pronosticar sus volatilidades.

En la Figura 5.9 se observa que los últimos veinte datos de la serie de retornos  $R_t$  se encuentran dentro de las bandas de volatilidad pronosticadas por cada uno de los modelos tratados, por tanto los modelos  $GARCH(1,1)$ ,  $EGARCH(1,1)$  e  $IGARCH(1,1)$  pronostican adecuadamente las volatilidades de los últimos veinte datos de la serie de retornos  $R_t$ .

Por otro lado, al tomar como referencia los precios de cierre comprendidos entre el 21/03/2017 hasta 10/04/2017 obtenidos por la bolsa de valores de Colombia (de Colombia, 2011), se analiza si los modelos de volatilidad condicional  $GARCH(1,1)$ ,  $IGARCH(1,1)$  y  $EGARCH(1,1)$  con el modelo de volatilidad estocástica  $SVt-AR(1)$  realizan un buen pronóstico de las volatilidades de la serie de retornos.

En la Figura 5.10 se observa que las bandas de volatilidad para cada modelo, con un nivel de confianza del 95 %, abarcan completamente los retornos comprendidos entre 21/03/2017 hasta 10/04/2017, esto es, los 15 pronósticos que se encuentran después de la línea verde  $t = 1165$ . Sin embargo, dado que todas las volatilidades pronósticos para cada uno de los modelos va en forma creciente, el modelo  $SVt-AR(1)$  posee menos incertidumbre con relación a los modelos  $IGARCH(1,1)$ ,  $EGARCH(1,1)$  y  $GARCH(1,1)$ , pues sus volatilidades pronosticadas son inferiores al resto de los modelos de volatilidad condicional, conforme a lo mostrado en la Figura 5.11.

En la Tabla 5.6 se muestran los primeros quince pronósticos, por medio de los modelos de volatilidad considerados, para los precios de cierre de acciones CEMARGOS

