



**UNIVERSIDAD  
SERGIO ARBOLEDA**

**Uso de cópulas bivariadas para el modelado de índices de renta  
variable en el período 2015-2020**

**Carlos Felipe Bedoya Riveros**

Universidad Sergio Arboleda  
Escuela de Ciencias Exactas e Ingeniería, Departamento de Matemáticas Bogotá,  
Colombia

2021

**Uso de cópulas bivariadas para el modelado de índices de renta variable en el período 2015-2020**

**Carlos Felipe Bedoya Riveros**

Trabajo de grado presentado como requisito para optar al título de: **Magister en Matemáticas Aplicadas**

Director:  
MSc. Alejandro Bello

Universidad Sergio Arboleda  
Escuela de Ciencias Exactas e Ingeniería, Departamento de Matemáticas Bogotá,  
Colombia  
2021

# Índice

<b>1. Resumen</b>	<b>5</b>	
<b>2. Introducción</b>	<b>6</b>	
<b>3. Justificación</b>	<b>8</b>	
<b>4. Hipótesis</b>	<b>9</b>	
<b>5. Objetivos</b>	<b>10</b>	
5.1. Objetivo general .....	10	
5.2. Objetivos específicos .....	10	
<b>6. Marco teórico</b>	<b>11</b>	
<b>7. Metodología</b>	<b>15</b>	
7.1. Cópulas .....	15	
7.2. Modelo GARCH.....	20	
7.3. Coeficientes de correlación .....	22	
<b>8. Análisis de los resultados</b>	<b>25</b>	
<b>9. Conclusiones</b>	<b>29</b>	
<b>10. Anexos</b>		<b>30</b>
10.1. Más sobre los índices utilizados .....	30	
10.2. Gráficas .....	31	
10.3. Código utilizado y resultados del código .....	33	

10.3.1. Co'odigo .....	33
10.3.2. Inspecci'ón del modelo ARMA .....	35
10.3.3. Prueba de estacionariedad .....	36
10.3.4. Resultados del la co'úpula acorde con el orden ARMA-GARCH ..	36
10.3.5. Coeficiente de Pearson .....	37
10.3.6. Sobre el modelo ARMA-GARCH: vector de innovaciones .....	38

## 1. Resumen

Los mercados de valores de Estados Unidos, Alemania y Jap'ón resultan ser fundamentales en el mercado de capitales moderno, dado que estas economías se encuentran entre las cinco más grandes del planeta y cuentan con sectores financieros vigorosos en cuanto a sus niveles de capitalización b'ursatil<sup>1</sup>. La estimación de distribuciones marginales v'ia un modelo ARMA-GARCH y el uso de co'úpulas t-student permiten concluir que existe un fuerte v'inculo entre los índices representativos de tales mercados de renta variable, siendo este un resultado totalmente coherente con el grado de globalización que existe en el mercado de capitales moderno.

## Palabras Clave

Co'úpulas, GARCH, ARMA,  $\rho$  de Spearman, dependencia estructural, S&P500, DAX, Nikkei.

## Abstract

The American, German and Japanese Stock markets are fundamental in the modern capital market given the fact that these economies are in the top five of the largest economies of the world and that they possess a vigorous financial sector. The estimation of univariate marginal distributions through an ARMA-GARCH model and the application of t-student copulas allows to conclude that there is a strong bond between the representative indices of those markets. This result is completely coherent with the

---

<sup>1</sup> Basado en cifras del Banco Mundial, puntualmente, en cifras del Producto Interno Bruto y del Índice de Capitalización B'ursatil

degree at which the modern capital markets is globalized and integrated as a single entity.

## 2. Introducción

La medición adecuada de la correlación entre dos variables aleatorias es un ejercicio totalmente necesario, dado que permite entender de manera precisa la relación existente entre las mismas, si esta existe. Este tipo de análisis se puede aplicar, por ejemplo, a estudios de diversificación de un portafolio de activos, entre otros.

La correlación entre variables aleatorias se suele estimar mediante procedimientos sencillos como la estimación del coeficiente de correlación de Pearson, el cual presenta limitantes en cuanto a su aplicación<sup>2</sup> que pueden ser superadas por otros coeficientes alternativos de correlación, por ejemplo, el coeficiente de Spearman.

El cálculo de coeficientes de correlación está asociado a las *cóputas*, las cuales son funciones multivariadas. En esencia, estas permiten transformar un grupo de distribuciones univariadas en un modelo multivariado.<sup>3</sup> Adicionalmente, las cóputas determinan unívocamente a coeficientes de correlación como el de Spearman y el de Kendall. Es por este proceso de unir variables distintas en un modelo conjunto, y por la relación existente con respecto a diversos coeficientes de correlación, que las cóputas se asocian con estudios de dependencia estructural<sup>4</sup>. Esto es, ofrecen una manera robusta de medir la correlación existente entre variables aleatorias, superando varias de las limitantes de los coeficientes “clásicos” como el de Pearson.

Las cóputas son funciones cuyo origen se asocian más a estudios de áreas de matemática como Probabilidad y Análisis que a aplicaciones específicas dentro del sector financiero. Sin embargo, es una metodología que se ha ido aplicando de manera ostensible durante los últimos años. Es un área versátil tanto en términos de teoría como de aplicaciones. Las cóputas son, en términos formales, distribuciones multivariadas acumuladas que operan sobre variables aleatorias que deben transformarse previamente a distribuciones uniformes en  $(0,1)$ . Por tanto, las cóputas no solo permiten cuantificar aspectos específicos como la correlación entre variables, en sí mismas modelan el vínculo existente entre tal grupo de variables, tal y como lo

---

<sup>2</sup> En la sección de *Metodología* se discuten algunas de ellas.

<sup>3</sup> De hecho, la palabra “Cópula” proviene del término “Couple”, cuyo significado en este contexto hace alusión a la unión de elementos.

<sup>4</sup> Note que el cálculo de correlación no determina necesariamente vínculos de causalidad. Por tanto, para hablar en propiedad de dependencia estructural se suele complementar lo previamente mencionado con otras técnicas estadísticas y con análisis teórico propio de cada contexto.

hace cualquier distribución multivariada. En parte, es esta dualidad la que la convierte en una metodología fértil. Su aplicación es virtualmente posible a cualquier campo o fenómeno que involucre el estudio de variables aleatorias.

La metodología basada en el uso de cópulas impone dos tareas preponderantes: la estimación adecuada de modelos univariados y la transformación de los mismos en un único modelo multivariado, en donde la cópula como función resulta ser fundamental.

