



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Estimación del Comportamiento de la Volatilidad a través
de Series de Tiempo y Medidas de Riesgo

Tesis presentada al

Colegio de Matemáticas

como requisito para la obtención del grado de

LICENCIADO EN MATEMÁTICAS

presenta

Victor Manuel Ortiz Rosas

Asesorado por

Dr. Hugo Adán Cruz Suárez

Puebla Pue.
Octubre de 2022



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Estimación del Comportamiento de la Volatilidad a través
de Series de Tiempo y Medidas de Riesgo

Tesis presentada al

Colegio de Matemáticas

como requisito parcial para la obtención del grado de

LICENCIADO EN MATEMÁTICAS

por

Victor Manuel Ortiz Rosas

Asesorado por

Dr. Hugo Adán Cruz Suárez

Puebla Pue.
Octubre de 2022

Título: Estimación del Comportamiento de la Volatilidad a través de Series de Tiempo y Medidas de Riesgo

Estudiante: VICTOR MANUEL ORTIZ ROSAS

COMITÉ

Dr. Rei Israel Ortega Gutiérrez
Presidente

Dr. José Juan Castro Alva
Secretario

Dr. Rubén Blancas Rivera
Vocal

Dr. Hugo Adán Cruz Suárez
Asesor

Agradecimientos

EL termino de esta etapa de mi vida no podría haberse llevado a cabo sin el apoyo incondicional por parte de mi familia, por lo que agradezco infinitamente a la misma, a mis hermanas, Marina, Araceli y Miriam, por la confianza que han puesto en mí a lo largo de estos años. Agradezco a mi padre Manuel Ortiz Nieto por sus enseñanzas a lo largo de mi vida y agradezco de manera especial a mi madre Ma. Antonia Rosas Jiménez por la vida, por brindarme apoyo, cariño y atención, además de demostrarme confianza en cada momento a lo largo de toda mi vida, sin su ayuda nada de esto podría haber sido posible. También agradezco a mis sobrinos, Alejandro y Karla, en ellos he encontrado la alegría y diversión que me ha hecho falta en aquellos momentos difíciles.

A mis amigos, conocidos de la universidad y aquellos con los que he pasado tiempo y hecho buenos momentos, con los cuales compartí horas en clases, en la biblioteca y en general las instalaciones de la universidad.

A la FCFM por permitirme en sus aulas conocer esta bella ciencia, maravillarme con sus conceptos e ideas y ayudarme a encontrar en las matemáticas un mundo en el que deseo quedarme.

A mi jurado por su valiosa retroalimentación para la mejora de este trabajo, Dr. Rei Israel Ortega Gutiérrez, Dr. José Juan Castro Alva, y Dr. Rubén Blancas Rivera.

A la Vicerrectoría de Investigación y Estudios de Posgrado por el apoyo para la elaboración de este trabajo.

Finalmente le agradezco al Dr. Hugo Adán Cruz Suárez por su apoyo y la aceptación para la dirección de este trabajo, además de sus enseñanzas y consejos, sin los cuales este trabajo no podría haberse concretado.

A todos ustedes, muchas gracias.

Índice general

Introducción	9
1. Medidas de Riesgo	11
1.1. Valor en Riesgo	14
1.2. Esperanza de Cola Condicional	17
2. Series de Tiempo	23
2.1. Modelo GARCH	24
2.2. Modelo HAR-RV	27
3. Aplicación al IPC y S&P500	31
3.1. Hechos Estilizados	31
3.2. Resultados	36
Conclusiones	43
Referencias	44
Anexo A. Códigos en R	47

Introducción

En el mundo financiero a menudo se quisiera resumir toda la información acerca de un activo a un sólo número, sin embargo esto podría llegar a ser muy difícil por no decir imposible. Ahora bien, hay ciertos aspectos o características particulares de interés acerca de ese activo que sí podría resumirse a un número. Para este trabajo nos gustaría conocer un número que nos indique cuales podrían ser los peores rendimientos al invertir en un activo, y otro que nos indique un valor de lo que se podría esperar como rendimiento, suponiendo que los rendimientos ahora serán mayores al número antes mencionado. Candidatos a este par de números podrían ser el **Valor en Riesgo** (VaR) [14] y la **Esperanza de Cola Condicional** (TCE) [13].

Además de estos números nos gustaría tener otro que nos indicara que tan dispersos podrían estar los rendimientos o en que intervalo podrían encontrarse. Esta idea es la que motiva el estudio del concepto de volatilidad. La forma más simple de medir la volatilidad se basa en el cálculo de la desviación típica asociada a un conjunto de rendimientos observados. Así, se obtiene un sólo valor que representa la dispersión global de los datos, pero con este número no podríamos saber como es que esta dispersión se comporta a lo largo del tiempo. Algo que se ha llegado a observar en las series de rendimientos es que esta variabilidad o dispersión de los datos parece no ser constante [21], sino que existen intervalos de tiempo en los que los datos se dispersan en un intervalo grande y otros periodos en lo que los datos parecen variar muy poco, es por eso que algunos autores han sugerido que la volatilidad debería pensarse como un proceso estocástico y se han propuesto algunos modelos como el **Modelo Generalizado de Heterocedasticidad Condicional Autoregresiva** (GARCH) [2] propuesto por Bollerslev en 1988 o el **Modelo de Volatilidad Relizada Autorregresivo Heterogéneo** (HAR-RV) [6] propuesto por Corsi en 2004.

Esta tesis está conformada por tres capítulos, en el Capítulo 1, se estudia de manera teórica las medidas de riesgo, en él se definen las medidas de ries-

go con las que se trabajará y se demuestran algunos resultados importantes sobre sus propiedades. En el Capítulo 2, estudiaremos los modelos de volatilidad, en este se presentarán los modelos a estudiar y veremos como estimar los parámetros para ajustar estos modelos a alguna serie de rendimientos. Finalmente en el Capítulo 3 a modo de ejemplo aplicamos las medidas de riesgo y los modelos de volatilidad a las series de rendimientos de los índices **Standard & Poor's 500 Index** (S&P500) e **Índice de Precios y Cotizaciones** (IPC) con el objetivo de ejemplificar el uso y aplicación de estas herramientas y hacer una comparación entre los mercados financieros americano y mexicano. Además, mencionamos que la razón del porque trabajar con los rendimientos en lugar de trabajar directamente con los precios del activo es que nos permite hacer comparaciones entre varios activos sin importar su moneda o si estos están en números muy grandes o demasiado pequeños.

Por último en esta parte introductoria cabe mencionar que para este trabajo se usó el lenguaje de programación **R**, y en el **Anexo A**, se pueden encontrar todos los códigos usados para obtener los diferentes resultados que iremos viendo, estos códigos se incluyen para una posible consulta o de existir algún interés en particular.

Capítulo 1

Medidas de Riesgo

Como mencionamos en la introducción, estamos interesados en resumir ciertos aspectos acerca de un activo a un número y para ello haremos uso de la teoría de medidas de riesgo que desarrollaremos a lo largo de este capítulo. En adelante, vamos a considerar fijo un espacio de probabilidad (Ω, \mathcal{F}, P) . A partir de este, sea $L^0(\Omega, \mathcal{F}, P)$ el conjunto de todas las variables aleatorias $X : \Omega \rightarrow \mathbb{R}$. Para definir los riesgos financieros, sobre este último conjunto, consideramos el subconjunto $\mathcal{M} \subseteq L^0(\Omega, \mathcal{F}, P)$ de variables aleatorias que representan las pérdidas de un portafolio de inversión en un horizonte de tiempo finito Δ fijo, a estas variables les llamaremos riesgos [10]. Además vamos a suponer que \mathcal{M} es cerrado bajo sumas y productos por escalares no negativos. Esto es, $X_1 + X_2 \in \mathcal{M}$ y $\lambda X_1 \in \mathcal{M}$ para todo $X_1, X_2 \in \mathcal{M}$ y para todo $\lambda \in \mathbb{R}$ con $\lambda \geq 0$.

Definición 1.0.1. *Una medida de riesgo, es una función $\rho : \mathcal{M} \rightarrow \mathbb{R}$.*

Observación 1.0.1. *Hay que notar en la definición anterior que en realidad el dominio \mathcal{M} de una medida de riesgo ρ se restringe a donde tenga sentido evaluar a la función.*

Ejemplo 1.0.1. *Dado $X \in \mathcal{M}$, algunas medidas de riesgo pueden ser:*

- a) $\rho_1(X) = E[X]$,
- b) $\rho_2(X) = \text{Var}[X]$,
- c) $\rho_3(X) = k$, con k una constante.

Todas las funciones anteriores pueden ser consideradas como medidas de riesgo, sin embargo dependiendo del fenómeno que estemos interesados en estudiar existen una serie de propiedades deseables que se espera cumpla una medida de riesgo para que nos ayude a cuantificar el riesgo.

Definición 1.0.2. Dada una medida de riesgo $\rho : \mathcal{M} \rightarrow \mathbb{R}$ se definen las siguientes propiedades que puede cumplir

- a) *Constancia:* $\rho(c) = c$, para toda $c \in \mathbb{R}$.
- b) *Sin estafa:* para todo $X \in \mathcal{M}$, se satisface que $\rho(X) \leq \sup[X]$, donde $\sup[X] = \sup\{X(w) | w \in \Omega\}$ si existe.
- c) *No negatividad:* para todo $X \in \mathcal{M}$, se satisface que $\rho(X) \geq E[X]$.
- d) *Monotonicidad:* Sean $X, Y \in \mathcal{M}$, si $X \leq Y$ P -casi seguramente (denotado $P - c.s.$), entonces $\rho(X) \leq \rho(Y)$.
- e) *Homogeneidad Positiva:* para todo $X \in \mathcal{M}$ y para todo $\lambda \geq 0$ se satisface $\rho(\lambda X) = \lambda \rho(X)$.
- f) *Invariante ante traslaciones:* para todo $X \in \mathcal{M}$ y para todo $\alpha \in \mathbb{R}$ se satisface $\rho(X + \alpha) = \rho(X) + \alpha$.
- g) *Subaditividad:* para todo $X, Y \in \mathcal{M}$ se satisface $\rho(X + Y) \leq \rho(X) + \rho(Y)$.
- h) *Invarianza en ley:* Si $X = Y$ en distribución, entonces $\rho(X) = \rho(Y)$.
- i) *Convexidad:* para todo $X, Y \in \mathcal{M}$ y $0 \leq k \leq 1$ se satisface $\rho(kX + (1 - k)Y) \leq k\rho(X) + (1 - k)\rho(Y)$.

Observación 1.0.2. Cada una de estas propiedades tiene una interpretación que justifica el porque es deseable que se cumplan:

- a) *Constancia:* Si la variable de pérdida en cualquier caso es constante, entonces el riesgo al que uno se expone, debería ser esa constante.
- b) *Sin estafa:* El riesgo al que uno se expone, no debería ser mayor a la peor pérdida posible.
- c) *No negatividad:* El riesgo al que uno se expone debería ser mayor al esperado.
- d) *Monotonicidad:* Este propiedad intuitiva y lo que nos dice es que si la variable de pérdida Y es mayor que X , entonces los riesgos correspondientes deben satisfacer la misma desigualdad.
- e) *Homogeneidad Positiva:* Este axioma significa que medir una proporción λ de un una variable de pérdida equivale a considerar la medida de esa variable por la misma proporción.

- f) *Invariante ante traslaciones:* Esta propiedad significa que agregar una cantidad fija al riesgo cambiará el valor de medida del riesgo en la misma cantidad.
- g) *Subaditividad:* El razonamiento detrás de este axioma es integrar la idea de diversificación, es decir, el riesgo global de un portafolio formado por dos o más activos es menor o igual que la suma de los riesgos individuales.
- h) *Invarianza en ley:* Si dos variables pierden lo mismo con la misma probabilidad, entonces serán igual de riesgosas.
- i) *Convexidad:* La convexidad se asegura de que al diversificar el riesgo no sea mayor. El riesgo de una posición diversificada $kX + (1-k)Y$ es menor o igual que la media ponderada de los riesgos individuales.

Cabe notar que todas estas propiedades pueden no tener relacion entre sí, y que es posible que una medida de riesgo cumpla todas, algunas o ninguna. Es por esto, que combinar algunas de estas propiedades pueden darnos medidas de riesgo que resulten más eficientes según el fenómeno a estudiar, algunas de estas posibles combinaciones tienen nombres especiales como ahora definiremos.