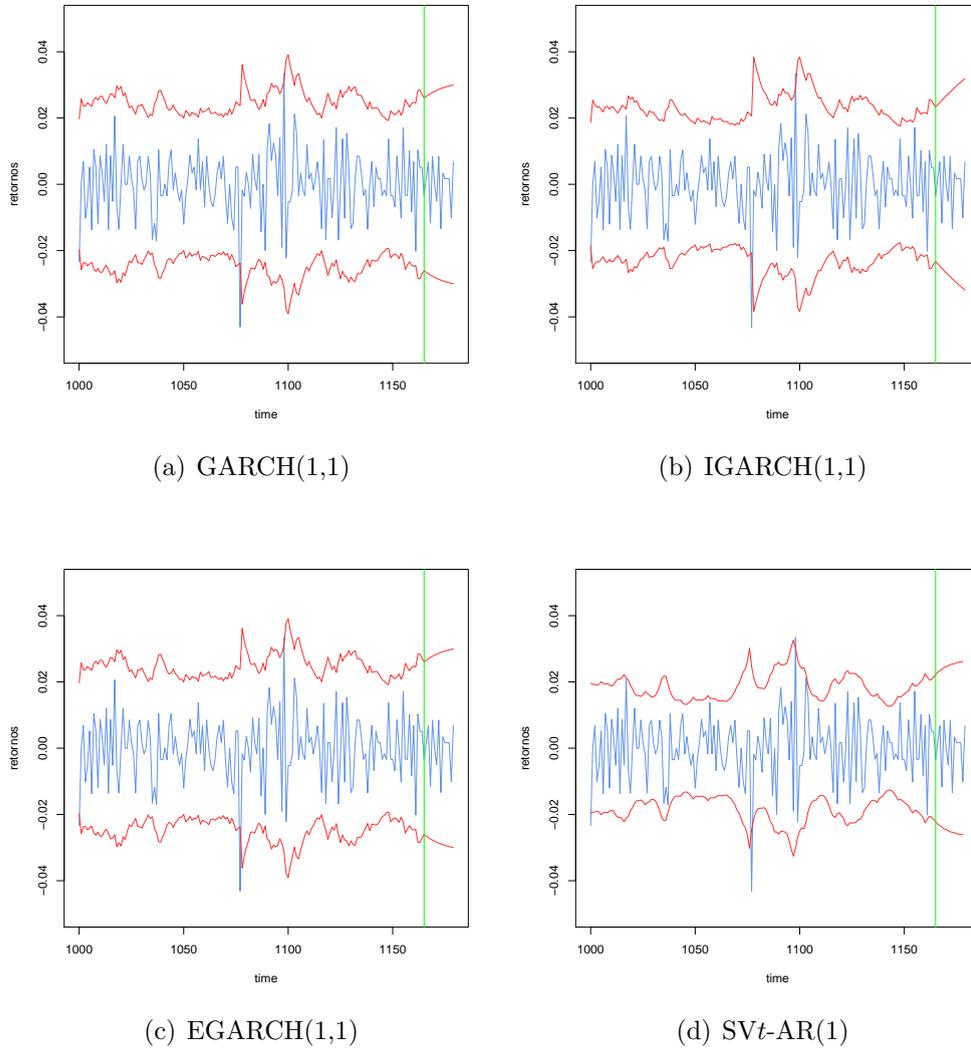


Figura 5.10: Bandas de volatilidad para la serie  $R_t$  con 15 datos posteriores a línea verde  $t = 1165$ .

entre el 21/03/2017 a 10/04/2017, donde

$$\begin{aligned} \text{MAE} &= \frac{1}{15} \sum_{t=1}^T |R_t - \widehat{R}_t|, \\ \text{RMSE} &= \frac{1}{15} \sum_{t=1}^T (R_t - \widehat{R}_t)^2, \\ \text{MAPE} &= \frac{1}{15} \sum_{t=1}^T \left| \frac{R_t - \widehat{R}_t}{R_t} \right|, \\ \text{AMAPE} &= \frac{1}{15} \sum_{t=1}^T \left| \frac{R_t - \widehat{R}_t}{R_t + \widehat{R}_t} \right|, \end{aligned}$$