El presente trabajo mide el coeficiente de Spearman entre los índices S&P-500, DAX y Nikkei para el período 2015-2020.<sup>5</sup> Para ello se necesita estimar una distribución univariada asociada a cada índice, y una función cópula de la cual se pueda calcular el coeficiente de correlación previamente mencionado. En cuanto a la estimación univariada, se utiliza un modelo ARMA-GARCH. Este modelo es particularmente útil en el manejo de series relacionadas a rendimiento de activos de renta variable (en particular, al observarse una volatilidad cambiante a lo largo de todo el intervalo). Adicionalmente, se calcula la cópula *t*-student que mejor se ajusta a los datos. De lo anterior se sigue el cálculo del coeficiente de correlación de Spearman.

El resultado final varía entre 0.6 y 0.7, aproximadamente, concluyendo así que existe una fuerte correlación positiva entre tales índices. Tal resultado es consistente con trabajos como [3] el cual señala claros síntomas de integración entre diversos e importantes mercados bursátiles. Esto permite evidenciar la existencia de vínculos entre los mercados de países como Estados Unidos, Alemania y Japón.

---

<sup>5</sup> La medición se realiza por pares, por ejemplo, se estima el coeficiente para DAX vs Nikkei.

### **3. Justificación**

En modelación financiera existen varias tareas fundamentales: Primero, definir el conjunto de elementos que se van a estudiar y seleccionar el modelo que se va a utilizar. Luego, realizar un análisis bajo supuesto *Ceteris Paribus*, esto es, estudiar a cada uno de los elementos del modelo como un mundo independiente. Por ejemplo, el valor de mercado de una acción específica, un commodity, un índice bursátil, entre otros. Para esto existe una batería bien conocida de técnicas, estas van desde modelos económicos hasta aplicaciones de redes neuronales. Otra tarea, no menos importante, es entender cómo estos componentes interactúan, cómo el cambio en una de ellas puede generar, o no, variaciones significativas en otra. Esta tarea es crítica, considere el lector que puede llegar a suceder si una entidad, que adquiere un determinado conjunto de activos o riesgos, no tiene en cuenta la dinámica intrínseca entre tales variables.

Asumiendo un contexto no determinístico, el conjunto de componentes del modelo va a ser representado por un conjunto de variables aleatorias. Así mismo, existen varias formas de cuantificar la correlación entre estas, siendo una de las conocidas el coeficiente de Correlación de Pearson. No obstante, esta forma tan sencilla en términos prácticos resulta estar condicionada a varios supuestos que matemáticamente no son triviales, por ejemplo, la existencia de momentos de orden dos. Adicionalmente, este coeficiente presenta limitaciones que serán discutidas más adelante.

Coómulas ofrece una metodología versátil y precisa para cuantificar tal correlación, permitiendo entender de mejor manera la dinámica existente entre un determinado conjunto de variables aleatorias.

Una aplicación específica del análisis de tal dependencia entre variables aleatorias radica en entender qué tan fuerte es el vínculo entre distintos mercados bursátiles, con base en ello es posible establecer un claro panorama respecto al nivel de diversificación de un portafolio dado. Otra aplicación común de cómulas, está en el análisis de riesgo de portafolios de créditos y otros mecanismos similares. En general, tal metodología resulta pertinente en cualquier contexto donde se necesite entender claramente el vínculo existente entre un grupo de variables.

### **4. Hipótesis**

Existe una clara correlación positiva entre los índices accionarios S&P-500, DAX y Nikkei tomando como referencia el período 2015-2020.

## 5. Objetivos

### 5.1. Objetivo general

Cuantificar la correlación existente entre los índices S&P-500, DAX y Nikkei, tomando como referencia el período 2015-2020 por medio del modelo GARCH-Co'pula y el subsecuente cálculo coeficiente  $\rho$  de Spearman.

### Objetivos espec'ficos

### 5.2. Objetivos espec'ficos

- Estimar distribuciones univariadas para cada uno de los índices mediante el uso del modelo GARCH.
- Estimar la distribución multivariada por medio de una cópula  $t$ -student.
- Calcular el coeficiente de correlación de Spearman por medio de la cópula  $t$ -student.



## 6. Marco teórico

Como ya se ha podido evidenciar, el concepto que subyace al presente trabajo es el de *cóputas*. Pero, ¿Cómo surge tal objeto? Este nace con el trabajo de Abe Sklar, cuyas implicaciones serían mucho más poderosas que lo evidenciado en su trabajo original [27]. Este teorema, el cual se enuncia más adelante, establece un vínculo directo entre un conjunto de distribuciones univariadas acumuladas y una distribución multivariada acumulada, cuyos componentes son los componentes univariados previamente mencionados. Este teorema fue inicialmente objeto de discusión desde el punto de vista meramente teórico. Trabajos un poco más recientes como el de [8] ofrecen una prueba de tipo topológico para tal teorema.

Posteriormente, en trabajos como [25] y [31] se descubre y discute el vínculo existente entre las *cóputas* como funciones multivariadas y coeficientes de correlación como el de Spearman. Esto, en conjunción con las evidentes asociadas al coeficiente de Pearson, lleva al uso de *cóputas* como medio para medir de manera más robusta y precisa la correlación entre variables aleatorias. No obstante, para la década de los noventa del siglo XX, las *cóputas* eran algo casi exclusivo del área de la matemática teórica. Es durante el corrido del siglo XXI que se empieza a volver un tema notorio, no solo a nivel de matemática teórica, sino a nivel de aplicaciones puntuales, en particular, dentro del campo de la modelación financiera.

El uso de *cóputas* es bastante amplio, dada su versatilidad y rigurosidad teórica. En general, permite estudiar la interacción entre cualquier  $n$ -tupla de variables aleatorias. Por tanto, su aplicación es potencialmente viable en casi cualquier campo.

En [19] se modela la correlación entre los retornos de portafolios de acciones y se compara con otros modelos. Así mismo, se usan *cóputas* de valores extremos, para así poder modelar escenarios particulares, que si bien no son comunes dentro de los mercados financieros, son relevantes toda vez que es indispensable estar preparado para los mismos.

En [22] se usan *cóputas* de Vine. construidas a partir de *cóputas* t-bivariadas. Esto permite darle un mejor manejo de las colas de la distribución del modelo, esto permite hacer predicciones y comparaciones entre distintos modelos de *cóputa*, con base en máxima verosimilitud y cuantiles extremos.

Esta metodología también resulta particularmente adecuada para modelar el contagio existente entre mercados, es por eso que [5] enfatiza como esta metodología es mucho más adecuada que simples índices como el de Pearson, que si bien es versátil, esta circunscrito a calcular correlación lineal entre variables aleatorias, el ejercicio se

aplica principalmente sobre las principales acciones negociadas en el mercado de Estados Unidos.

Un trabajo similar es realizado en [24], en tal trabajo se propone un modelo cruzado que analiza la correlación entre algunos índices financieros como lo son S&P 500, NASDAQ 100, DAX 30, FTSE 100, and Nikkei 225 y su correspondiente volatilidad. Se propone también un modelo de cópulas dinámicas que capturan ciertos coeficientes de correlación.