Definición 1.0.3. *Una medida de riesgo $\rho : \mathcal{M} \rightarrow \mathbb{R}$ se dice que es una medida de riesgo monetaria si cumple con ser monótona e invariante ante traslaciones.*

Un ejemplo para este tipo de medidas de riesgo es el *VaR* como veremos más adelante.

Definición 1.0.4. *Una medida de riesgo $\rho : \mathcal{M} \rightarrow \mathbb{R}$ se dice que es una medida de riesgo convexa si satisface las propiedades de monotonidad, invarianza ante traslaciones y convexidad.*

Las medidas de riesgo convexas se presentaron y estudiaron por primera vez en 2002 y podemos encontrar más al respecto en [7].

Definición 1.0.5. *Una medida de riesgo $\rho : \mathcal{M} \rightarrow \mathbb{R}$ se dice que es una medida de riesgo coherente si satisface las siguientes propiedades: monotonidad, homogeneidad positiva, invarianza ante traslaciones y subaditividad.*

Este tipo de medidas de riesgo son las más relevantes, pues sus propiedades las hacen aptas para lo que queremos estudiar [3].

Para verificar algunas propiedades de las medidas de riesgo no resultará útil el siguiente lema.

Lema 1.0.1. Sean $X, Y \in \mathcal{M}$, si $X \leq Y$ (P -c.s.), entonces $F_Y(x) \leq F_X(x)$ para todo $x \in \mathbb{R}$.

Demostración. Como $X \leq Y$ P -c.s., esto significa que $P(X \leq Y) = 1$ y $P(X > Y) = 0$. Sea $x \in \mathbb{R}$

$$\begin{aligned} F_Y(x) &= P(Y \leq x) \\ &= P(Y \leq x, X \leq Y) + P(Y \leq x, X > Y) \\ &= P(Y \leq x, X \leq Y) \end{aligned}$$

y

$$\begin{aligned} P(X \leq x, X \leq Y) &= P(X \leq x, X \leq Y) + P(X \leq x, X > Y) \\ &= P(X \leq x) \\ &= F_X(x). \end{aligned}$$

Pero, observe que $\{Y \leq x, X \leq Y\} \subseteq \{X \leq x, X \leq Y\}$, por lo que $P(Y \leq x, X \leq Y) \leq P(X \leq x, X \leq Y)$ y con esto finalmente se tiene que $F_Y(x) \leq F_X(x)$. \square

1.1. Valor en Riesgo

Vamos a trabajar con dos medidas de riesgo, una de ellas cumple con ser monetaria y la otra cumple con ser coherente, estas medidas las definiremos a continuación y verificaremos que efectivamente cumplen con las propiedades que señalaremos en cada una. Veamos una de las medidas de riesgo más conocida y ampliamente usado en diferentes contextos.

Definición 1.1.1. El valor en riesgo (*VaR* por sus siglas en inglés) a un nivel de confianza $\alpha \in (0, 1)$ es una función $VaR_\alpha : \mathcal{M} \rightarrow \mathbb{R}$ tal que para cada $X \in \mathcal{M}$:

$$VaR_\alpha(X) = \inf\{l \in \mathbb{R} | P(X > l) < 1 - \alpha\} \quad (1.1)$$

$$= \inf\{l \in \mathbb{R} | P(X \leq l) \geq \alpha\} \quad (1.2)$$

$$= \sup\{l \in \mathbb{R} | F_X(l) < \alpha\}. \quad (1.3)$$

Lo que nos indica el VaR es por debajo de que valor podemos encontrar los rendimientos más pequeños con probabilidad α .

Proposición 1.1.1. El VaR_α para cualquier nivel de confianza $\alpha \in (0, 1)$ satisface las siguientes propiedades:

- a) *Constancia*
- b) *Sin estafa*
- c) *Monotonicidad*
- d) *Invariante ante translaciones*
- e) *Homogeneidad positiva*
- f) *Invarianza en ley*

Demostración. Sea $\alpha \in (0, 1)$

- a) Sea $c \in \mathbb{R}$,

$$\begin{aligned} VaR_\alpha(c) &= \sup\{l \in \mathbb{R} | F_c(l) < \alpha\} \\ &= c. \end{aligned}$$

- b) Sea $\sup[X] := \sup\{X(w) | w \in \Omega\}$, entonces $P(X \leq \sup[X]) = 1$, entonces $\sup[X]$ es cota superior de $\{l \in \mathbb{R} | P(X \leq l) \geq \alpha\}$, y así

$$VaR_\alpha(X) = \inf\{l \in \mathbb{R} | P(X \leq l) \geq \alpha\} \leq \sup[X].$$

- c) Sean $X, Y \in \mathcal{M}$ tales que $X \leq Y$ P -c.s., por el Lema 1.0.1 se tiene que $F_Y(l) \leq F_X(l)$, es decir

$$P(Y \leq l) \leq P(X \leq l).$$

Así, para cada $l \in \mathbb{R}$ si $P(Y \leq l) \geq \alpha$, entonces $P(X \leq l) \geq \alpha$, es decir

$$\{l \in \mathbb{R} | P(Y \leq l) \geq \alpha\} \subseteq \{l \in \mathbb{R} | P(X \leq l) \geq \alpha\}$$

y por lo tanto

$$\inf\{l \in \mathbb{R} | P(X \leq l) \geq \alpha\} \leq \inf\{l \in \mathbb{R} | P(Y \leq l) \geq \alpha\}$$

es decir $VaR_\alpha(X) \leq VaR_\alpha(Y)$.

- d) Ahora, sea $c \in \mathbb{R}$,

$$\begin{aligned} VaR_\alpha(X + c) &= \inf\{l \in \mathbb{R} | P(X + c \leq l) \geq \alpha\} \\ &= \inf\{l \in \mathbb{R} | P(X \leq l - c) \geq \alpha\} \\ &= \inf\{r + c \in \mathbb{R} | P(X \leq (r + c) - c) \geq \alpha\} \\ &= \inf\{r + c \in \mathbb{R} | P(X \leq r) \geq \alpha\} \\ &= \inf\{r \in \mathbb{R} | P(X \leq r) \geq \alpha\} + c \\ &= VaR_\alpha(X) + c. \end{aligned}$$

e) Ahora, sea $\lambda > 0$, tenemos:

$$\begin{aligned}
 VaR_\alpha(\lambda X) &= \inf\{l \in \mathbb{R} | P(\lambda X \leq l) \geq \alpha\} \\
 &= \inf\{l \in \mathbb{R} | P(X \leq l/\lambda) \geq \alpha\} \\
 &= \inf\{r\lambda \in \mathbb{R} | P(X \leq (r\lambda)/\lambda) \geq \alpha\} \\
 &= \inf\{r\lambda \in \mathbb{R} | P(X \leq r) \geq \alpha\} \\
 &= \inf\{r \in \mathbb{R} | P(X \leq r) \geq \alpha\}\lambda \\
 &= \lambda VaR_\alpha(X).
 \end{aligned}$$

Si $\lambda = 0$ entonces $VaR_\alpha(\lambda X) = VaR_\alpha(0) = 0 = \lambda VaR_\alpha(X)$.

f) Sean $X, Y \in \mathcal{M}$ tales que $F_X = F_Y$, entonces:

$$VaR_\alpha[X] = \sup\{l \in \mathbb{R} | F_X(l) < \alpha\} = \sup\{l \in \mathbb{R} | F_Y(l) < \alpha\} = VaR_\alpha[Y].$$

□

A pesar de cumplir todas estas propiedades antes demostradas tenemos que el VaR_α no es una medida de riesgo coherente, pues no cumple con ser subaditiva como veremos a continuación:

Ejemplo 1.1.1. Consideremos dos variables aleatorias X_1, X_2 independientes e idénticamente distribuidas Bernoulli, $B(0.03)$, en este caso su función de masa es:

$$f_{X_1}(x) = f_{X_2}(x) = \begin{cases} 0.97 & \text{si } x = 0 \\ 0.03 & \text{si } x = 1 \\ 0 & \text{en otro caso.} \end{cases}$$

Se puede observar que $X_1 + X_2 \sim Bin(2, 0.03)$ ([16]) y su función de masa está dada por:

$$f(x) = \begin{cases} 0.9409 & \text{si } x = 0 \\ 0.0582 & \text{si } x = 1 \\ 0.0009 & \text{si } x = 2 \\ 0 & \text{si en otro caso.} \end{cases}$$

De aquí se tiene que $VaR_{0.95}(X_1) = 0 = VaR_{0.95}(X_2)$ y por lo tanto $VaR_{0.95}(X_1) + VaR_{0.95}(X_2) = 0$, mientras que $VaR_{0.95}(X_1 + X_2) = 1$, es decir que $VaR_{0.95}(X_1 + X_2) > VaR_{0.95}(X_1) + VaR_{0.95}(X_2)$ lo cual demuestra que el VaR_α en general no cumple con ser subaditiva y por lo tanto no es una medida de riesgo coherente.

Observación 1.1.1. *A pesar de no ser una medida de riesgo coherente, tiene ciertas ventajas por las cuales es frecuentemente usado, a continuación se mencionan algunas*

- a) *Es aplicable a portafolios con instrumentos no lineales, tales como opciones, con distribuciones de pérdidas no necesariamente simétricas, ni normales.*
- b) *El VaR es probabilístico y le proporciona al gestor del riesgo información útil sobre las probabilidades asociadas con un monto específico de pérdidas.*
- c) *El VaR se expresa en la unidad de medida más simple y fácil de entender en el contexto del riesgo financiero, esto es, pérdida en dinero.*

Existen diferentes métodos para calcular el VaR_α y en particular para este trabajo se usó el método histórico. Este método consiste en que si queremos calcular el VaR diario a un nivel de confianza $\alpha \in (0, 1)$ tenemos que calcular el α -cuantil de los rendimientos diarios observados de la cartera. Una de las ventajas de este método es que no impone supuestos distributivos; la distribución de los rendimientos de la cartera se construye a partir de información histórica. Por lo tanto, a veces este método se llama método no paramétrico. Las observaciones se recopilan a partir de una ventana de tiempo predeterminada, que podría ser un periodo en el que de manera particular nos interesa conocer el comportamiento de los rendimientos o algo más sencillo como tomar la serie de rendimientos en el año más reciente. Además de este método existen otros como el método híbrido o Monte Carlo, para más detalles puede consultarse [14] p. 165.

1.2. Esperanza de Cola Condicional

Debido al inconveniente que presenta el VaR al no ser una medida de riesgo coherente, además de no dar información acerca de lo que pasa en la cola derecha de la distribución, se propone una medida alternativa que es la siguiente.

Definición 1.2.1. *La esperanza de cola condicional (TCE por sus siglas en inglés) a un nivel de confianza $\alpha \in (0, 1)$ es una función $TCE_\alpha : \mathcal{M} \rightarrow \mathbb{R}$ tal que para cada $X \in \mathcal{M}$:*

$$TCE_\alpha(X) = E[X|X > VaR_\alpha(X)]. \quad (1.4)$$

Observación 1.2.1. Por definición de esperanza condicional [18] sabemos lo siguiente

$$\begin{aligned} TCE_\alpha(X) &= E[X|X > VaR_\alpha(X)] \\ &= \int_{-\infty}^{\infty} x dF_{X|X > VaR_\alpha(X)}(x). \end{aligned}$$

Pero notemos que

$$\begin{aligned} F_{X|X > VaR_\alpha(X)}(x) &= \frac{P(X \leq x, X > VaR_\alpha(X))}{P(X > VaR_\alpha(X))} \\ &= \frac{P(VaR_\alpha(X) < X \leq x)}{P(X > VaR_\alpha(X))} \end{aligned}$$

luego

$$F_{X|X > VaR_\alpha(X)}(x) = \begin{cases} \frac{F_X(x) - F_X(VaR_\alpha(X))}{1 - F_X(VaR_\alpha(X))} = \frac{F_X(x) - \alpha}{1 - \alpha} & \text{si } x \geq VaR_\alpha(X) \\ 0 & \text{si } x < VaR_\alpha(X). \end{cases}$$

Por lo que

$$\begin{aligned} \int_{-\infty}^{\infty} x dF_{X|X > VaR_\alpha(X)}(x) &= \frac{1}{1 - \alpha} \int_{VaR_\alpha(X)}^{\infty} x dF_X \\ &= \frac{E[X \cdot 1_{\{X \geq VaR_\alpha(X)\}}]}{1 - \alpha}. \end{aligned}$$

Entonces,

$$TCE_\alpha(X) = \frac{E[X \cdot 1_{\{X \geq VaR_\alpha(X)\}}]}{1 - \alpha}.$$

Antes de derivar las propiedades del TCE_α , demostraremos el siguiente lema que nos será de utilidad más adelante.