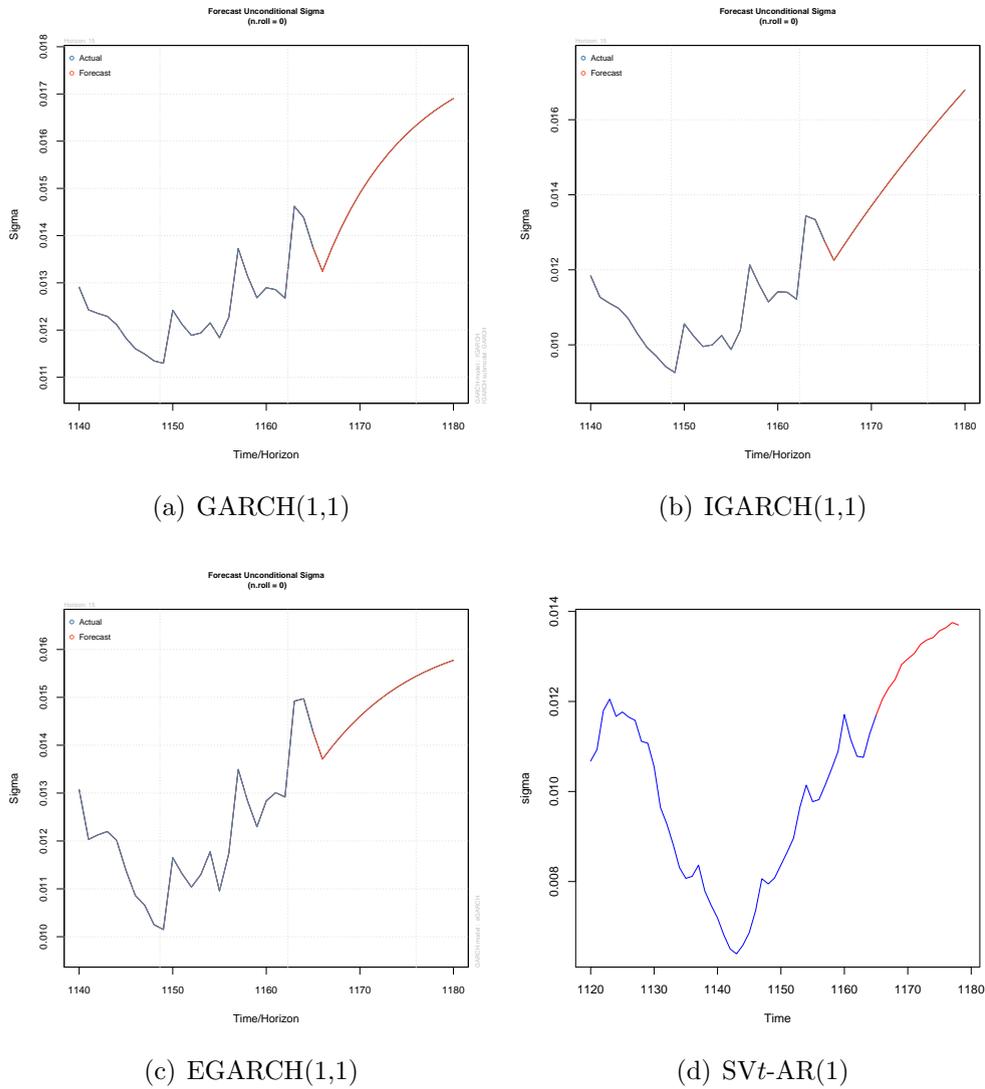


Figura 5.11: Pronóstico volatilidad de la serie  $R_t$  para 15 datos.

son los errores de dichos pronósticos, con  $R_t$  los retornos de precios de cierre reales y  $\widehat{R}_t$  los pronosticados.

El modelo SVt-AR(1) posee los menores valores MAE, RMSE, MAPE y AMAPE con relacion a los modelos GARCH(1,1), IGARCH(1,1) y EGARCH(1,1), confirmando que dicho modelo, ademas de pronosticar adecuadamente las volatilidades, pronostica con mínimo error los precios de acciones CEMARGOS.

En la Tabla 5.7 se muestran los intervalos de confianza, con un nivel de confianza del 95 %, para los pronósticos obtenidos por el modelo SVt-AR(1), el cual se evidencia que los precios reales se encuentran dentro de dichas bandas.

Fecha	Precio	Precio pronosticado			
		GARCH(1,1)	IGARCH(1,1)	EGARCH(1,1)	SVt-AR(1)
21/03/2017	11860	11993.1395	11986.1752	11996.4459	11979.3693
22/03/2017	11880	12119.4770	12102.3235	12124.9351	12086.8144
23/03/2017	11960	11814.5930	11823.0092	11819.2680	11828.1940
24/03/2017	11820	12038.6706	12028.7527	12041.2774	12017.8682
27/03/2017	11920	11934.8446	11933.3671	11939.5198	11931.0009
28/03/2017	11920	11918.4807	11918.2540	11923.6018	11917.2071
29/03/2017	11860	12019.2199	12011.7678	12020.8686	12001.4441
30/03/2017	11960	11789.1527	11796.8987	11800.2781	11810.6725
31/03/2017	11800	11754.9784	11764.7173	11767.6585	11782.1956
03/04/2017	11840	11739.4142	11749.9407	11752.8686	11769.2840
04/04/2017	11860	11737.0253	11747.6522	11750.6074	11767.3158
05/04/2017	11880	11606.5347	11621.6346	11627.5667	11659.7316
06/04/2017	11900	11761.3398	11772.6379	11773.0585	11787.1002
07/04/2017	11780	11843.0108	11853.0232	11849.5490	11854.1848
10/04/2017	11860	12013.3466	12022.3934	12008.5833	11992.3784
MAE		132.0127	125.5372	127,5650	112.0906
RMSE		152.2193	144.7215	147.5396	130.1742
MAPE		0.0111	0.0106	0.0107	0.0094
AMAPE		0.0056	0.0053	0.0054	0.0047

Tabla 5.6: Pronósticos precios de cierre CEMARGOS.

Fecha	Precio	Intervalo de confianza del 95 %	
		Limite Inferior	Limite Superior
21/03/2017	11860	9107.704	14851.035
22/03/2017	11880	9214.844	14958.785
23/03/2017	11960	8956.933	14699.455
24/03/2017	11820	9146.095	14889.641
27/03/2017	11920	9059.468	14802.534
28/03/2017	11920	9045.712	14788.703
29/03/2017	11860	9129.717	14873.171
30/03/2017	11960	8939.456	14681.889
31/03/2017	11800	8911.051	14653.340
03/04/2017	11840	8898.172	14640.396
04/04/2017	11860	8896.209	14638.423
05/04/2017	11880	8788.886	14530.578
06/04/2017	11900	8915.944	14658.257
07/04/2017	11780	8982.856	14725.513
10/04/2017	11860	9120.677	14864.080

Tabla 5.7: Intervalos de confianza para los pronósticos precios de cierre CEMARGOS dados por el modelo SVt-AR(1).