En [15] se estudia el comportamiento del mercado de renta variable, se encuentra que la correlación es más persistente entre acciones del mercado europeo. Se usa el modelo GARCH para estimar el modelo univariado. En este trabajo se analiza el comportamiento de cuatro índices bursátiles: S&P-500, DAX, CAC y FTSE.<sup>6</sup>

En [1] se usa un enfoque similar, desde el punto de vista del modelo GARCH-Co'pula, para estudiar el vínculo existente entre el precio del petróleo y el valor de distintas tasas de cambio. Estudio que es muy relevante toda vez que dimensiona correctamente el claro vínculo entre divisas y sector minero-energético. Este vínculo se hace evidente al considerar argumentos de flujo neto de divisas asociadas a tal sector.

---

<sup>6</sup> En la sección de Anexos se hace una breve descripción de cómo se calculan estos índices.

Por otro lado, en [17] se propone un algoritmo iterativo para la estimación de máxima verosimilitud de modelos basados en cópulas. El algoritmo lleva a una descomposición natural de una función de probabilidad que es potencialmente difícil de manejar. Se comparan estos resultados con los obtenidos en un modelo GARCH-Cópula.

En [29] también se modela la correlación entre series financieras, valor de mercado de acciones de Shanghai, mediante una cópula  $t$ -student, llegando a conclusiones similares en cuanto a la existencia de una fuerte correlación positiva entre las mismas.

Esta metodología es fácilmente extendible a conclusiones respecto a la diversificación de un portafolio dado, tal y como se plantea en [30]. En el mismo se plantea una interrogante interesante: ¿cómo se puede identificar el modelo óptimo para pronosticar el riesgo de la cartera a través del análisis en la muestra? Esto se aborda mediante un extenso estudio empírico de 1.500 carteras bivariadas que contienen datos sobre mercado de renta variable, materias primas y futuros de divisas.

El modelo GARCH es en sí mismo bastante versátil, tal y como se puede observar en [4]. En tal trabajo se ajustan doce modelos GARCH a cada criptomoneda, y sus ajustes se evalúan en términos de cinco criterios.

Un trabajo ciertamente similar se realiza en [16]. En este se busca el modelo óptimo de heterocedasticidad condicional con respecto a la bondad de ajuste a los datos de precios de Bitcoin. Se encuentra que el mejor modelo es el modelo AR-CGARCH,<sup>7</sup> una variante teórica del modelo GARCH.

Esta metodología también se suele usar para análisis de corte macroeconómico, como es el caso de [23]. En este se utilizan datos de series de tiempo anuales, se estudia la inflación en Kenia durante período 1960 - 2017 utilizando enfoques ARIMA y GARCH. estables. El mejor modelo se escoge mediante criterios como el AIC, entre otros. Los resultados de este estudio predicen que se espera que la persistencia de la trayectoria inflacionaria se mantenga.

Cabe resaltar que tanto el modelo de cópulas como el modelo GARCH, son ampliamente versátiles. El primero es utilizado en diversas áreas, algunas totalmente lejanas al objeto del presente trabajo. Por ejemplo en hidrología. La razón radica en que la cópula subyacente a un determinado grupo de variables aleatorias existe y es única si los datos provienen de distribuciones continuas. Con base en ello, el cálculo de

---

<sup>7</sup> Conocido como el modelo *Componente GARCH*, diseñado cuantificar de manera más precisa el componente de largo plazo de la volatilidad, una especificación de tal modelo puede hallarse en [11]

coeficientes de correlación es inmediato y fácilmente calculable computacionalmente. Es por ello que no es sorpresa observar diversas aplicaciones de tal metodología.

En [12] se presenta varios métodos que se enfocan en el modelado de estructuras de correlación, incluyendo distribución paramétrica, entropía, cópula y no paramétrica.

Se discuten las fortalezas y limitaciones de los diferentes métodos y vías de investigación futura, como el modelado de la correlación en un clima cambiante, para ayudar a los planificadores y administradores de los recursos hídricos en la selección y aplicación de técnicas adecuadas.

Así mismo, en [2] se muestra una muy amplia gama en cuanto al uso de cópulas para estudios hidrológicos, toda vez que en los mismos es fundamental entender la interacción entre diferentes actores en un contexto dado.

Finalmente, en [7] se establece que la previsión de sequías hidrológicas se considera un componente clave en la gestión de riesgos de los recursos hídricos. Dado que la sequía meteorológica sostenida puede conducir a una sequía hidrológica a lo largo del tiempo, es conceptualmente factible capitalizar la correlación entre las sequías meteorológicas e hidrológicas mientras se intenta pronosticar esta última. Para tales efectos, se usa una cópula paramétrica.

Luego de haber dado una breve muestra del amplio poder de uso de las funciones cópula se procederá entonces a mostrar sus principales rasgos teóricos.

## 7. Metodología

La metodología a seguir en el presente documento consiste en lo siguiente: Primero, se realiza la estimación univariada mediante el modelo ARMA-GARCH. Luego se transforman las variables aleatorias hacia variables uniformemente distribuidas en  $(0,1)$ . Posteriormente se estima el modelo multivariado mediante una cópula  $t$ -student. Finalmente, se calcula el coeficiente de correlación de Spearman toda vez que este coeficiente está determinado por la función cópula subyacente al modelo.

### 7.1. Cópulas

A continuación se mostrarán algunos rasgos fundamentales de una cópula. Para ello, debemos introducir algunos conceptos previos. Los siguientes lemas (y sus pruebas) pueden ser encontrados en varios textos, Por ejemplo en libros avanzados como [26], así como en otros con un enfoque aplicado, como es el caso de [6]. En lecturas como [13], se enuncian también en las siguientes propiedades.

**Teorema 1.** *Probability Transformation.* Sea  $F$  una función de distribución acumulada y sea  $X \sim F$ . Entonces  $F(X) \sim U(0,1)$ .

La transformación anterior, resultará ser bastante útil. El coeficiente de correlación de Pearson capta de manera adecuada la correlación lineal existente entre variables aleatorias. Esto resulta adecuado para variables normalmente distribuidas, variables con distribución uniforme, entre otras. No obstante, acorde con [13], este coeficiente resulta inadecuado para cuantificar la dependencia estructural existente entre variables con otros tipos de distribuciones, por ejemplo, exponenciales. En términos concretos, el coeficiente obtenido puede ser relativamente bajo no porque no exista un fuerte vínculo, sino porque este no es lineal en cuanto a su dinámica se refiere.

Así mismo, la información relativa a la relación entre un par de variables aleatorias y la dependencia entre sus respectivas transformaciones es la misma, por lo que se logra estandarizarlas y no se pierde información significativa respecto a cómo estas interactúan. Recuerde el lector que se está estudiando dicha relación, se asume que ya se conocen bien las formas univariadas, por tanto, la manipulación de estas últimas resulta irrelevante para los propósitos del estudio en cuestión.