Lema 1.2.1. Sean $X, Y \in \mathcal{M}$ con esperanza finita y tales que $X \leq Y$ P -c.s., entonces $E[X] \leq E[Y]$.

Demostración. Primero, recordemos que para una variable aleatoria X con esperanza finita, se tiene que

$$E[X] = \int_0^{\infty} [1 - F_X(x)] dx - \int_{-\infty}^0 F_X(x) dx,$$

la identidad anterior puede ser consultada en [15] página 120. Ahora, como $X \leq Y$ $P - c.s.$, por el Lema 1.0.1 tenemos que $F_Y(x) \leq F_X(x)$ para todo $x \in \mathbb{R}$, pero esto también significa que $1 - F_X(x) \leq 1 - F_Y(x)$ para todo $x \in \mathbb{R}$.

Esto implica que

$$\int_{-\infty}^0 F_Y(x)dx \leq \int_{-\infty}^0 F_X(x)dx$$

y

$$\int_0^{\infty} [1 - F_X(x)]dx \leq \int_0^{\infty} [1 - F_Y(x)]dx.$$

De estas dos desigualdades se sigue que

$$\int_0^{\infty} [1 - F_X(x)]dx - \int_{-\infty}^0 F_X(x)dx \leq \int_0^{\infty} [1 - F_Y(x)]dx - \int_{-\infty}^0 F_Y(x)dx.$$

Esto es

$$E[X] \leq E[Y].$$

□

Proposición 1.2.1. *El TCE_{α} para cualquier nivel de confianza $\alpha \in (0, 1)$ satisface las siguientes propiedades [1]:*

- a) *Monotonidad*
- b) *Invariante ante traslaciones*
- c) *Homogeneidad positiva*
- d) *Subaditividad*
- e) *Invarianza en ley*

Demostración. Sea $\alpha \in (0, 1)$

- a) Si $X \leq Y$ $P - c.s.$, entonces $X \cdot 1_{\{X \geq VaR_{\alpha}(X)\}} \leq Y \cdot 1_{\{Y \geq VaR_{\alpha}(Y)\}}$ $P - c.s.$ y luego $E[X \cdot 1_{\{X \geq VaR_{\alpha}(X)\}}] \leq E[Y \cdot 1_{\{Y \geq VaR_{\alpha}(Y)\}}]$.

Por lo tanto

$$\frac{E[X \cdot 1_{\{X \geq VaR_{\alpha}(X)\}}]}{1 - \alpha} \leq \frac{E[Y \cdot 1_{\{Y \geq VaR_{\alpha}(Y)\}}]}{1 - \alpha}$$

$$TCE(X) \leq TCE(Y).$$

b) Sea $Y = X + m$ con $m \in \mathbb{R}$, entonces:

$$\begin{aligned} F_Y(x) &= P(Y \leq x) \\ &= P(X + m \leq x) \\ &= P(X \leq x - m) \\ &= F_X(x - m). \end{aligned}$$

Por lo que

$$F_{Y|Y > VaR_\alpha(Y)}(x) = \begin{cases} \frac{F_X(x-m)-\alpha}{1-\alpha} & \text{si } x - m \geq VaR_\alpha(X) \\ 0 & \text{si } x - m < VaR_\alpha(X). \end{cases}$$

Así

$$\begin{aligned} TCE(X + m) &= TCE(Y) \\ &= \int_{-\infty}^{\infty} x dF_{Y|Y > VaR_\alpha(Y)}(x) \\ &= \int_{-\infty}^{\infty} x dF_{X|X > VaR_\alpha(X)}(x - m) \\ &= \int_{-\infty}^{\infty} (z + m) dF_{X|X > VaR_\alpha(X)}(z) \\ &= \int_{-\infty}^{\infty} z dF_{X|X > VaR_\alpha(X)}(z) + \int_{-\infty}^{\infty} m dF_{X|X > VaR_\alpha(X)}(z) \\ &= \int_{-\infty}^{\infty} z dF_{X|X > VaR_\alpha(X)}(z) + m \int_{-\infty}^{\infty} dF_{X|X > VaR_\alpha(X)}(z) \\ &= TCE(X) + m. \end{aligned}$$

c) Sea $Y = \lambda X$ con $\lambda > 0$, entonces:

$$\begin{aligned} F_Y(x) &= P(Y \leq x) \\ &= P(\lambda X \leq x) \\ &= P(X \leq x/\lambda) \\ &= F_X(x/\lambda). \end{aligned}$$

Por lo que

$$F_{Y|Y > VaR_\alpha(Y)}(x) = \begin{cases} \frac{F_X(x/\lambda)-\alpha}{1-\alpha} & \text{si } x/\lambda \geq VaR_\alpha(X) \\ 0 & \text{si } x/\lambda < VaR_\alpha(X). \end{cases}$$

Así

$$\begin{aligned}
TCE(\lambda X) &= TCE(Y) \\
&= \int_{-\infty}^{\infty} x dF_{Y|Y > VaR_{\alpha}(Y)}(x) \\
&= \int_{-\infty}^{\infty} x dF_{X|X > VaR_{\alpha}(X)}(x/\lambda) \\
&= \int_{-\infty}^{\infty} (z \cdot \lambda) dF_{X|X > VaR_{\alpha}(X)}(z) \\
&= \lambda \int_{-\infty}^{\infty} z dF_{X|X > VaR_{\alpha}(X)}(z) \\
&= \lambda TCE(X).
\end{aligned}$$

d) Sean $X_1, X_2 \in \mathcal{M}$ y sea $X_3 = X_1 + X_2$, notemos que

$$X_1(1_{\{X_1 > VaR_{\alpha}(X)\}} - 1_{\{X_3 > VaR_{\alpha}(X_3)\}}) \geq VaR_{\alpha}(X_1)(1_{\{X_1 > VaR_{\alpha}(X)\}} - 1_{\{X_3 > VaR_{\alpha}(X_3)\}}),$$

pues dado $w \in \Omega$:

Si $X_1(w) > VaR_{\alpha}(X_1)$, entonces $1_{\{X_1 > VaR_{\alpha}(X)\}}(w) = 1$ por lo que

$$(1_{\{X_1 > VaR_{\alpha}(X)\}} - 1_{\{X_3 > VaR_{\alpha}(X_3)\}})(w) \geq 0$$

y luego

$$X_1(w) \cdot (1_{\{X_1 > VaR_{\alpha}(X)\}} - 1_{\{X_3 > VaR_{\alpha}(X_3)\}})(w) \geq VaR_{\alpha}(X_1) \cdot (1_{\{X_1 > VaR_{\alpha}(X)\}} - 1_{\{X_3 > VaR_{\alpha}(X_3)\}})(w).$$

Por otro lado si $X_1(w) \leq VaR_{\alpha}(X_1)$, entonces $1_{\{X_1 > VaR_{\alpha}(X)\}}(w) = 0$ por lo que

$$(1_{\{X_1 > VaR_{\alpha}(X)\}} - 1_{\{X_3 > VaR_{\alpha}(X_3)\}})(w) \leq 0$$

y luego

$$X_1(w) \cdot (1_{\{X_1 > VaR_{\alpha}(X)\}} - 1_{\{X_3 > VaR_{\alpha}(X_3)\}})(w) \geq VaR_{\alpha}(X_1) \cdot (1_{\{X_1 > VaR_{\alpha}(X)\}} - 1_{\{X_3 > VaR_{\alpha}(X_3)\}})(w).$$

Por lo tanto efectivamente

$$X_1(1_{\{X_1 > VaR_{\alpha}(X)\}} - 1_{\{X_3 > VaR_{\alpha}(X_3)\}}) \geq VaR_{\alpha}(X_1)(1_{\{X_1 > VaR_{\alpha}(X)\}} - 1_{\{X_3 > VaR_{\alpha}(X_3)\}}).$$

Y análogamente

$$X_2(1_{\{X_2 > VaR_{\alpha}(X)\}} - 1_{\{X_3 > VaR_{\alpha}(X_3)\}}) \geq VaR_{\alpha}(X_2)(1_{\{X_2 > VaR_{\alpha}(X)\}} - 1_{\{X_3 > VaR_{\alpha}(X_3)\}}).$$

Así,

$$(1 - \alpha)(TCE(X_1) + TCE(X_2) - TCE(X_3)) =$$

$$\begin{aligned}
&= E[X_1 \cdot 1_{\{X_1 \geq VaR_\alpha(X_1)\}}] + E[X_2 \cdot 1_{\{X_2 \geq VaR_\alpha(X_2)\}}] - E[X_3 \cdot 1_{\{X_3 \geq VaR_\alpha(X_3)\}}] \\
&= E[X_1 \cdot 1_{\{X_1 \geq VaR_\alpha(X_1)\}} + X_2 \cdot 1_{\{X_2 \geq VaR_\alpha(X_2)\}} - X_3 \cdot 1_{\{X_3 \geq VaR_\alpha(X_3)\}}] \\
&= E[X_1 \cdot 1_{\{X_1 \geq VaR_\alpha(X_1)\}} + X_2 \cdot 1_{\{X_2 \geq VaR_\alpha(X_2)\}} - (X_1 + X_2) \cdot 1_{\{X_3 \geq VaR_\alpha(X_3)\}}] \\
&= E[X_1 \cdot (1_{\{X_1 \geq VaR_\alpha(X_1)\}} - 1_{\{X_3 \geq VaR_\alpha(X_3)\}}) + \\
&\quad X_2 \cdot (1_{\{X_2 \geq VaR_\alpha(X_2)\}} - 1_{\{X_3 \geq VaR_\alpha(X_3)\}})] \\
&= E[X_1 \cdot (1_{\{X_1 \geq VaR_\alpha(X_1)\}} - 1_{\{X_3 \geq VaR_\alpha(X_3)\}})] + \\
&\quad E[X_2 \cdot (1_{\{X_2 \geq VaR_\alpha(X_2)\}} - 1_{\{X_3 \geq VaR_\alpha(X_3)\}})] \\
&\geq E[VaR_\alpha(X_1) \cdot (1_{\{X_1 \geq VaR_\alpha(X_1)\}} - 1_{\{X_3 \geq VaR_\alpha(X_3)\}})] + \\
&\quad E[VaR_\alpha(X_2) \cdot (1_{\{X_2 \geq VaR_\alpha(X_2)\}} - 1_{\{X_3 \geq VaR_\alpha(X_3)\}})] \\
&= VaR_\alpha(X_1)E[1_{\{X_1 \geq VaR_\alpha(X_1)\}} - 1_{\{X_3 \geq VaR_\alpha(X_3)\}}] + \\
&\quad VaR_\alpha(X_2)E[1_{\{X_2 \geq VaR_\alpha(X_2)\}} - 1_{\{X_3 \geq VaR_\alpha(X_3)\}}] \\
&= VaR_\alpha(X_1)(E[1_{\{X_1 \geq VaR_\alpha(X_1)\}}] - E[1_{\{X_3 \geq VaR_\alpha(X_3)\}}]) + \\
&\quad VaR_\alpha(X_2)(E[1_{\{X_2 \geq VaR_\alpha(X_2)\}}] - E[1_{\{X_3 \geq VaR_\alpha(X_3)\}}]) \\
&= VaR_\alpha(X_1)(P(X_1 \geq VaR_\alpha(X_1)) - P(X_3 \geq VaR_\alpha(X_3))) + \\
&\quad VaR_\alpha(X_2)(P(X_2 \geq VaR_\alpha(X_2)) - P(X_3 \geq VaR_\alpha(X_3))) \\
&= VaR_\alpha(X_1)((1 - \alpha) - (1 - \alpha)) + VaR_\alpha(X_2)((1 - \alpha) - (1 - \alpha)) \\
&= VaR_\alpha(X_1) \cdot 0 + VaR_\alpha(X_2) \cdot 0 = 0.
\end{aligned}$$

Por lo tanto $(1 - \alpha)(TCE(X_1) + TCE(X_2) - TCE(X_1 + X_2)) \geq 0$ y así $TCE(X_1 + X_2) \leq TCE(X_1) + TCE(X_2)$.