# Conclusiones

- El modelo EGARCH(1,1) es adecuado para ajustar la heterocedasticidad condicional de los retornos de los precios de la acción de Cementos Argos S.A, pues cumple con los supuestos en sus residuales estandarizados y captura adecuadamente los efectos de apalancamiento producidos por las buenas o malas noticias que acarrearán las alzas o las bajas de precio de cierre, respectivamente.
- Los modelos de volatilidad estocástica no son adecuados para explicar la volatilidad diaria de los retornos de los precios de la acción de Cementos Argos S.A, pues sus residuales estandarizados no se ajustan correctamente a la distribución normal y  $t$ -student, respectivamente. Sin embargo, dichos modelos logran capturar parcialmente los retornos.
- Los modelos de volatilidad GARCH(1,1), IGARCH(1,1), EGARCH(1,1) y SVt-AR(1) pronostican adecuadamente la incertidumbre diaria de los precios de la acción de Cementos Argos S.A. Sin embargo, el modelo SVt-AR(1) resultó ser el mejor modelo para realizar los pronósticos de volatilidades y precios de cierre diario con mínimo error.
- Al considerar variables dummies para los datos inusuales de la serie de retornos  $R_t$ , esto es, para  $t = 96, 625, 869, 924$ , e incluirlas en ecuación de la media y/o varianza de los modelos GARCH(1,1), IGARCH(1,1) y EGARCH(1,1), solamente en la ecuación de la media dichas variables fueron significativas. Sin embargo, a pesar que dichos modelos poseen menores BIC con relación a los estimados a lo largo de este trabajo, no lograron capturar completamente los retornos y sus pronósticos no son relativamente significativos respecto a los estimados por los modelos anteriores. Por lo tanto, estos modelos no se consideraron en este trabajo.
- Las autocorrelaciones significativas proporcionado por gráfica de la función de autocorrelación (ACF) en la serie de retornos  $R(t)$  determina si es posible modelar su media y, los rezagos significativos de la serie de retornos al cuadrado  $R_t^2$  es un factor determinante para modelar la volatilidad condicional.
- El número de observaciones también puede ser un limitante en el ajuste del modelo, tanto condicional como estocástica, impidiendo una mejor captura de la volatilidad en los retornos de la acción diaria de Cementos Argos.
- El modelo ajustado resultante es de gran ayuda para los inversionistas, pues da una idea del rendimiento medio de los retornos de la acción diaria de CEMARGOS facilitando la toma de decisiones a la hora de invertir.

# Bibliografía

- Argos, C. (2012). Reporte integrado 2012, cementos argos sa.
- Awartani, B. M. and Corradi, V. (2005). Predicting the volatility of the s&p-500 stock index via garch models: the role of asymmetries. *International Journal of Forecasting*, 21(1):167–183.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.
- de Colombia, B. d. V. (2011). Bolsa de valores de colombia. *Un país todos los valores*.
- Engle, R. (1982). Autorregresive conditional heterocedasticity with estimates of the variance of the.
- Franco, P. B. Á., Restrepo, D. A., and Pérez, F. O. (2007). Estudio de efectos asimétricos y día de la semana en el índice de volatilidad'vix'. *Revista Ingenierías Universidad de Medellín*, 6(11):125–147.
- Ghalanos, A. (2013). Introduction to the rugarch package. *R vignette*.
- Gutiérrez, R. d. J. and Gutiérrez, L. C. (2013). Modelación y predicción de la volatilidad con innovaciones de colas pesadas distribuidas: evidencia empírica para los petróleos maya y mezcla mexicana de exportación. (1):67–105.
- Jiménez Moscoso, J. A. (2014). *Distribuciones de probabilidad alternativas para la gestión de riesgo en mercados financieros*. PhD thesis, Universidad de Valencia.
- Kastner, G. (2015). Heavy-tailed innovations in the r package stochvol.
- Kastner, G. (2016a). Dealing with stochastic volatility in time series using the r package stochvol. *Journal of Statistical Software*, 69(5):1–30.
- Kastner, G. (2016b). Efficient bayesian inference for stochastic volatility (sv) models r package stochvol version 1.2.3.
- Tsay, R. S. (2014). *An introduction to analysis of financial data with R*. John Wiley & Sons.
- Wuertz, D., Chalabi, Y., and Miklovic, M. (2009). fgarch: Rmetrics-autoregressive conditional heteroskedastic modelling. *R package version*, 2110.