**Teorema 2.** *Transformación cuantílica.* Sea  $U \sim U(0,1)$  y  $F$  cualquier distribución acumulada. Entonces  $F^{-1}(U) \sim F$ . Donde  $F^{-1}$  denota la inversa generalizada, esta se define así:

$$F^{-1}(y) = \inf\{x \in \mathbb{R} : F(x) \geq y\}$$

$$y \in [0,1]$$

La anterior transformaci3n tambi3n resulta fundamental. En general, es bastante 3til hallar la funci3n inversa a una aplicaci3n espec3fica. No obstante, para que esto suceda deben cumplirse requisitos estrictos, por ejemplo, biyectividad. La transformaci3n cuant3lica permite hallar una funci3n, que para efectos estad3sticos ofrece las mismas bondades que una funci3n inversa, aun cuando los restrictivos requerimientos mencionados no se cumplan.

Con base en lo anterior, se tiene el teorema de Sklar [27], demostrado en tal art3culo. Por otro lado, algunas interpretaciones interesantes se hacen expl3citas en [13]. Se mencionaran algunas de ellas. La existencia de las transformaciones previamente mencionadas resultan fundamentales para tal teorema.

**Teorema 3.** *Las siguientes proposiciones son verdaderas:*

1. *Para cualquier  $n$ -dimensional distribuci3n acumulada  $H$  con marginales univariadas  $F_1, \dots, F_n$ , existe una c3pula  $n$ -dimensional  $C$ , tal que*

$$H(x) = C(F_1(x_1), \dots, F_n(x_n)), x \in \mathbb{R}^n \quad (1)$$

*La c3pula  $C$  est3a un3vocamente determinada sobre  $\prod_{i=1}^n \text{ran}(F_i)$ , y est3a dada por*

$$C(u) = H(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)), u \in \prod_{i=1}^n \text{ran}(F_i)$$

2. *Rec3procamente, dada una c3pula  $n$ -dimensional  $C$  con distribuciones univariadas  $F_1, \dots, F_n$ , entonces  $H$  definida como en (1) es una funci3n de distribuci3n acumulada con  $F_1, \dots, F_n$  como distribuciones marginales.*

Respecto al teorema (3) se pueden hacer las siguientes observaciones:

- Como se mencion3 anteriormente, una c3pula es un modelo que permite convertir  $n$  variables univariadas en un modelo multivariado.
- Si  $F_i$ ,  $i \in [n]$  es una distribuci3n continua, entonces la c3pula es 3nica, lo que justifica que estad3sticamente se realicen esfuerzos para encontrarla y dar as3 una representaci3n multivariada al conjunto de variables.

- De las cópulas paramétricas se pueden obtener coeficientes que describen de manera cuantitativa la relación entre las variables, sin circunscribirse exclusivamente a relaciones lineales.

En resumen, una cópula es un modelo multivariado cuyas distribuciones marginales son uniformes en  $(0,1)$ . Note que esta elección es totalmente arbitraria. Bien se podría buscar que cada uno de los elementos del vector se distribuya uniformemente en otro intervalo. Incluso, se puede transformar la distribución de cada uno de los vectores a otra distribución, siempre y cuando sean la misma. De esta manera es mucho más coherente el análisis de dependencia entre variables.

En cuanto a los tipos de cópulas existente, se puede hablar de dos grandes familias: paramétricas y no paramétricas. Las del segundo tipo están fuera del objetivo del presente trabajo. Entre las cópulas paramétricas, se encuentran las Elípticas y las Arquimedianas. Las primeras están asociadas al concepto de distribución elíptica, que es una generalización de distribuciones como la normal y la t-student multivariada. Por otro lado, las Arquimedianas son muy populares en su uso dada la capacidad que tienen de llevar a una fórmula explícita, lo cual no ocurre con las elípticas. Finalmente, en algunos casos y por cuestiones de simplicidad, se presentará únicamente el caso bivariado.

Entre las cópulas paramétricas más conocidas se encuentran:

1. *Cópula de Frank*. Sea  $\theta \in \mathbb{R} \setminus \{0\}$ , se define la cópula como

$$C_{\theta}^F(u) = -\frac{1}{\theta} \log \left( 1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right)$$

El parámetro  $\theta$  controla la magnitud y dirección de la relación entre las variables aleatorias. Esto es, entre mayor sea su nivel absoluto, más fuerte es la dependencia entre las dos variables en cuestión. Por otro lado,  $\theta < 0$  indica un vínculo inverso entre ellas.

2. *Cópula de Clayton*. Para  $\theta \in [-1, \infty) \setminus \{0\}$

$$C_{\theta}^C = \max\{u_1^{-\theta} + u_2^{-\theta} - 1, 0\}$$

Se observa que en tanto en la cópula de Frank como en la de Clayton  $\theta \rightarrow 0$  resulta en que la función multivariada tienda a la cópula independiente<sup>8</sup>. Esto es totalmente natural. Piense el lector, a manera de analogía el vínculo que existe

---

<sup>8</sup> La cópula independiente, es la que establece que el modelo multivariado es igual a la productoria de las distribuciones marginales univariadas, algo que debe resultar evidente al observar la palabra "independiente".

entre dos variables independientes y el hecho de que el coeficiente de correlación (Pearson) sea igual a cero.

3. Co'pula Gumbel-Hougaard. Sea  $\theta \in [1, \infty)$ , y  $u \in [0, 1]^n$

$$c_{\theta}^{GH}(u) = e^{\left(-(\sum_{j=1}^n (-\log(u_j))^{\theta})^{\frac{1}{\theta}}\right)}$$

Para el caso bivariado, se presenta una caracter'istica bastante notoria. Las co'pulas como funciones tienen un m'aximo y un m'ınimo, dado el orden usual de los n'umeros reales. Esto es, para toda co'pula  $C$ , 2- dimensional se tiene que

$$W(u) \leq C(u) \leq M(u) \\ u \in [0, 1]^2$$

Donde  $W(u)$  y  $M(u)$  son las co'pulas de Hoeffding y Frchet, respectivamente. Estas se definen como

$$W(u) = \max\{u_1 + u_2 - 2, 0\} \\ M(u) = \min\{u_1, u_2\}$$

Los resultados originales y la versio'n generalizada puede hallarse en [9], [14], [13], [21], entre otros.

Ahora bien, para entender el conceto de co'pula el'ıptica es necesario entender el concepto subyacente, el cual es es el de *distribuci'ın el'ıptica*. Estas se pueden caracterizar mediante la matriz dispersi'ın o escala. Por ejemplo, esta es una matriz definida semipostivia y sim'etrica. No obstante, el rasgo que m'as resalta de estas distribuciones viene de un hecho f'acil de describir desde lo geom'etrico y muy sencillo en cuanto a su conceptualizaci'ın; estas distribuciones multivariadas <sup>9</sup> se caracterizan por el hecho de que al graficar sus curvas de nivel se obtiene una elipse o elipsoide. Piense el lector en los datos que siguen esta distribuci'ın mediante la siguiente analog'ia: considere una nube de polvo que gravita alrededor de un centro gravitacional, en tal caso, el centro es el conjunto de valores esperados mientras que cada dato lo orbita, as'ı como el polvo se mueve alrededor del centro de masa. Ejemplos cl'asicos son la distribuci'ın normal y  $t$ -student bivariadas. Obs'ervese la figura 1 en al secci'ın de anexos.