□

Del mismo modo que para el VaR , también tenemos un método histórico para calcular el TCE_α , para calcularlo, tomamos los rendimientos diarios y se ordenan los rendimientos de manera creciente $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ y tomar la observación (k) tal que $X_{(k)} = VaR_\alpha$. Entonces, el TCE_α se calcula haciendo una media muestral de los valores que exceden el VaR . Este método tiene la ventaja de que no hay que asumir nada acerca de la distribución de la variable X , aún así, existen otros métodos que pueden verse en [17].

Capítulo 2

Series de Tiempo

En este capítulo se hará el desarrollo de series de tiempo y se presentarán los modelos de series de tiempo con los que se trabajará.

Definición 2.0.1. *Una serie de tiempo es un proceso estocástico $(X_t)_{t \in \mathbb{Z}_+}$, es decir, una familia de variables aleatorias indexadas por \mathbb{Z}_+ y definidas sobre un espacio de probabilidad (Ω, \mathcal{F}, P) .*

Se sabe que las series de tiempo financieras cumplen una serie de propiedades que facilita su manejo y son conocidas como hechos estilizados, las cuales son una colección de observaciones empíricas e inferencias extraídas de estas observaciones, que parecen aplicarse a la mayoría de series diarias de cambios en los factores de riesgo, como rendimientos logarítmicos de acciones, índices o tasas de cambio. Una versión para los hechos estilizados es la siguiente [8]:

Definición 2.0.2. *Los hechos estilizados se definen como las siguientes condiciones:*

- a) *La volatilidad parece variar con el tiempo.*
- b) *Los rendimientos extremos aparecen en grupos.*
- c) *Las series de rendimientos muestran poca correlación serial sin embargo no son independientes.*
- d) *Las series de rendimientos son leptocúrticas o de cola pesada.*

Esta lista de enunciados se toman como hipótesis que se deben satisfacer para aplicar los modelos de series de tiempo que más adelante mencionaremos. Más adelante en el Capítulo 3 evidenciaremos que los datos con los que se trabajará cumplen esta serie de enunciados.

2.1. Modelo GARCH

Definición 2.1.1. *La serie de tiempo $(X_t)_{t \in \mathbb{Z}_+}$ es estrictamente estacionaria si $(X_{t_1}, \dots, X_{t_n})$ tiene la misma distribución que $(X_{t_1+k}, \dots, X_{t_n+k})$ para todo $t_1, \dots, t_n, k \in \mathbb{Z}_+$ y para todo $n \in \mathbb{N}$.*

Estamos interesados en crear un modelo para calcular la volatilidad de un activo. Para ello primero definamos el rendimiento: Si p_t es el precio de un activo en el día t , definimos el log-rendimiento al día t como $X_t = \ln\left(\frac{p_t}{p_{t-1}}\right)$.

Usualmente la volatilidad se define como la desviación estándar de los rendimientos sin embargo, en este modelo se piensa como un proceso estocástico $(\sigma_t)_{t \in \mathbb{Z}_+}$ que se ajusta a través del tiempo como veremos adelante.

Definición 2.1.2. *La serie de tiempo $(X_t)_{t \in \mathbb{Z}_+}$ es un proceso IID($0, \sigma^2$) si está conformada por variables aleatorias independientes e idénticamente distribuidas con media cero y varianza σ^2 .*

Definición 2.1.3. *Sea $(X_t)_{t \in \mathbb{Z}_+}$ un proceso IID($0, \sigma^2$). El proceso $(X_t)_{t \in \mathbb{Z}_+}$ es un proceso GARCH(p, q) si es estrictamente estacionario y si satisface que, para todo $t \in \mathbb{Z}_+$ y para algún proceso de valores positivos $(\sigma_t)_{t \in \mathbb{Z}_+}$, se cumplen las siguientes propiedades:*

$$\begin{aligned} X_t &= \sigma_t Z_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \end{aligned} \quad (2.1)$$

donde $\alpha_0 > 0$, $\alpha_i \geq 0$ para $i = 1, \dots, p$ y $\beta_j \geq 0$ para $j = 1, \dots, q$ [19].

En particular, en este trabajo se usará el modelo GARCH(1,1) pues en la práctica este tiende a ser el más usado y es suficiente para obtener buenos resultados [8]. Este modelo tiene la siguiente forma:

$$\begin{aligned} X_t &= \sigma_t Z_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \end{aligned}$$

donde $\alpha_0 > 0$, $\alpha_1 \geq 0$ y $\beta_1 \geq 0$.

Lo que nos dice este modelo es que el rendimiento de un activo puede verse como el producto de la volatilidad en ese día por un factor de aleatoriedad en ese mismo día, mientras que, la volatilidad en ese día se describe a través del rendimiento de un día anterior y la volatilidad del día anterior.

Dada una serie de observaciones x_0, x_1, \dots, x_n nos interesa estimar los parámetros desconocidos en (2.1) que son $\alpha_0, \alpha_1, \beta_1$, para ello, lo haremos a través del método de máxima verosimilitud, pero antes enunciaremos el siguiente lema que nos será de utilidad.

Lema 2.1.1. *Sean X_0, \dots, X_n variables aleatorias, cada una con función de densidad positiva en los reales, entonces la función de densidad conjunta se puede expresar de la siguiente forma:*

$$f_{X_0, \dots, X_n}(x_0, \dots, x_n) = f_{X_0}(x_0) \prod_{t=1}^n f_{X_t|X_0, \dots, X_{t-1}}(x_t|x_0, \dots, x_{t-1}).$$

Demostración. Para demostrar esto procederemos por inducción:
Para $n = 1$, tenemos que

$$f_{X_1|X_0}(x_1|x_0) = \frac{f_{X_0, X_1}(x_0, x_1)}{f_{X_0}(x_0)}$$

y luego

$$f_{X_0, X_1}(x_0, x_1) = f_{X_0}(x_0) f_{X_1|X_0}(x_1|x_0).$$

Ahora, supongamos que para $k \in \mathbb{N}$ es cierto que

$$f_{X_0, \dots, X_k}(x_0, \dots, x_k) = f_{X_0}(x_0) \prod_{t=1}^k f_{X_t|X_0, \dots, X_{t-1}}(x_t|x_0, \dots, x_{t-1}).$$

Observemos que:

$$f_{X_{k+1}|X_0, \dots, X_k}(x_{k+1}|x_0, \dots, x_k) = \frac{f_{X_0, \dots, X_{k+1}}(x_0, \dots, x_{k+1})}{f_{X_0, \dots, X_k}(x_0, \dots, x_k)}.$$

Usando nuestra hipótesis de inducción llegamos a que:

$$f_{X_{k+1}|X_0, \dots, X_k}(x_{k+1}|x_0, \dots, x_k) = \frac{f_{X_0, \dots, X_{k+1}}(x_0, \dots, x_{k+1})}{f_{X_0}(x_0) \prod_{t=1}^k f_{X_t|X_0, \dots, X_{t-1}}(x_t|x_0, \dots, x_{t-1})}.$$

Finalmente, despejando $f_{X_0, \dots, X_{k+1}}$ llegamos a que

$$f_{X_0, \dots, X_{k+1}}(x_0, \dots, x_{k+1}) = f_{X_0}(x_0) \prod_{t=1}^{k+1} f_{X_t|X_0, \dots, X_{t-1}}(x_t|x_0, \dots, x_{t-1}).$$

□

La aplicación del método de máxima verosimilitud exacta [4] requiere el conocimiento de la distribución de la variable X_0 , que en general desconocemos. Por este motivo, los estimadores se suelen obtener maximizando la función de verosimilitud condicionada a X_0 . Para ello, se seleccionan los valores de α_0 , α_1 y β_1 que maximizan a la función de verosimilitud condicionada a X_0 , la cual es $f_{X_1, \dots, X_n | X_0}(x_1, \dots, x_n | x_0)$ pero observe que por el Lema 2.1.1

$$\begin{aligned} f_{X_1, \dots, X_n | X_0}(x_1, \dots, x_n | x_0) &= \frac{f_{X_0, \dots, X_n}(x_0, \dots, x_n)}{f_{X_0}(x_0)} \\ &= \frac{f_{X_0}(x_0) \prod_{t=1}^n f_{X_t | X_0, \dots, X_{t-1}}(x_t | x_0, \dots, x_{t-1})}{f_{X_0}(x_0)} \\ &= \prod_{t=1}^n f_{X_t | X_0, \dots, X_{t-1}}(x_t | x_0, \dots, x_{t-1}). \end{aligned}$$

Ahora, tomando $1 \leq t \leq n$ tenemos que calcular la función de densidad condicional $f_{X_t | X_0, \dots, X_{t-1}}(x_t | x_0, \dots, x_{t-1})$, pero nótese que X_t , depende de X_0, \dots, X_{t-1} sólo a través de X_{t-1} , o de forma equivalente de σ_t , esto significa que dada la muestra x_0, \dots, x_{t-1} se tiene que σ_t pasa a ser un valor observado y así, la función de densidad a calcular, es la densidad de la variable $\sigma_t Z_t$ con σ_t ahora una constante. Así, usando el Teorema de cambio de variable ([12] página 200) con $g(z_t) = \sigma_t z_t$ y sabiendo que $X_t = \sigma_t Z_t = g(Z_t)$, tenemos:

$$\begin{aligned} f_{X_t | X_0, \dots, X_{t-1}}(x_t | x_0, \dots, x_{t-1}) &= f_{X_t | X_{t-1}}(x_t | x_{t-1}) \\ &= f_Z(g^{-1}(x_t)) \cdot \left| \frac{dg^{-1}(x_t)}{dx_t} \right| \\ &= f_Z\left(\frac{x_t}{\sigma_t}\right) \cdot \left| \frac{d\frac{x_t}{\sigma_t}}{dx_t} \right| \\ &= f_Z\left(\frac{x_t}{\sigma_t}\right) \cdot \frac{1}{\sigma_t}. \end{aligned}$$

En donde f_Z es la densidad del ruido blanco $(Z_t)_{t \in \mathbb{Z}_+}$, que se considera que sigue una distribución normal estándar. De este modo, tenemos que la función de verosimilitud condicional se transforma en:

$$\begin{aligned}
L(\alpha_0, \alpha_1, \beta_1 | X_0) &= f_{X_1, \dots, X_n | X_0}(x_1, \dots, x_n | x_0) \\
&= \prod_{t=1}^n f_{X_t | X_0, \dots, X_{t-1}}(x_t | x_0, \dots, x_{t-1}) \\
&= \prod_{t=1}^n \frac{1}{2\pi} e^{-\frac{1}{2} \cdot \left(\frac{x_t}{\sigma_t}\right)^2} \cdot \frac{1}{\sigma_t} \\
&= \frac{1}{\sqrt{\prod_{t=1}^n (2\pi\sigma_t^2)}} \cdot \exp\left(-\sum_{t=1}^n \frac{x_t^2}{2\sigma_t^2}\right).
\end{aligned}$$

Así, luego podemos obtener la *log* función de verosimilitud condicional

$$\begin{aligned}
\ln L(\alpha_0, \alpha_1, \beta_1 | X_0) &= \frac{1}{\sqrt{\prod_{t=1}^n (2\pi\sigma_t^2)}} \cdot \exp\left(-\sum_{t=1}^n \frac{x_t^2}{2\sigma_t^2}\right) \\
&= -\left(\frac{1}{2} \sum_{t=1}^n \ln(2\pi\sigma_t^2)\right) - \sum_{t=1}^n \frac{x_t^2}{2\sigma_t^2} \\
&= \frac{1}{2} \sum_{t=1}^n \left[-\ln(2\pi) - \ln\sigma_t^2 - \frac{x_t^2}{\sigma_t^2}\right].
\end{aligned}$$

Ahora, recordemos que $\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2$, por lo que el estimador de máxima verosimilitud $\hat{\theta} = (\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1)$ será:

$$\hat{\theta} = \arg \max_{\alpha_0, \alpha_1, \beta_1} \frac{1}{2} \sum_{t=1}^n \left[-\ln(2\pi) - \ln(\alpha_0 + \alpha_1 x_{t-1}^2 + \beta_1 \sigma_{t-1}^2) - \frac{x_t^2}{\alpha_0 + \alpha_1 x_{t-1}^2 + \beta_1 \sigma_{t-1}^2}\right].$$

Observación 2.1.1. *Ahora tenemos un inconveniente, pues el valor de σ_0^2 en realidad no se observa, esto puede ser resuelto eligiendo un valor inicial, como la varianza muestral de x_1, \dots, x_n , o incluso simplemente cero [8].*

2.2. Modelo HAR-RV

A continuación, se presenta el modelo HAR-RV. Los datos de alta frecuencia contienen información de transacciones que pasan en el día, por lo que se puede utilizar para medir la volatilidad. Es importante notar que entre más datos conozcamos por día, nuestras aproximaciones sobre la volatilidad serán más exactas.