# Anexos

## Anexo 1. Código Programación R

### Paquetes a utilizar

```
library(TSA)
library(forecast)
library(tseries)
library(lmtest)
library(car)
library(FinTS)
library(rugarch)
library(fGarch)
library(arfima)
library(stochvol)
```

### Datos

```
ARGOS<-read.table("C:/ARGOS.txt", header=TRUE)
attach(ARGOS)
y=ts(ARGOS,frequency=1,start=c(1,01))
ts.plot(y, xlab="Dia", ylab="Precio acción")
```

### Análisis Precio de cierre CEMARGOS

```
ARGOS<-read.table("C:/ARGOS.txt", header=TRUE)
attach(ARGOS)
y=ts(ARGOS,frequency=1,start=c(1,01))
summary(y) ## Estadístico descriptivo
ts.plot(y, xlab="Día", ylab="Precio acción")
```

### Análisis Retornos

```
r=diff(y)
h=y[1:1164,]
r.cref<-r/h
summary(r.cref) ##Estadístico descriptivo
t.test(r.cref) ## One Sample t-test
adf.test(r.cref, alternative="stationary",k=1) ##Test Dickey-Fuller
```

## Retornos vs precios de cierre

```
par(mfrow=c(2,1))
plot(y,xlab="Día", ylab="Precio acción", main="Precio de cierre CEMARGOS")
plot(r.cref, xlab="Dia", ylab="retornos")
```

```
##ACF - PACF retornos y retornos a cuadrado
```

```
par(mfrow=c(3,2))
acf(r.cref,main="retornos",lag.max=40)
pacf(r.cref,main="retornos")
acf(r.cref^2, main="retornos^2")
pacf(r.cref^2, main="retornos^2")
```

```
###Prueba de Lyng - Box
```

```
Box.test(as.vector(r.cref),lag=1,type="Ljung-Box")
Box.test(as.vector(r.cref^2),lag=20,type="Ljung-Box")
```

## Histograma retornos vs precios de cierre

```
plotForecastErrors1 <- function(forecasterrors,titlex) {
mybinsize <- IQR(forecasterrors)/4
mysd <- sd(forecasterrors)
mymean <- mean(forecasterrors)
mymin <- min(forecasterrors) - mysd*5
mymax <- max(forecasterrors) + mysd*3
mynorm <- rnorm(10000, mean=mymean, sd=mysd)
mymin2 <- min(mynorm)
mymax2 <- max(mynorm)
if (mymin2 < mymin) { mymin <- mymin2 }
if (mymax2 > mymax) { mymax <- mymax2 }
mybins <- seq(mymin, mymax, mybinsize)
hist(forecasterrors, col="white", freq=FALSE, breaks=mybins,
main=titlex,xlab="Precio",ylab="Densidad")
myhist <- hist(mynorm, plot=FALSE, breaks=mybins)
points(myhist$mids, myhist$density, type="l", col="blue", lty=2, lwd=2)}
```

```
plotForecastErrors2 <- function(forecasterrors,titlex) {
mybinsize <- IQR(forecasterrors)/4
mysd <- sd(forecasterrors)
mymin <- min(forecasterrors) - mysd*5
mymax <- max(forecasterrors) + mysd*3
mynorm <- rnorm(10000, mean=0, sd=mysd)
mymin2 <- min(mynorm)
mymax2 <- max(mynorm)
if (mymin2 < mymin) { mymin <- mymin2 }
if (mymax2 > mymax) { mymax <- mymax2 }
```

```

mybins <- seq(mymin, mymax, mybinsize)
hist(forecasterrors, col="white", freq=FALSE,
breaks=mybins,main=titlex,xlab="retorno",ylab="Densidad")
myhist <- hist(mynorm, plot=FALSE, breaks=mybins)
points(myhist$mids, myhist$density, type="l", col="blue", lty=2, lwd=2)}

par(mfrow=c(2,1))
plotForecastErrors1(y,"Precios de cierre")
plotForecastErrors2(r.cref,"retornos")