---

<sup>9</sup> Considere el lector el caso sencillo de dos variables



Como complemento respecto a las cópulas elípticas, considere lo siguiente:

1. Considere un vector  $d$ -dimensional de variables aleatorias tal que se tiene una distribución normal multivariada  $N_d(0,P)$  donde  $0$  representa la matriz de ceros y  $P$  es la matriz de covarianzas, entonces la cópula es

$$C_P^m = \int_{-\infty}^{\Phi^{-1}(u_d)} \cdots \int_{-\infty}^{\Phi^{-1}(u_1)} \frac{e^{-\frac{1}{2}x'P^{-1}x}}{(2\pi)^{\frac{d}{2}} \det(P)^{\frac{1}{2}}} dx_1 \cdots dx_d$$

Donde que  $\Phi^{-1}$  es la transformación cuantílica de la distribución normal.

2. Por otro lado, la  $t$ -cópula es igual a

$$C_{(P,\nu)}^t(u) = \int_{-\infty}^{t_v^{-1}(u_d)} \cdots \int_{-\infty}^{t_v^{-1}(u_1)} \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})(\pi\nu)^{\frac{d}{2}} \sqrt{\det(P)}} \left(1 + \frac{x'P^{-1}x}{\nu}\right)^{-\frac{\nu+d}{2}} dx_1 \cdots dx_d$$

Obsérvese que la expresión analítica resulta bastante robusta. Más allá de ello considere el lector el rol que juega la cópula elíptica. Tómese el caso de distribución normal bivariada cuya expresión es ciertamente comprensible y sencilla. Se sabe también cómo luce la distribución normal multivariada, la cópula es el mecanismo que permite transformar un conjunto de variables que se distribuyen de forma univariada acorde con una expresión ya conocida, y que resultan distribuyéndose de manera multivariada de una forma también conocida y familiar. En ese orden de ideas, la cópula permite la transformación de un conjunto de variable univariadas en un modelo multivariado práctico, útil y sujeto a ser base de toma de decisiones.

Hasta este punto, no se ha tocado el tema de la correlación. ¿Por qué las cópulas permiten caracterizar la dependencia estructural? La respuesta es impactante, los coeficientes usados para medir la correlación entre variables, como el de Spearman y como la  $\tau$  de Kendall resultan estar totalmente determinados por la cópula asociada a las observaciones. Es por ello que determinar la cópula equivale a determinar el valor de coeficientes de correlación mucho más robustos y versátiles que el de Pearson, caracterizando así la dependencia estructural entre variables aleatorias.

En general, las cópulas se caracterizan por lo siguiente:

- Dado que una cópula es una función

$$C : [0,1]^n \Rightarrow [0,1]$$

se debe entonces garantizar que en efecto sea una distribución multivariada. Para ello se formula una noción de medida que dota de estructura a la función. Sea  $x,y$

$\in (0,1)^n$  con  $x_i \leq y_i$  para  $i \in [n]$ . El hiper-rectángulo formado por  $x,y$  se define como  $\{u \in (0,1)^n : x_i < u_i \leq y_i, 1 \leq i \leq n\}$ . Con base en ello se define un  $C$ -volumen:

$$\Delta_{[x,y]}C = \sum_{i \in [0,1]^n} (-1)^{\sum_{j=1}^n i_j} C(x_1 y_1^{i_1-1}, \dots, x_n y_n^{i_n-1})$$

Para el caso de  $n = 2$ , esto se puede sintetizar como:

$$\Delta_{[x,y]}C = C(y_1, y_2) - C(y_1, x_2) - C(x_1, y_2) + C(x_1, x_2)$$

Una condición necesaria para la función en cuestión sea una cópula, debe ser cierto que  $\Delta_{[x,y]}C \geq 0$  para todo  $u$  en el hiper-rectángulo  $[x,y]$ . La interpretación del  $C$ -volumen resulta impactante, finalmente equivale a la probabilidad de que un vector  $u$  esté en un hiper-rectángulo específico, de ahí su importancia. Mayor profundidad sobre el cosntructo teórico puede encontrarse en [21].

- $C(u_1, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_n) = 0$
- $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$

Las pruebas de las últimas dos propiedades se desprende de simple uso de desigualdad triangular y propiedades de la probabilidad como medida. Más allá de ello, cabe comentar un poco al respecto. Recuerde el lector que una cópula es en sí mismo un modelo multivariado, luego  $u_i = 0$  equivale a  $F_i(X_i) \leq 0$ , lo cual tiene un valor nulo a nivel univariado dado que  $F_i$  es una distribución acumulada, este razgo se extiende naturalmente al modelo multivariado. La última propiedad no es más que la forma natural de expresar el supuesto Ceteris Paribus para una variable específica, dado  $F_i(1) \leq 1 = 1$ .

Finalmente, se dice que una cópula es absolutamente continua si admite una función de densidad, esto es:

$$c(u) = \frac{\partial^n}{\partial u_n \dots \partial u_1} C(u_1, \dots, u_n)$$

El presente trabajo, se enfoca en el uso de cópulas  $t$ -student.

## 7.2. Modelo GARCH

Hasta este punto ya se tiene claro cómo se van a unir las distribuciones univariadas en un modelo multivariado. Ahora bien, ¿Cómo se van a estimar las distribuciones marginales? Se hará mediante el uso del modelo GARCH. Para tales fines, se asume que el lector posee nociones básicas de econometría y series de tiempo, y que está familiarizado con conceptos como estacionariedad, modelo ARMA y ARIMA, en caso contrario, puede encontrar tal información en varias fuentes, como por ejemplo [10].

Recuérdese que uno de los supuestos fundamentales de las regresiones es el de homocedasticidad. Este establece que las innovaciones resultantes del modelo deben tener varianza constante. Esto está fuertemente relacionado con la varianza que presenta el modelo predictivo, por ejemplo, el modelo ARIMA. Si este no es el caso, el modelo debe complementarse con alguna otra técnica que se encargue de modelar la volatilidad cambiante. Es en este contexto donde el modelo GARCH resulta importante.

El modelo GARCH es una generalización del modelo ARCH, que significa *Autoregressive Conditional Heteroscedasticity*. Este es un modelo de series de tiempo que describe la varianza de las innovaciones o vector de errores como función de las innovaciones observadas en períodos pasados, en particular, en función del cuadrado de los errores. La generalización mencionada incluye como variables independientes a los mismos rezagos de la varianza.