Definición 2.2.1. *La volatilidad realizada (RV) es la suma de los cuadrados de la rendimiento intradía de los activos financieros, que se puede expresar como:*

$$RV_t = r_{t2}^2 + r_{t3}^2 + \cdots + r_{tm}^2 = \sum_{i=2}^m r_{ti}^2,$$

en donde m es el tamaño de observaciones al día.

Notemos, de la definición anterior que el tamaño de muestra m dependerá de las actualizaciones de los precios que tengamos al día, por ejemplo, si tenemos actualizaciones cada 5 minutos, entonces nuestra muestra diaria de datos será de 78 observaciones, este criterio se utilizará en lo subsecuente.

Vamos a trabajar con el modelo **HAR-RV** (Heterogeneous Autoregressive-Realized Volatility) que propone escribir a la volatilidad realizada del día $t + 1$, RV_{t+1}^d como una función lineal de las volatilidades realizadas diaria RV_t^d , semanal RV_t^w y mensual RV_t^m del día t y se puede expresar como

$$RV_{t+1}^d = b_0 + b_1 RV_t^d + b_2 RV_t^w + b_3 RV_t^m. \quad (2.2)$$

Las constantes b_0, b_1, b_2, b_3 son valores reales que representan la importancia de las volatilidades realizadas en los distintos periodos y RV_t^d, RV_t^w, RV_t^m son valores que calcularemos a partir de la muestra de datos y se calculan como:

$$RV_t^d = \sum_{i=1}^{78} r_{ti}$$

$$RV_t^w = \frac{1}{5} \sum_{k=0}^4 RV_{t-k}^d$$

$$RV_t^m = \frac{1}{20} \sum_{k=0}^{19} RV_{t-k}^d.$$

Recordemos que cada r_{ti} corresponde al rendimiento de la i -ésima actualización del día t y que se calcula como $r_{ti} = \ln(P_{ti}) - \ln(P_{t(i-1)})$ con $i = 2, \dots, 78$ (para $i = 1$ es habitual tomar el rendimiento $r_{t1} = 0$), donde P_{ti} es el precio del activo en el periodo i del día t .

Notemos que de las variables en (2.2), nos hace falta conocer los coeficientes b_0, b_1, b_2 y b_3 . Para calcular tales valores tenemos que aplicar algún método, el más habitual es del mínimas desviaciones absolutas (**LAD**, por sus siglas en inglés), que consiste en encontrar los valores óptimos de b_0, b_1, b_2, b_3

tales que, para dada una muestra de $n \geq 21$ observaciones de RV^d , esto es, $RV_1^d, RV_2^d, \dots, RV_n^d$ podamos minimizar la expresión:

$$\sum_{t=21}^n |RV_{t+1}^d - b_0 - b_1RV_t^d - b_2RV_t^w - b_3RV_t^m|.$$

Además de este método, podemos usar algunos otro como el de mínimos cuadrados ponderados (WLS) o mínimos cuadrados ordinarios (OLS), que pueden verse en [5].

Capítulo 3

Aplicación al IPC y S&P500

En esta sección veremos el cálculo de las medidas de riesgo y la aplicación de los modelos de volatilidad a datos de los índices *S&P500* y del *IPC*.

Se tomaron los precios de cierre del S&P500 con una muestra de 128 observaciones que comprenden el periodo del 08 de Mayo de 2020 y el 06 de Noviembre de 2020 y lo mismo para el IPC. Para esta parte del trabajo los datos fueron extraídos del sitio Yahoo Finance [22].

Únicamente para la parte del modelo HAR-RV tomamos una muestra con 9984 observaciones que comprenden el periodo del 08 de Mayo de 2020 al 06 de Noviembre de 2020, con actualizaciones de 5 minutos, y así tener 78 observaciones por día (esto por la necesidad del tipo de datos para el modelo HAR-RV). Los datos fueron extraídos del sitio Stooq [20].

La razón del porque elegir este periodo de tiempo están en que nos interesa conocer el comportamiento de los mercados americano y mexicano durante los primeros meses de desarrollo de la pandemia ocasionada por el SARS-CoV-2 [11].

3.1. Hechos Estilizados

A continuación veremos algunos resultados para evidenciar el cumplimiento de los hechos estilizados mencionados en la definición 2.0.2 del Capítulo 2. Primero, observamos las gráficas de precios en la Figura 3.1 entre S&P500 e IPC y se puede notar que S&P500 sigue una tendencia alcista en este periodo de tiempo, mientras el IPC varía en su tendencia.

Ahora bien, en las gráficas de rendimientos de la Figura 3.2 podemos ver

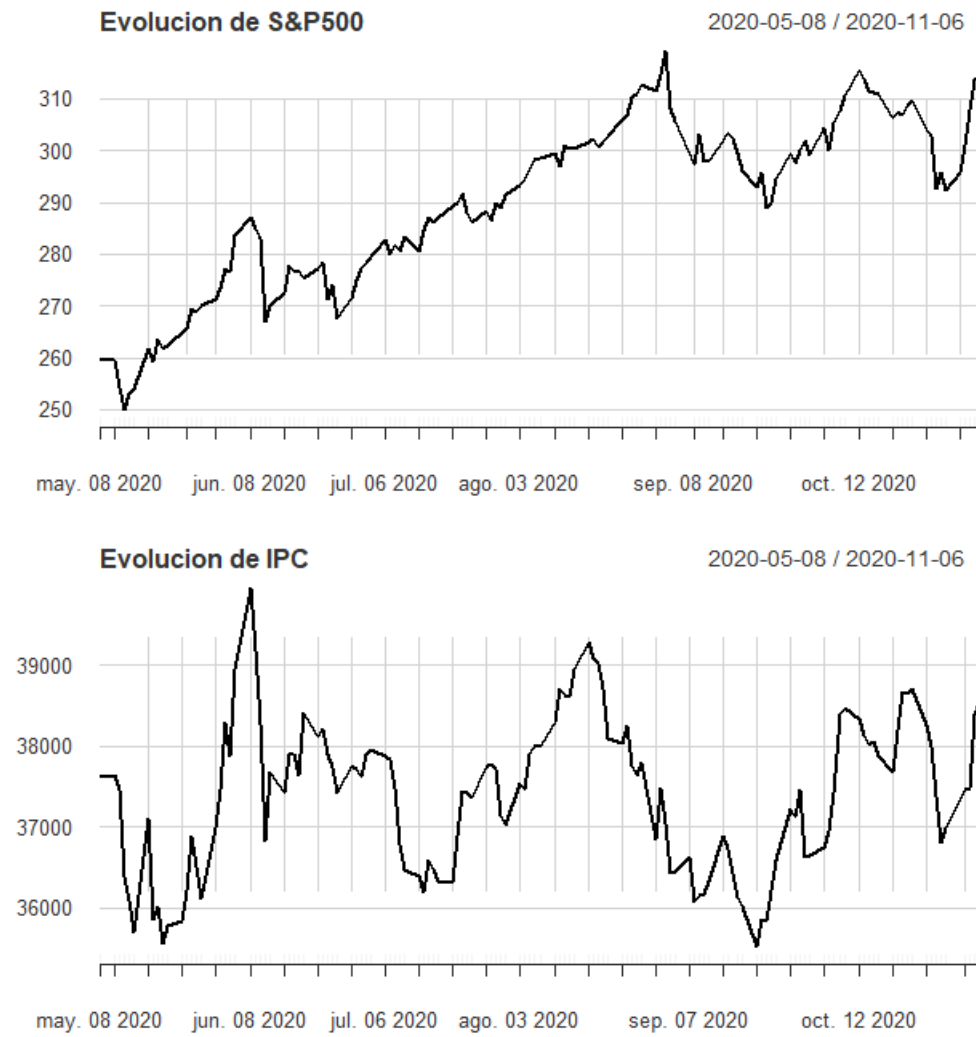


Figura 3.1: Evolución de los precios.

que ambos tienen datos poco frecuentes.

En las gráficas anteriores de los rendimientos podemos observar que la volatilidad parece variar en el tiempo, pues durante los meses de mayo y junio podemos observar más valores extremos, mientras que en julio y agosto los valores no varían en un rango demasiado grande, esto aplica para ambas series y esto evidencia el primer hecho estilizado acerca de que la volatilidad parece variar en el tiempo.

Además, en la misma Figura 3.2 se puede observar la existencia de clusters de volatilidad, pues como se nota, rendimientos extremos vienen seguidos de rendimientos extremos de igual o contrario signo y lo mismo pasa para rendimientos pequeños pues vienen seguidos de rendimientos pequeños de igual o contrario signo, esto respalda que se cumpla el segundo hecho estilizado sobre que los rendimientos extremos aparecen en grupos. Además esto también indica que las variables no son independientes, pues el valor de una afecta a la siguiente.

A continuación observemos las siguientes gráficas de la Figura 3.3 de las funciones de autocorrelación de la serie de rendimientos. Esto respalda el tercer hecho estilizado acerca de la poca correlación serial de las series de rendimientos.

Además, al hacer los cálculos tenemos que la serie de rendimientos del IPC tiene curtosis igual a 0.6311219 mientras que para el S&P500 es de 3.15239, en ambos casos se tiene curtosis mayor a cero lo cual indica que la distribución de rendimientos es leptocúrtica o de cola pesada, esto verifica el cumplimiento último hecho estilizado y lo que esto nos indica es que la probabilidad de obtener un resultado o valor extremo es mayor que la que se encontraría en una distribución normal de resultados.

Con esto hemos verificado que efectivamente las series con las que estamos trabajando cumplen los hechos estilizados y podemos aplicar los modelos de volatilidad que desarrollamos en el Capítulo 2.

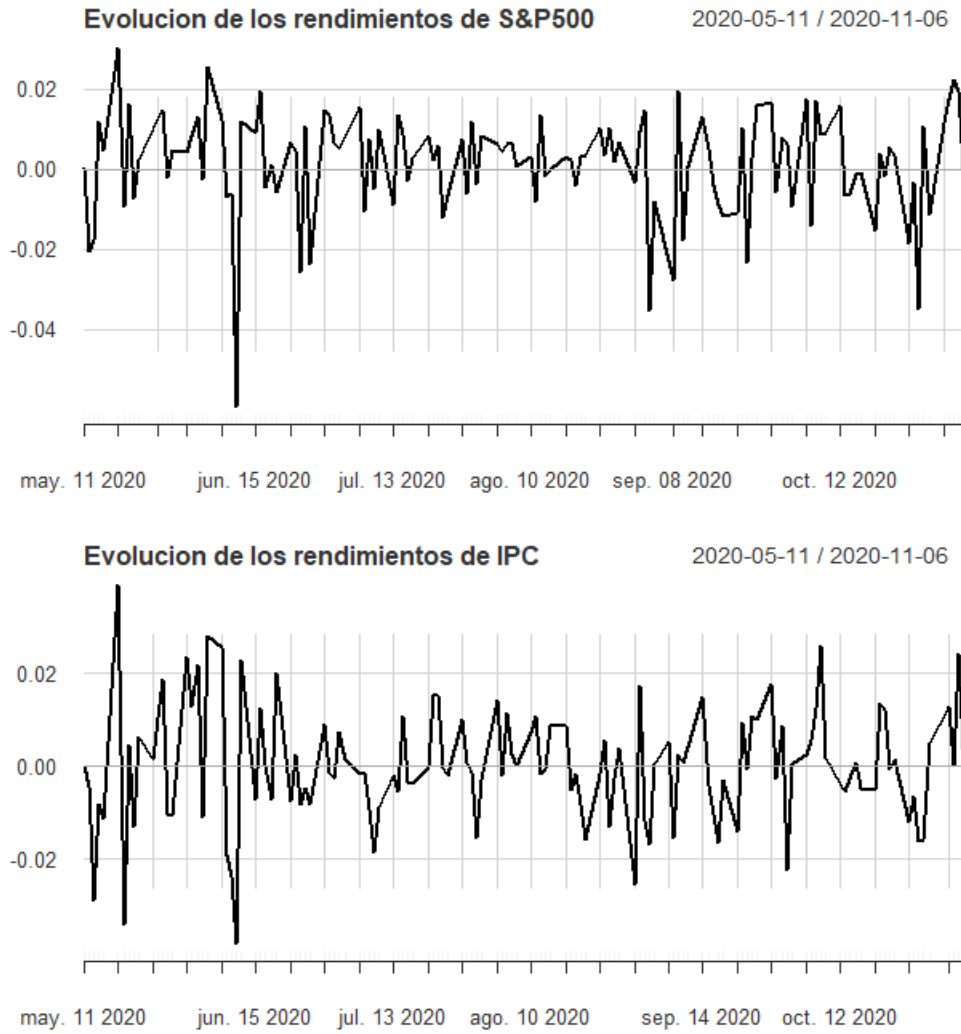


Figura 3.2: Evolución de los rendimientos.