```

## Arch Test

```

"archTest" <- function(rtn,m=10){
y=(r.cref-mean(r.cref))^2
T=length(r.cref)
atsq=y[(m+1):T]
x=matrix(0,(T-m),m)
for (i in 1:m){
x[,i]=y[(m+1-i):(T-i)]}
md=lm(atsq~x)
summary(md)}

```

```
archTest(r.cref,24)
```

## Modelos de volatilidad condicional

```

ARCH<-function (x)
{Zak = ugarchspec(mean.model = list(armaOrder = c(0, 0),
include.mean = FALSE,
arfima = FALSE), variance.model = list(garchOrder = c(1,
0), model = "fGARCH", submodel = "GARCH"),
distribution.model = "ged")
model <- ugarchfit(x, spec = Zak)
return(model)}
ARCH(r.cref)
plot(ARCH(r.cref))

```

```

GARCH<-function (x)
{Zak = ugarchspec(mean.model = list(armaOrder = c(0, 0),
include.mean = FALSE,
arfima = FALSE), variance.model = list(garchOrder = c(1,
1), model = "fGARCH", submodel = "GARCH"),
distribution.model = "ged")
model <- ugarchfit(x, spec = Zak)
return(model)}
ARCH(r.cref)
plot(ARCH(r.cref))

```

```

EGARCH<-function (x)
{Zak = ugarchspec(mean.model = list(armaOrder = c(0, 0),
include.mean = FALSE,
      arfima = FALSE), variance.model = list(garchOrder = c(1,
      1), model = "eGARCH"), distribution.model = "ged",
fixed.pars=list(alpha1=0))
model <- ugarchfit(x, spec = Zak)
return(model)}
EGARCH(r.cref)
plot(EGARCH(r.cref))

```

```

IGARCH<-function (x)
{Zak = ugarchspec(mean.model = list(armaOrder = c(0, 0),
include.mean = FALSE,
      arfima = FALSE), variance.model = list(garchOrder = c(1,
      1), model = "iGARCH"),distribution.model = "ged")
model <- ugarchfit(x, spec = Zak)
return(model)}
IGARCH(r.cref)
plot(IGARCH(r.cref))

```

```

TGARCH<-function (x)
{Zak = ugarchspec(mean.model = list(armaOrder = c(0, 0),
include.mean = FALSE,
      arfima = FALSE), variance.model = list(garchOrder = c(1,
      1), model = "fGARCH", submodel = "TGARCH"),
      distribution.model = "ged")
model <- ugarchfit(x, spec = Zak)
return(model)}
TGARCH(r.cref)
plot(TGARCH(r.cref))

```

```

APARCH<-function (x)
{Zak = ugarchspec(mean.model = list(armaOrder = c(0, 0),
include.mean = FALSE,
      arfima = FALSE), variance.model = list(garchOrder = c(1,
      1), model = "fGARCH", submodel = "APARCH"),
      distribution.model = "ged")
model <- ugarchfit(x, spec = Zak)
return(model)}
TGARCH(r.cref)
plot(TGARCH(r.cref))

```

### Pronóstico modelo GARCH para 15 días posteriores y 20 días anteriores

```
pred = ugarchforecast(GARCH(r.cref), n.ahead = 15, n.roll=20)
```

## Modelos de volatilidad estocástica

```
rtn<-r.cref-mean(r.cref)
svar<-svsample(rtn, designmatrix = "ar1")
svtar <- svsample(rtn, priornu = c(2, 100), designmatrix = "ar1")

##Análisis modelo SVt-AR(1)

summary(svtar, showlatent = FALSE)
plot(svtar, showobs = FALSE)

### Bandas de volatilidad estimadas modelo SV-AR(1) en la serie de retornos

dim(res$latent)
ht=apply(svar$latent,2,median)
v1=exp(ht/2)
upp1=0+1.96*v1
low1=0-1.96*v1
plot(r.cref,xlab="time",ylab="log-retornos",type="l", col="cornflowerblue")
lines(upp1,col="red1",type="l")
lines(low1,col="red1",type="l")

### Pronósticos para 15 días

fore <- predict(svar, 15)
summary(fore)
plot(svar, forecast = fore)
```