Ahora bien, ¿Cuándo resulta útil aplicar tal modelo? Considere el lector el siguiente ejemplo: El comportamiento del valor de mercado de acciones o de coeficientes de renta variable. En períodos de relativa calma, suele existir baja volatilidad. Si hoy el precio es estable, es posible que esto sea en parte consecuencia de un movimiento también estable del día de ayer. Así mismo, si hoy los mercados entran en una gran turbulencia, es entonces probable que el día de mañana continúe tal comportamiento errático. Note como en general la volatilidad no es la misma a lo largo de períodos extensos de tiempo y cómo la volatilidad específica en cierto período puede influir fuertemente en el comportamiento del activo el día de mañana.

El modelo GARCH resulta ser aplicable al presente trabajo toda vez que el rendimiento o variación porcentual del valor de cierre de los índices bursátiles S&P-500, DAX y Nikkei tienen medias muestrales cercanas a cero, y una dispersión con respecto a la misma que no resulta ser constante.

En cuanto a la formulación matemática, un modelo ARCH(p) es un modelo que expresa la varianza  $\sigma_t^2$  como se sigue:

$$\sigma_t^2 = \lambda_0 + \lambda_1 u_{t-1}^2 + \lambda_2 u_{t-2}^2 + \dots + \lambda_p u_{t-p}^2$$

Donde  $U_t$  es el vector de innovaciones.

Un modelo GARCH, o la versión generalizada, incluye rezagos de la misma varianza. Por ejemplo, un modelo GARCH(p,q) se estructura de la siguiente forma:

$$\sigma_t^2 = \lambda_0 + \lambda_1 u_{t-1}^2 + \lambda_2 u_{t-2}^2 + \dots + \lambda_p u_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \dots + \beta_q \sigma_{t-q}^2$$

Con respecto a la última ecuación, note que los parámetros  $(p,q)$  denotan el nivel de rezagos del vector de innovaciones y del vector de varianzas. Es decir, se explica la varianza actual del modelo en función de las innovaciones y la varianza de períodos anteriores. Esta sería la forma de modelar la varianza observada.

Para la media, se usará un modelo ARMA. Recuerde el lector que un modelo ARMA plantea a la variable dependiente en función de rezagos propios y del vector de innovaciones. Un modelo ARIMA hace lo mismo, solo que como paso previo se requiere la  $i$ -ésima diferenciación de las series. Esto es, un modelo ARIMA( $p,i,q$ ) es un modelo que explica a la variable dependiente  $Y_t$  en función de  $p$  rezagos de sí misma y en función de  $q$  rezagos del vector de perturbaciones. El término  $i$  hace alusión al número de veces que deba diferenciarse la serie para obtener estacionariedad, supuesto clave en el ejercicio.

### 7.3. Coeficientes de correlación

No tésese que ya se ha caracterizado y se han mostrado los rasgos fundamentales de las cópulas. No obstante, no se ha discutido cómo estas modelan la correlación existente entre variables aleatorias. Empecemos por discutir qué medidas de correlación existen.

Una de las más famosas y ampliamente usadas es el coeficiente de correlación de Pearson. Este se define como se sigue

$$\rho(x,y) = \frac{\text{Cov}(x,y)}{\sigma_x \sigma_y}$$

Con  $x, y$  variables aleatorias y  $\sigma_x, \sigma_y$  sus respectivas desviaciones. El lector resultará familiarizado con tal coeficiente dado que este caracteriza la dirección y fortaleza de la correlación entre las variables aleatorias. Es sencillo ver que  $-1 \leq \rho(x, y) \leq 1$ .

Este coeficiente es ampliamente usado. Sin embargo, se deben hacer algunos comentarios al respecto. Primero, entre más cercano sea  $|\rho(x, y)|$  a 1 resulta más fuerte la relación *lineal* entre las variables aleatorias. ¿Se sigue entonces que entre más próximo sea  $|\rho(x, y)|$  a cero se tiene independencia entre las variables aleatorias? La respuesta es un rotundo *No*. Un valor cercano a cero implica que la relación lineal es débil, esto no necesariamente implica que las variables no estén relacionadas, toda vez que alguna dinámica más compleja puede subyacer al contexto. Segundo, si los momentos de orden dos no existen, tampoco existe el coeficiente de correlación en cuestión. Estos, junto otros potenciales problemas teóricos, muestran la necesidad de al menos considerar otras opciones.

Una aplicación puntual de otros coeficientes de correlación es mostrada en [13]. Considérese entonces las siguientes medidas de correlación de rango.

- *Versión poblacional del  $\rho(x, y)$  de Spearman*: Sea  $(X_1, X_2)$  un vector aleatorio bivariado con distribuciones marginales  $F_1$  y  $F_2$ . El coeficiente se define como:  $\rho_s(X_1, X_2) = \text{Cor}(X_1, X_2)$

Donde *Cor* es el coeficiente usual de correlación. Note que el problema de existencia queda automáticamente resuelto, toda vez que  $F(X_i)$  con  $i = 1, 2$  se distribuye uniformemente en el intervalo  $(0, 1)$  y la varianza de estas distribuciones es bien conocida. ■  *$\tau$  de Kendall*: Acorde con [20], se puede definir el coeficiente  $\tau$  como:

$$\tau = \frac{(\text{núm. de pares concordantes}) - (\text{núm. de pares discordantes})}{\binom{n}{2}}$$

Entiéndase por un par concordante de la siguiente forma: Sean  $(X_i, Y_i)$  y  $(X_j, Y_j)$  con  $j \neq i$ . Se dice que es un par concordante si  $X_i > X_j$  y  $Y_i > Y_j$ , o bien,  $X_i < X_j$  y  $Y_i < Y_j$ . En otro caso, es un par discordante.

Finalmente, acorde con [21], tanto  $\rho$  de Spearman como  $\tau$  de Kendall están determinadas unívocamente por la cópula, de la siguiente forma

$$\rho(C) = 12 \int_{[0,1]^2} C(u) du - 3$$

$$\tau(C) = 4 \int_{[0,1]^2} C(u) dC(u) - 1$$

En resumen, las distribuciones univariadas asociadas a las variables pertinentes serán calculadas mediante el modelo ARMA-GARCH. Para el cual, la media se modelará mediante el modelo ARMA, mientras que la varianza será estudiada mediante el modelo GARCH. Posteriormente, se transforman las observaciones en elementos del intervalo (0,1) a través de lo enunciado en el primer teorema de esta sección. Con base en ello, se calcula una cópula *t*-student, de la cual se obtendrá el valor estimado del coeficiente  $\rho(X,Y)$  de Spearman, aplicado dos a dos a los índices de renta variable previamente mencionados.