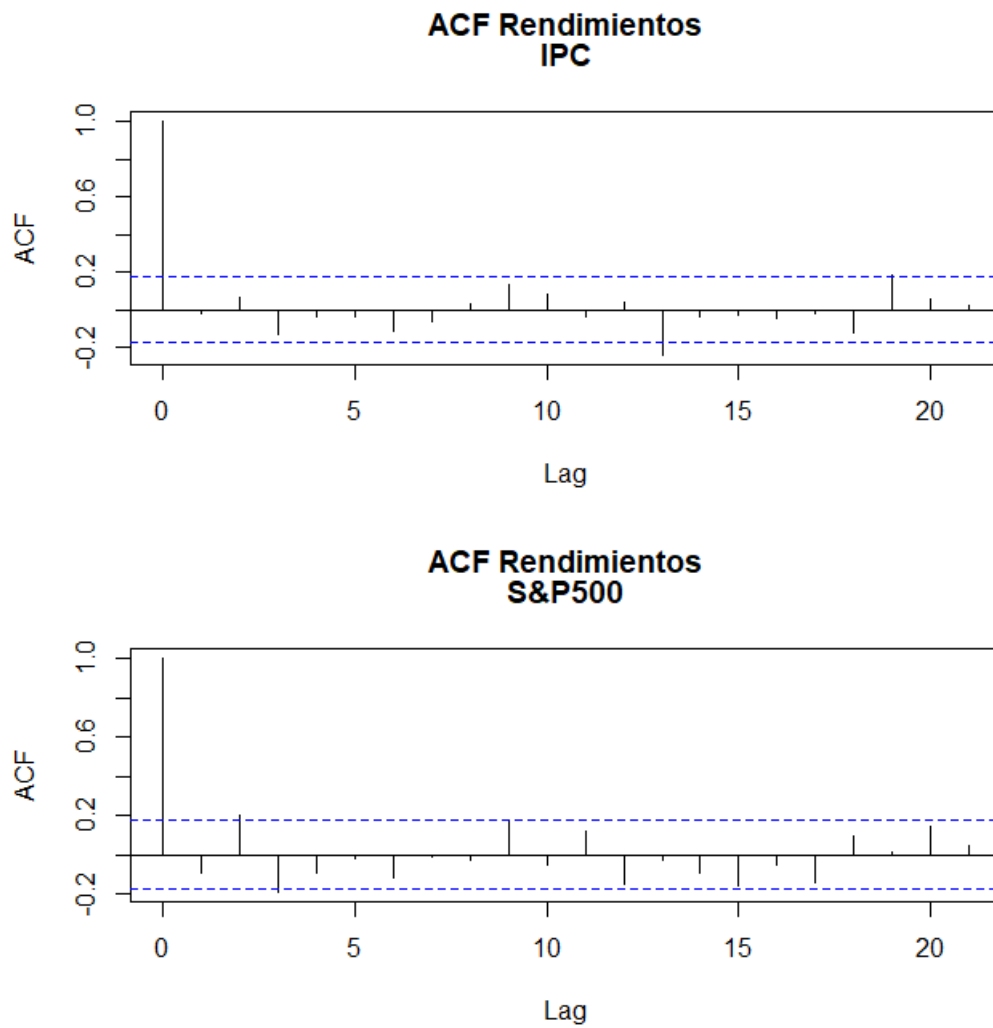


Figura 3.3: Autocorrelación de rendimientos.

3.2. Resultados

Calculamos el VaR sobre el IPC y S&P500 usando el método histórico que describimos en el Capítulo 1, tomando como variable de pérdidas a los log-rendimientos diarios y $\alpha = 0.01$, obtuvimos lo siguiente.

Tras obtener los valores del VaR en ambos casos, en la Figura 3.4 podemos compararlos en los histogramas de rendimientos. Observando los resultados y comparando los histogramas de los rendimientos de S&P500 e IPC, podemos notar que el VaR de IPC (-0.02928265) es menor que en S&P500 (-0.0287529), esto significa que de cada 100 días, sólo en uno podríamos esperar una pérdida mayor a 0.029 en el caso del IPC, y 0.028 en el caso de S&P500. Con esto podemos ver que invertir en el mercado americano es ligeramente un poco más seguro que invertir en el mercado mexicano si queremos evitar cierto riesgo o al menos eso nos indica el periodo de tiempo que estamos considerando.

A continuación vamos a calcular el TCE con el mismo valor $\alpha = 0.01$ para saber que pasa una vez que hemos pasado el *VaR*. En las Figuras 3.5 y 3.6 presentamos los histogramas de rendimientos junto con los valores VaR, TCE (con $\alpha = 0.01$) y la esperanza en cada caso. Observando los resultados, podemos notar que el VaR de IPC es ligeramente mayor que S&P500 y que en ambos casos los valores del TCE son muy cercanos a las medias.

Ahora, podríamos preguntarnos ¿De qué nos sirve conocer al TCE al mismo nivel de confianza que el VaR?, la respuesta está en que, conocer únicamente el VaR puede darnos una falsa sensación de seguridad si no es muy bajo, pues podríamos pensar que eso significa que podríamos esperar rendimientos muy altos, y esto no tiene porque ser así, pues incluso con un VaR muy cercano a cero (pero no mayor), puede que todos los rendimientos sean negativos, es decir, pérdidas. Saber el TCE nos ayudará a saber si aún teniendo las peores pérdidas podemos esperar un rendimiento que nos ayude a compensar esa pérdida.

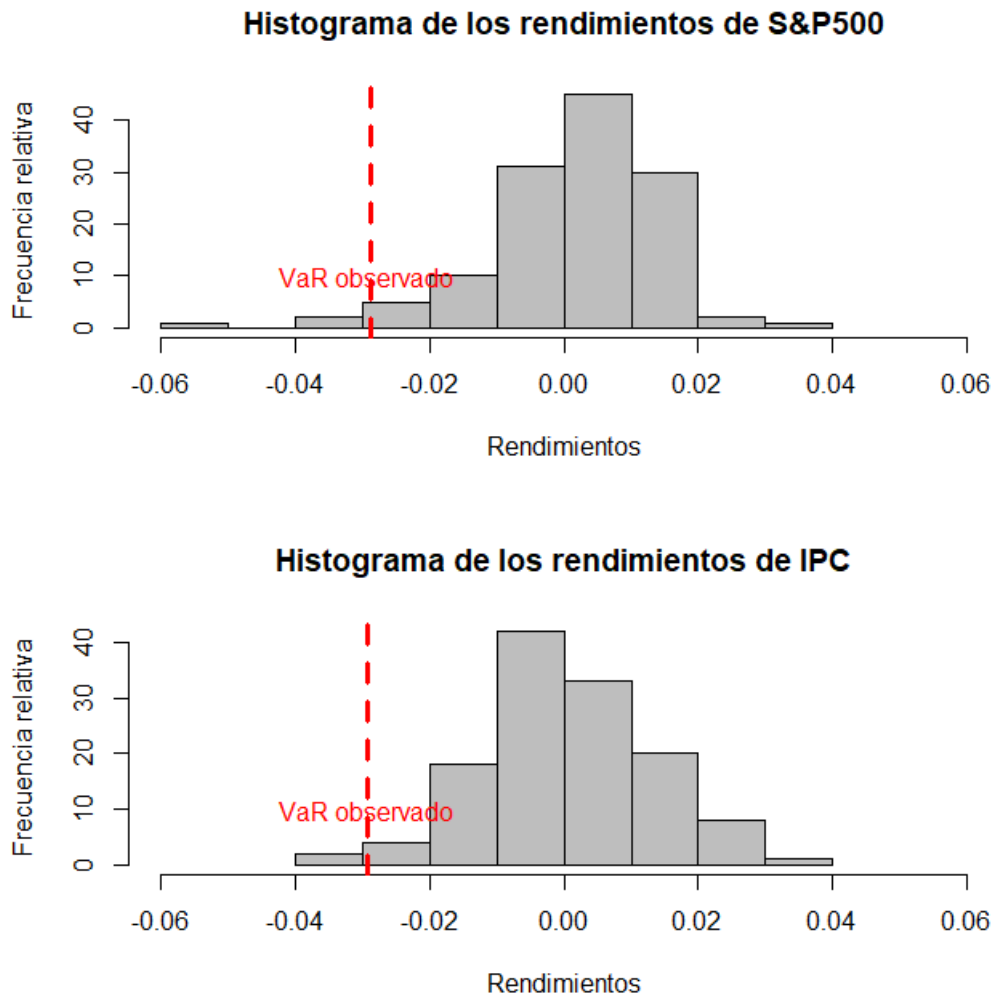
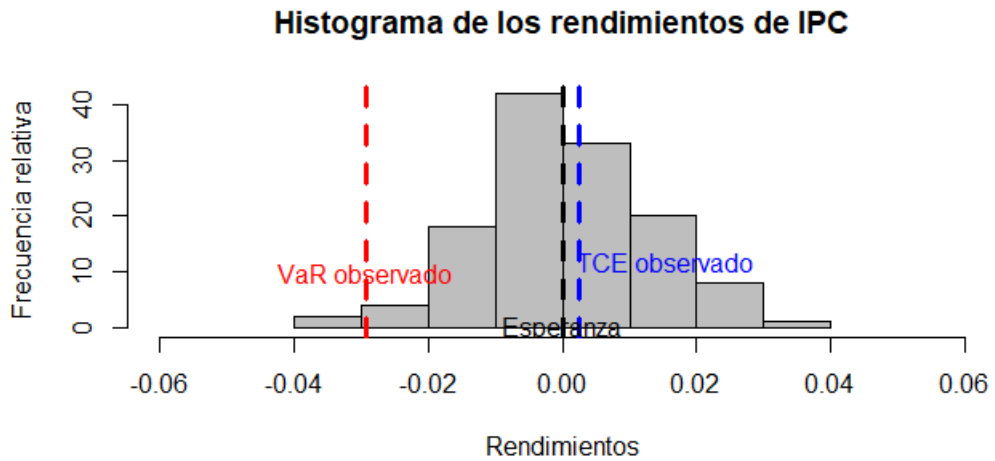
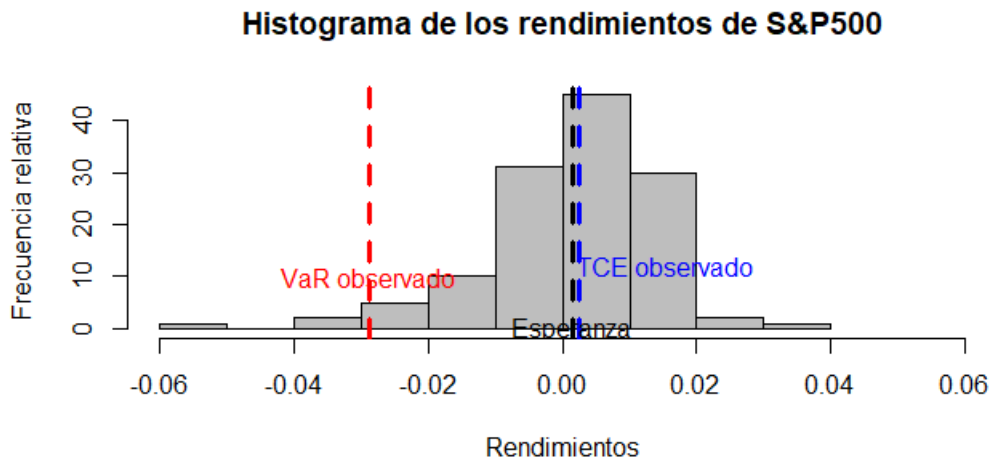


Figura 3.4: Histograma de rendimientos y VaR.



$$VaR_{0.01} = -0.02928265, TCE_{0.01} = 0.002271311, \mu = 0.001490711.$$

Figura 3.5: TCE, VaR y esperanza del IPC.



$$VaR_{0.01} = -0.02875287, TCE_{0.01} = 0.002572188, \mu = 0.0001860664.$$

Figura 3.6: TCE, VaR y esperanza del S&P500.

Ahora, recordamos que contamos con 9984 observaciones de los precios del activo S&P500 con actualizaciones de 5 minutos, y así tener 78 observaciones por día, lo cual nos permite tener una muestra de 128 observaciones de RV^d . Para calcular los coeficientes de b_0, b_1, b_2, b_3 haremos uso de la función **HAREstimate** de la paquetería **HARMModel** de R (véase Anexo A), de donde obtuvimos como resultado que los valores para los coeficiente deben ser lo que pueden verse en la Figura 3.7. Así, nuestro modelo para la

```

Coefficients:
      Estimate Std. Error t value Pr(>|t|)
beta0  4.836e-05  4.111e-05   1.176   0.243
beta1  5.768e-01  1.169e-01   4.933 4.65e-06 ***
beta5  3.178e-03  1.818e-01   0.017   0.986
beta20 1.174e-01  2.483e-01   0.473   0.638
---

```

Figura 3.7: Valores de los coeficientes.

volatilidad se convierte en:

$$RV_{t+1}^d = 0.00004836 + 0.5768RV_t^d + 0.003178RV_t^w + 0.1174RV_t^m. \quad (3.1)$$

Tomando las 128 observaciones de RV^d , para cada día $t + 1$ con $t \geq 20$ calculamos RV_{t+1}^d con los datos reales de la muestra, es decir con los datos de los 20 días anteriores. Dicha información es ilustrada en la Figura 3.8 en la cual podemos comparar los valores estimados contra los datos reales.

En la anterior, los puntos rojos representan la volatilidad real, mientras que los puntos azules representan la volatilidad calculada con el modelo a partir de los datos reales de los días anteriores. Como puede observarse, los valores estimados no están muy alejados de los reales, lo cual es un buen indicador de que el modelo funciona para predecir la volatilidad del día siguiente.

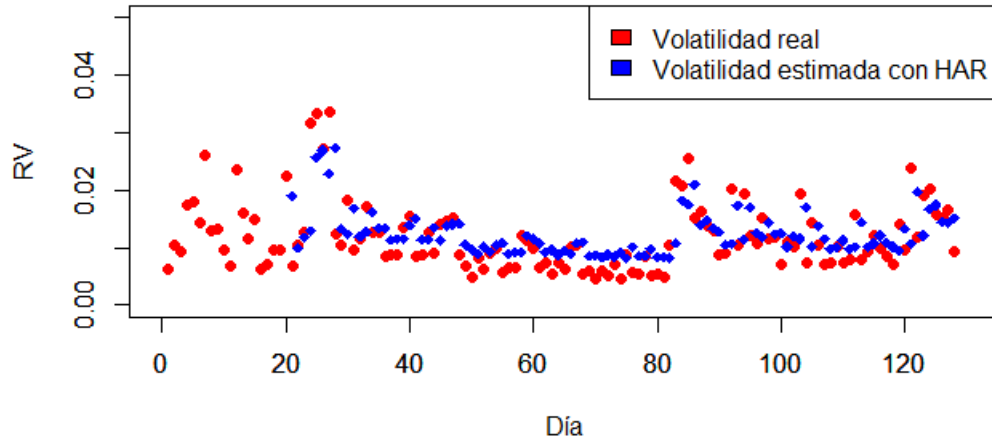


Figura 3.8: Volatilidad real y estimada con HAR.

Ahora, para el modelo GARCH y del mismo modo poder crear un modelo para predecir la volatilidad, tenemos que encontrar los valores de los coeficientes α_0 , α_1 y β_1 , para ello, usaremos la función `garchFit` de la paquetería `fGarch` de R (véase Anexo A).

Nótese que como mencionamos en el Capítulo 2 hemos hecho el supuesto de que $X_0 = 0 = \sigma_0$.

Para el caso del S&P500 los valores para los coeficientes son $\alpha_0 = 1.956706 * 10^{-05}$, $\alpha_1 = 1.383778 * 10^{-01}$ y $\beta_1 = 7.409681 * 10^{-01}$. Así, la ecuación para predecir la volatilidad queda expresada como:

$$\sigma_t^2 = 1.956706 * 10^{-05} + 1.383778 * 10^{-01} X_{t-1}^2 + 7.409681 * 10^{-01} \sigma_{t-1}^2.$$

Tomando las 128 observaciones de los rendimientos, para cada día t con $t \geq 1$ calculamos σ_t con los datos reales de la muestra. Dicha información se representa en la gráfica de la Figura 3.9. En la anterior, los puntos rojos representan la volatilidad real, mientras que los puntos verdes representan la volatilidad calculada con el modelo a partir de los datos reales de los días anteriores. Como puede observarse, los valores estimados no difieren en gran medida de los valores reales, y esto nos indica que el modelo funciona en buena parte para predecir la volatilidad del día siguiente.

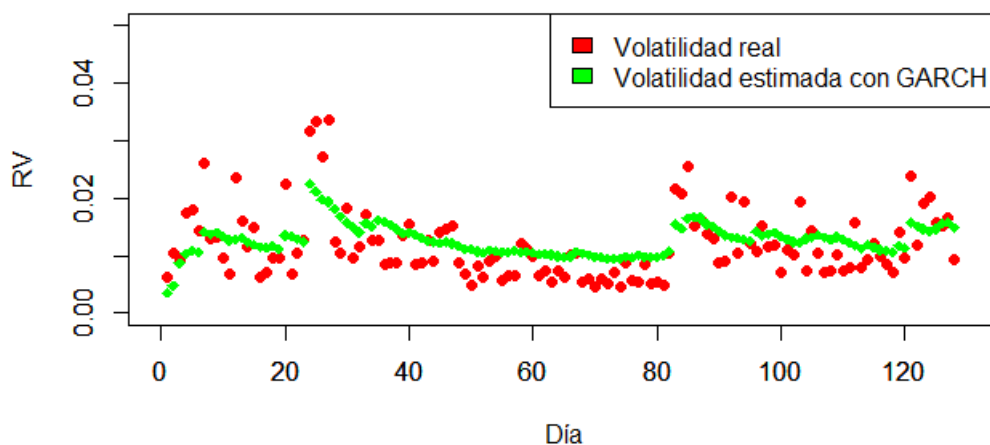


Figura 3.9: Volatilidad real y estimada con GARCH.

Ahora, podemos ver la comparación de los resultados del modelo HAR-RV y del modelo GARCH en la Figura 3.10.

Para comparar que modelo es mejor haremos uso del error cuadrático medio (MSE) que mide el promedio de los errores al cuadrado, la manera de calcularlo es la siguiente:

$$MSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{\sigma}_i - \sigma_i)^2}.$$

En donde n es el tamaño de la muestra de datos a comparar, $\hat{\sigma}_i$ es la volatilidad predecida y σ_i corresponde a la volatilidad real.

Con esto tenemos que el MSE para el modelo HAR-RV es 0.005044121, mientras que para el modelo GARCH es 0.004487477, por lo que el modelo GARCH hace mejores estimaciones para la volatilidad.

Ahora, veamos como el predecir la volatilidad nos puede ayudar a complementar nuestra parte de las medidas de riesgo, Para ello, recordemos que lo que nos dice la volatilidad es, que tan dispersos podríamos obtener nuestros rendimientos, Sabemos que el VaR nos da un valor en el que dada una frecuencia podemos encontrar los rendimientos inferiores a este, sin embargo, no nos dice que tan pequeños podrían llegar a ser esos rendimientos, para esto nos sirve conocer la volatilidad, pues podríamos establecer una cota inferior para los peores rendimientos que podríamos obtener y así establecer si

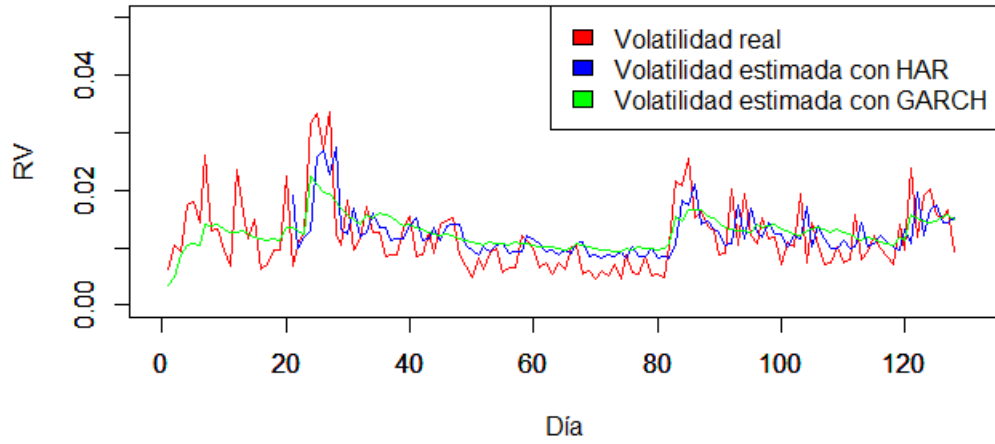


Figura 3.10: Volatilidad real y estimada con GARCH y HAR.

estamos en una posición segura conservando un activo, o si por el contrario la posición es muy riesgosa y sea mejor retirarse.

Conclusiones

Como conclusión acerca de nuestro ejemplo de aplicación pudimos ver que de los cálculos tanto del VaR como el TCE, pudimos observar que en el mercado americano se corren menos riesgos a la hora de invertir, sin embargo, esto también significa que se podrían esperar una menor ganancia, por lo que si se quiere maximizar la posibilidad de obtener ganancia, se debería invertir en el mercado mexicano. Por tanto, lo ideal sería tener una cartera equilibrada con inversión en el mercado americano para no exponer esa cantidad a pérdida grande, y también tener inversión en el mercado mexicano si se quiere tener una mayor posibilidad de obtener ganancias asumiendo un mayor riesgo. Cabe mencionar que esto concluimos de los resultados obtenidos a partir de la base de datos considerada, pues si tomáramos otro intervalo de tiempo esto podría cambiar.

Ahora, a manera de conclusión general tenemos que a lo largo de estas páginas hemos observado como las medidas de riesgo pueden cuantificar de una buena manera cierta característica de interés acerca del riesgo, aunque, nos hemos limitado a la aplicación de la teoría de medidas de riesgo a activos financieros, en realidad esta es una teoría muy amplia que se puede aplicar a otros contextos como seguros, pero esto queda fuera del alcance de este trabajo. Del mismo modo, para los modelos de volatilidad los hemos aplicado al contexto de finanzas, sin embargo no negamos la posibilidad de poder ajustar estos modelos a cualquier fenómeno que involucre una varianza no constante a través del tiempo.

Como hemos visto, usar los modelos de volatilidad estocástica puede complementar de buena manera la parte de teoría de riesgo, pues podemos cuantificar de mejor manera las peores pérdidas al poder acotarlas en un intervalo y tener un punto de equilibrio que nos permite saber si aún teniendo las peores pérdidas aún podemos tener rendimientos que compensen estas.

A partir de lo anterior, se espera como trabajo futuro estudiar algunas otras medidas de riesgo o algún otro modelo de volatilidad que nos permita extraer más información de un activo o no sólo enfocarnos en los peores

escenarios como lo hemos hecho en este trabajo sino ahora pensar en las posibilidades para los mejores escenarios.