Como comentario final, es importante resaltar la razón de ser del uso de la cópula *t* para el presente trabajo. Primero, los tipos de cópulas más frecuentemente utilizadas dada su versatilidad son las cópulas elípticas y las arquimedianas. Las primeras son bastante útiles para modelar rendimientos de activos financieros, dado que estas observaciones tienden a estar cerca a su media muestral, la cual suele ser cercana a cero. Dentro de esta familia de cópulas se encuentran varios tipos específicos de Cópulas, las principales son la normal y la *t*. Esta última tiende a utilizarse frecuentemente dado que el modelo multivariado presenta colas más “pesadas” y por tanto se adapta mejor a la naturaleza del problema.

## 8. Análisis de los resultados

Las figuras 2, 3 y 4 de la sección de anexos muestra el comportamiento de los índices DAX, S&P-500 y Nikkei, representando así los mercados de valores de Estados Unidos, Alemania y Japón. El período de referencia es 2015-2020 y la frecuencia es semanal.

Los tres índices tienen un comportamiento esencialmente similar. Por un lado, se observa un sostenido incremento en el período 2015-2018. Durante el 2018 se presenta una mayor volatilidad en los mercados, mientras que para 2019 se retoma la senda de crecimiento observada anteriormente. Finalmente, en el año 2020 ocurre un gran y fuerte desplome durante el primer trimestre del año, siendo esto totalmente consecuente con la debacle económica que resultó ser la pandemia asociada al COVID-19 y la parálisis que sufrió el sector productivo, acorde con lo expuesto en [18]. No obstante, estos índices recuperaron rápidamente su trayectoria, dadas las expectativas existentes sobre algunos sectores claves. Por otro lado, cabe aclarar que las tendencias durante períodos específicos de tiempo pueden ser respaldadas por análisis de factores fundamentales y bursátiles. No obstante, trabajos como [28] muestran que en general el comportamiento de índices como el S&P-500 es caótico.

En cuanto a los datos, estos cubren el período 2015-2020 y la frecuencia es semanal. Este período se escoge por principalmente por dos razones: Primero, refleja un repunte de los mercados de capitales a nivel mundial, dado que en el año 2015 varias economías, como las europeas, empiezan a recuperar niveles de crecimiento observados similares a los existentes previo al período 2008-2010. Este regreso a una senda de crecimiento estable permite un análisis más confiable en cuanto al vínculo existente entre tales mercados. Por otro lado, el período 2015-2020 permite recolectar una muestra adecuada de observaciones para el estudio en cuestión.

Por otro lado, la frecuencia es semanal por dos motivos: Primero, la frecuencia mensual o más extensa implicaría una cantidad insuficiente de datos. Segundo, y más importante, es fácil ver que existe una baja correlación entre tales índices a nivel diario. Esto resulta evidente dado que las mismas fuerzas determinan a tales bolsas en períodos largos de tiempo. No obstante, el ruido generado a nivel diario por factores locales, y ciertamente minúsculos, lleva a que se pierda de vista tales factores que impulsan el comportamiento observado. Los datos fueron obtenidos del portal *Yahoo Finance*.<sup>10</sup>

---

<sup>10</sup> Menos del 1% de los datos eran nulos, esto es, en vez del valor de cierre se registró "Null" como valor, posiblemente debido a inactividad o problemas con la fuente primaria de datos. Estos fueron reemplazados con el dato no nulo inmediatamente anterior. Por ejemplo, si en la primera semana de Marzo se registra "Null", entonces esto es reemplazado por lo observado en la última semana de Febrero, asumiendo que esta última sí presente información.

Los cálculos fueron hechos en R-Project. El código principal se encuentra en la sección de *Anexos*. Al igual que los gráficos previamente mencionados. Las librerías utilizadas fueron *xts*, *qrmtools*, *rugarch*, *forecast*, *tseries*, *copulas* y *dygraphs*. Los datos, que por defecto están en términos de valor de cierre, se transforman en variación porcentual, dado que se desea estudiar la correlación entre el rendimiento de los índices involucrados, esto se logra con la librería *qrmtools*. Adicionalmente, es útil unificar los datos en un único objeto, con propiedades de series de tiempo, esto se logra mediante el uso de *xts*. Como se menciona más adelante, la librería *forecast* se usó para indagar respecto al orden escogido del modelo ARMA, mediante la función “*auto.arima*”, esto también permite verificar la estacionariedad de las series. La librería *tseries* fue clave para ejecutar varias pruebas de estacionariedad, esto se discute más adelante. Finalmente, *rugarch* se utiliza para la estimación de la varianza, al aplicar el modelo GARCH previamente mencionado. Finalmente, *copulas* fue útil en la conversión de los datos en una muestra uniformemente distribuida en  $(0,1)$ , en la estimación de la cópula  $t$ , al estimar los parámetros asociados por el método de Máxima Verosimilitud. Finalmente, esta librería permite calcular numericamente el coeficiente de correlación de Spearman, tomando como base la información resultante de la cópula calculada.

Se realizaron pruebas de estacionariedad sobre los retornos de los índices en cuestión. Con un nivel de significancia de 5% se rechaza la hipótesis de que no seas estacionarios, la prueba usada fue la ADF<sup>11</sup>, aunque otras llevan a resultados similares. El análisis gráfico de las series de los retornos muestran que, en efecto y exceptuando inicios del 2020, la volatilidad no varía fuertemente a lo largo del tiempo. Por lo anterior se procede a estimar la distribución univariada mediante el modelo ARMA-GARCH. La media siendo analizada mediante un enfoque ARMA y la varianza mediante el enfoque GARCH. Con base en esto, se realiza la conversión de las series en vectores definidos sobre el intervalo  $(0,1)$ .

En textos como [13] se usa por defecto un modelo ARMA de orden  $(1,1)$ . No obstante, se usa la librería “*Forecast*” para determinar el mejor orden de ARMA para las series, mediante la función “*auto.arima*” Cabe resultar que el resultado para índices como el S&P-500 arroja un orden  $(2,3)$ . Se contrasta con el orden inicialmente mencionado, se observa que el valor de coeficiente varía, no obstante, tales cambios son leves. Para los índices Nikkei y DAX sí se utiliza el orden  $(1,1)$ . La función previamente mencionada sugiere que el mejor orden del modelo ARMA para estos dos casos es  $(0,0)$ . Se calcula el modelo  $(1,1)$  y se obtiene que los índices de información son similares. No obstante, el uso de órdenes  $(0,0)$ ,  $(1,0)$  y  $(0,1)$  también resultan viables, aunque pueden obtenerse diferencias en cuanto al cálculo del coeficiente de Spearman con respecto al orden  $(1,1)$

---

<sup>11</sup> Esto se refiere a la prueba aumentada de Dickey-Fuller, aplicada para estudiar la posible estacionariedad de las series. En [10], entre otros, se realiza una descripción de la misma.



de hasta  $\pm 0,05$ . Ahora, el orden sugerido por la librería, (0,0), puede sugerir el tratamiento de las series como ruido blanco.