Bibliografía

- [1] Carlo Acerbi y Dirk Tasche. “On the Coherence of Expected Shortfall”. En: *Journal of Banking & Finance* 26 (2002), págs. 1487-1503.
- [2] Carol Alexander. *Market risk analysis, practical financial econometrics*. John Wiley & Sons, 2008.
- [3] Philippe Artzner y col. “Coherent measures of risk”. En: *Mathematical finance* 9.3 (1999), págs. 203-228.
- [4] George Casella y Roger L Berger. *Statistical inference*. Cengage Learning, 2021.
- [5] Adam Clements y Daniel PA Preve. “A practical guide to harnessing the HAR volatility model”. En: *Journal of Banking & Finance* 133 (2021), págs. 1-16.
- [6] Fulvio Corsi. “A simple approximate long-memory model of realized volatility”. En: *Journal of Financial Econometrics* 7.2 (2009), págs. 174-196.
- [7] Hans Föllmer y Alexander Schied. “Convex measures of risk and trading constraints”. En: *Finance and stochastics* 6.4 (2002), págs. 429-447.
- [8] Christian Francq y Jean-Michel Zakoian. *GARCH models: structure, statistical inference and financial applications*. John Wiley & Sons, 2019.
- [9] Zinoviy Landsman y Emiliano A Valdez. “Tail conditional expectations for exponential dispersion models”. En: *ASTIN Bulletin: The Journal of the IAA* 35.1 (2005), págs. 189-209.
- [10] Alexander J McNeil, Rüdiger Frey y Paul Embrechts. *Quantitative risk management: concepts, techniques and tools-revised edition*. Princeton university press, 2015.
- [11] Ricardo J Mendoza-Rivera, José Antonio Lozano-Díez y Francisco Venegas-Martínez. “Impacto de la pandemia Covid-19 en variables financieras relevantes en las principales economías de Latinoamérica”. En: *Economía: teoría y práctica* SPE5 (2020), págs. 125-144.

- [12] Alexander McFarlane Mood. *Introduction to the Theory of Statistics*. McGraw-hill, 1950.
- [13] Abdelhakim Necir, Abdelaziz Rassoul y Ričardas Zitikis. “Estimating the conditional tail expectation in the case of heavy-tailed losses”. En: *Journal of Probability and Statistics* (2010).
- [14] Svetlozar T Rachev, Stoyan V Stoyanov y Frank J Fabozzi. *A probability metrics approach to financial risk measures*. John Wiley & Sons, 2011.
- [15] Luis Rincon Solís. *Curso Intermedio de Probabilidad*. Universidad Nacional Autónoma de México, 2010.
- [16] Sheldon Ross. *A first course in probability*. Pearson, 2010.
- [17] Ana Maite Sesma. *Expected Shortfall: Modelo de investigación de aplicación práctica y metodologías de Backtesting bajo FRTB*. Tesis de Maestría en la Universidad Complutense de Madrid, 2016.
- [18] Albert N Shiryaev. *Probability-1*. Vol. 95. Springer, 2016.
- [19] Robert H Shumway y David S Stoffer. *Time series: a data analysis approach using R*. Chapman y Hall/CRC, 2019.
- [20] “Stooq”. En: (22 de Junio de 2022). URL: <https://stooq.com/db/h/>.
- [21] Francisco Venegas-Martínez y col. *Riesgos financieros y económicos*. Universidad de Medellín, 2008.
- [22] “Yahoo Finance”. En: (22 de Junio de 2022). URL: <https://finance.yahoo.com/lookup?s=DATA>.
- [23] Rongjun Yang y col. “Big data analytics for financial Market volatility forecast based on support vector machine”. En: *International Journal of Information Management* 50 (2020), págs. 452-462.

Anexo A. Códigos en R

Código usado para el cálculo del VaR

```
library(quantmod)
library(PerformanceAnalytics)

#Obtenemos los datos con los que vamos a trabajar
getSymbols("^MXX", from="2020-05-08", to="2020-11-07")

getSymbols("V00", from="2020-05-08", to="2020-11-07")

# Extraemos los precios de cierre
V00 <- V00[, "V00.Adjusted"]

IPC <- MXX[, "MXX.Adjusted"]

#Graficas de evolucion de los precios
chart.TimeSeries(V00, element.color = "blue",
                 main="Evolucion del etf V00")

chart.TimeSeries(IPC, element.color = "blue",
                 main="Evolucion del IPC")

#Obtenemos los lo-rendimientos
V00.r <- diff(log(V00))
V00.r<- V00.r[-1] #Se quita el dato inicial porque es vacío

IPC.r <- diff(log(IPC))
IPC.r<- IPC.r[-1] #Se quita el dato inicial porque es vacío
```

```
#Graficas de la evolucion de los rendimientos
chart.TimeSeries(V00.r,
                 main="Evolucion de los rendimientos de V00")

chart.TimeSeries(IPC.r,
                 main="Evolucion de los rendimientos de IPC")

#Calculamos el VaR con el metodo histórico y un nivel de confianza 99
V00.Var <- VaR(V00.r, p=.99, method="historical")

IPC.Var <- VaR(IPC.r, p=.99, method="historical")

#Histogramas de los rendimientos la linea roja nos indica en VAR
hist(V00.r, main="Histograma de los rendimientos de S&P500",
     xlim=c(-0.06,0.06),
     col="grey",
     xlab="Rendimientos",
     ylab="Frecuencia relativa")
abline(v=V00.Var, col="red", lwd=3, lty=2)

hist(IPC.r, main="Histograma de los rendimientos de IPC",
     xlim=c(-0.06,0.06),
     col="grey",
     xlab="Rendimientos",
     ylab="Frecuencia relativa")
abline(v=IPC.Var, col="red", lwd=3, lty=2)
```

Código usado para el cálculo del TCE

```
#Librerias a usar
library(quantmod)
library(PerformanceAnalytics)

#Obtenemos los datos con los que vamos a trabajar
getSymbols("^MXX",src = "yahoo", from="2020-05-08", to="2020-11-07")
getSymbols("V00", from="2020-05-08", to="2020-11-07")

# Extraemos los precios de cierre
V00 <- V00[, "V00.Adjusted"]
IPC <- MXX[, "MXX.Adjusted"]

#Graficas de evolucion de los precios
chart.TimeSeries(V00, element.color = "blue", main="Evolucion del etf V00")
chart.TimeSeries(IPC, element.color = "blue", main="Evolucion del IPC")

#Obtenemos los log-rendimientos
V00.r <- diff(log(V00))
V00.r<- V00.r[-1] #Se quita el dato inicial porque es vacío

IPC.r <- diff(log(IPC))
IPC.r<- IPC.r[-1] #Se quita el dato inicial porque es vacío

#Graficas de la evolucion de los rendimientos
chart.TimeSeries(V00.r, main="Evolucion de los rendimientos de V00")
chart.TimeSeries(IPC.r, main="Evolucion de los rendimientos de IPC")

#Calculamos el VaR con \alpha=0.01
V00.var<-VaR(V00.r, p=.99, method="historical")
IPC.var<-VaR(IPC.r, p=.99, method="historical")

#Calculamos el TCE con \alpha=0.01
V00.tce=mean(V00.r[V00.r>V00.var[1,1]])
IPC.tce=mean(V00.r[V00.r>IPC.var[1,1]])

#Histogramas de los rendimientos la linea roja nos indica en VAR
hist(V00.r, main="Histograma de los rendimientos de S&P500",
      xlim=c(-0.06,0.06),
```

```
        col="grey",
        xlab="Rendimientos",
        ylab="Frecuencia relativa")
abline(v=V00.var, col="red", lwd=3, lty=2)
text(V00.var,10, "VaR observado", col="red")
abline(v=V00.tce, col="blue", lwd=3, lty=2)
text(V00.tce,12, "                TCE observado", col="blue")
abline(v=mean(V00.r), col="black", lwd=3, lty=2)
text(mean(V00.r), 0, "Esperanza", col="black")

hist(IPC.r, main="Histograma de los rendimientos de IPC",
      xlim=c(-0.06,0.06),
      col="grey",
      xlab="Rendimientos",
      ylab="Frecuencia relativa")
abline(v=IPC.var, col="red", lwd=3, lty=2)
text(IPC.var,10, "VaR observado", col="red")
abline(v=IPC.tce, col="blue", lwd=3, lty=2)
text(IPC.tce,12, "                TCE observado", col="blue")
abline(v=mean(IPC.r), col="black", lwd=3, lty=2)
text(mean(IPC.r), 0, "Esperanza", col="black")
```

Código usado para el modelo GARCH

```

#librerías a usar
library(quantmod)
library(PerformanceAnalytics)
library(fGarch)
#Obtenemos los datos con los que vamos a trabajar
getSymbols("^MXX", from="2020-05-08", to="2020-11-07")
getSymbols("V00", from="2020-05-08", to="2020-11-07")
# Extraemos los precios de cierre
V00 <- V00[, "V00.Adjusted"]
IPC <- MXX[, "MXX.Adjusted"]

datos=as.ts(V00) #Acá definimos con que datos vamos a trabajar
nombre="S&P500"

#función que genera gráficas para describir la serie
niveles<- datos
#par(mfrow=c(2,3))
plot.ts(niveles, main=c("Niveles", nombre), col="blue")
rendimientos<- diff(log(niveles))
plot.ts(rendimientos, main=c("Rendimientos", nombre), col="blue")
plot.ts((rendimientos)^2, main=c("Rendimientos cuadrados", nombre), col="blue")
acf(rendimientos, main=c("ACF Rendimientos", nombre))
acf((rendimientos)^2, main=c("ACF Rendimientos cuadrados", nombre))
acf(abs(rendimientos), main=c("ACF Rendimientos absolutos", nombre))

dist='std'
modelo<- garchFit(formula=~garch(1,1), data=rendimientos,
                  cond.dist=dist, include.mean=FALSE)
#cond.dist puede ser std o norm
estimados<- modelo@fit$par[1:3] #Coeficientes para la ecuacion

vr=rep(1:length(rendimientos)+1)

vr[1]<- estimados[1]
for(i in 2:128 ) {
  sc<- estimados[1]+estimados[2]*((rendimientos[i-1])^2)+

```

```
    estimados[3]*(vr[i-1])^2  
    vr[i]<- sc  
  }
```

```
pred<- estimados[1]+estimados[2]*  
((rendimientos[length(rendimientos)])^2)+estimados[3]*vr  
sqrt(pred)
```

Código usado para el modelo HAR

```

library(readr) #Nos permite leer la base de datos (Archico CSV)
library(HARModel) #Para obtener el Modelo HAR-RV

Libro1 <- read_csv("RUTA") #Ruta de los datos descargados de Stooq
OC=as.vector(Libro1[,c(5)]) #Extraemos la parte que nos interesa
OC2=rbind(OC,0)
OC1=rbind(0,OC)

LogRet=log(OC2[1])-log(OC1[1]) #Obtener los rendimientos intradía
LogRet[1,]=0

RVd=rep(0,128)
RVd=cbind(RVd)
for (i in 1:128) {
  a=i*78-77
  b=i*78
  RVd[i,1]=sum((LogRet[c(a:b),1])**2)
}#Este ciclo nos permite calcular la RV de cada día

VOOHAR = HAREstimate(RM = RVd[1:100], periods = c(1,5,20)) #Modelo HAR-RV

RVw=rep(0,128)
RVw=cbind(RVw)

for (i in 5:128) {
  a=i-4
  b=i
  RVw[i,1]=(sum(RVd[c(a:b),1]))/5
}
RVm=rep(0,128)
RVm=cbind(RVm)
for (i in 20:128) {
  a=i-19
  b=i
  RVm[i,1]=(sum(RVd[c(a:b),1]))/20
}

```

```
}
```

```
MRVd =function(t){  
0.00004836+0.5768*RVd[t]+0.003178*RVw[t] + 0.1174*RVm[t]  
}
```

```
N=rep(1:128)  
Res=RVd[1:100]  
for (i in 101:128) {  
  ad=i-1  
  aw=i-5  
  bw=i-1  
  am=i-20  
  bm=i-1  
  Res[i]=0.00004836+0.5768*Res[ad]+0.003178*(sum(Res[aw:bw])/5) +  
    0.1174*(sum(Res[am:bm])/20)  
}
```

```
NN=rep(2:128)  
plot(N,sqrt(RVd),xlim=c(0,130),ylim=c(.0,.05),col='red',pch=16,  
      xlab='Día', ylab='RV')  
par(new=T)  
plot(N,sqrt(vr) ,xlim=c(0,130),ylim=c(.0,.05),col='green',pch=18,  
      xlab=' ', ylab=' ')
```