Por otro lado, al modificar el orden del modelo GARCH, el resultado final no cambia de manera dramática. Cabe mencionar que al establecer (1,0) como orden para la varianza, los resultados del coeficiente de Spearman y Pearson difieren por cerca de 0.05, este escenario plantea la mínima diferencia observada entre tales coeficientes. Se establece el orden (1,1) para el modelo GARCH dado que órdenes superiores no generan cambios significativos en el resultado final, pero sí generan mayor complejidad a nivel computacional.

Todos los parámetros de los modelos univariados se especifican en la sección del código comentada bajo ese mismo nombre. Esto, en conjunción con los datos ya transformados, se utilizan como insumo para la función “ugarchfit”, la cual realiza recursivamente los cálculos asociados al modelo GARCH deseado.

Así mismo, se hallan los residuos del modelo y se utilizan los datos obtenidos para obtener los vectores uniformemente distribuidos en (0,1) que son clave para el cálculo de la cópula. Esto se hace mediante la función “pobs” de la librería *copula*, la cual permite transformar las pseudo-observaciones involucradas.<sup>12</sup>

El paso final es el cálculo del coeficiente de Spearman mediante el ajuste o cálculo de la mejor cópula vía Máxima Verosimilitud. Esto se puede lograr con la función “fitcop”.<sup>13</sup>

Con base en tal descripción, se procede a estudiar los resultados y particularidades arrojados por el código.

Se comparan los índices de correlación de Pearson y el de Spearman, con base en las especificaciones de los modelos univariados previamente mencionadas. Al analizarse las series dos a dos, por ejemplo, S&P500 vs Nikkei. Se llega a que el primero es aproximadamente 0.6873, mientras que este último es 0.61739. Al estudiar la relación entre los índices S&P500 y DAX se obtiene un coeficiente de Spearman de 0.68218. En este ejercicio, los resultados arrojados por tales coeficientes de correlación sugieren que en efecto existe una fuerte vínculo entre estos índices bursátiles. No obstante, el lector debe tener en cuenta las limitantes del coeficiente de Pearson.

---

<sup>12</sup> Para mayor información sobre el tratamiento de pseudo-observaciones para la estimación de cópulas, se recomienda consultar [13].

<sup>13</sup> Los resultados explícitos de las cópulas, así como de pruebas de estacionariedad, inspección del modelo ARMA, entre otros, se encuentran en la sección de anexos.

Ahora bien, ¿Qué indicativo se tiene de que la aplicación del modelo es correcta? En principio, se reduce al análisis del modelo ARMA-GARCH, en particular, al vector de innovaciones resultante. En específico, se pide que este se comporte como ruido blanco. Para tal propósito, se realiza la prueba Ljung-Box y se realiza un análisis gráfico mediante la función PACF. Los resultados se encuentran en la sección de anexos, pero en general, el vector de errores estandarizados sí se comporta como ruido blanco. Luego de ellos, se transforman tales innovaciones en observaciones uniformemente distribuidas, el lector puede encontrar el histograma de tales observaciones en la sección de anexos. Habiendo asegurado la calidad del modelo GARCH y de que las observaciones resultantes están uniformemente distribuidas en  $(0,1)$ , el uso de cópulas  $t$  está totalmente justificado. Cabe también aclarar que el cálculo del coeficiente de Spearman no presupone ningún tipo de distribución específica, toda vez que consiste en el cálculo del coeficiente de correlación usual sobre vectores de la forma  $(F(X_1), F(X_2))$ , donde  $F$  es la función de densidad acumulada y  $X_{1,2}$  son variables aleatorias. Claramente este vector tiene distribución uniforme sobre  $(0,1)$ . Por tanto, no se necesita ningún supuesto sobre la distribución subyacente.

Finalmente es relevante mencionar que el uso de cópulas  $t$  se ajusta de manera adecuada al estudio de rendimiento de activos en renta variable, dado el carácter muestral de la información y dado el comportamiento de tales rendimientos. Dada la volatilidad variable de las series el modelo ARMA-GARCH resulta ser una herramienta útil para la estimación univariada, en caso de estacionariedad. En caso contrario, una simple extensión ARIMA para la media puede funcionar. No obstante, la escogencia del orden de tales modelos es importante. Si bien el resultado en cuanto a la correlación no cambia dramáticamente, sí es importante definir tal orden a priori, o por lo menos, comparar modelos, para tener mayor precisión en cuanto a la correlación existente.

## 9. Conclusiones

La medición de la dependencia estructural entre activos u otras variables relativas al mercado de capitales resulta ser un ejercicio fundamental. Cópulas ofrece una manera robusta y práctica de cuantificar la correlación existente entre dos variables. Esto es posible dado que coeficientes como el de Spearman, están unívocamente determinados por la cópula subyacente. Esto es totalmente coherente con la naturaleza misma de una cópula como objeto matemático, es una función que permite transformar modelos univariados disconexos en un modelo multivariado completo. Por tanto, el reto consiste en determinar la forma de establecer los modelos univariados y en escoger la cópula que mejor se ajuste a los datos. En el caso de cópulas paramétricas, esta estimación se realiza usualmente por el método de Máxima Verosimilitud.

En este trabajo se escogió el modelo ARMA-GARCH para hallar los modelos univariados, para luego ser transformados en una variable aleatoria uniformemente distribuida en el (0,1). Luego se estimó la cópula  $t$ -student que mejor se ajusta a las observaciones para de esta manera cuantificar la correlación entre los índices S&P-500, DAX y Nikkei, de manera pareada

El análisis muestra que existe en efecto una alta correlación entre los índices S&P-500, DAX y Nikkei. En general, el resultado arrojado por la estimación del coeficiente de Spearman está entre 0,6 y 0.7. En general, esto es consistente con el nivel de integración de los mercados de renta variable en la actualidad. También se observa que la escogencia del orden de los modelos univariados generan leves cambios en cuanto a la estimación del coeficiente de Spearman, toda vez que tales cambios se ven reflejadas por la cópula ajustada a los datos.

## 10. Anexos

### 10.1. Más sobre los índices utilizados

- S&P-500: Este índice se calcula con base en las 500 firmas más representativas y líquidas de las principales industrias del mercado de Estados Unidos. La metodología específica en cuanto al cálculo numérico y el cómo se escogen tales firmas se encuentra en la página web de su operador *S&P Dow Jones Indices*.
- DAX: Este índice de acciones se basa en el valor de mercado de las acciones de las 30 compañías más grandes de Alemania que cotizan en la Bolsa de Frankfurt. Su operador y fuente de información oficial es *Deutsche Börse*.

- Nikkei: Este índice se calcula a partir de los 225 valores más líquidos que cotizan en la Bolsa de Tokio, su operador oficial es *Nihon Keizai Shimbun*.



Figura 3: Período 2015-2020. Cálculos propios en R.

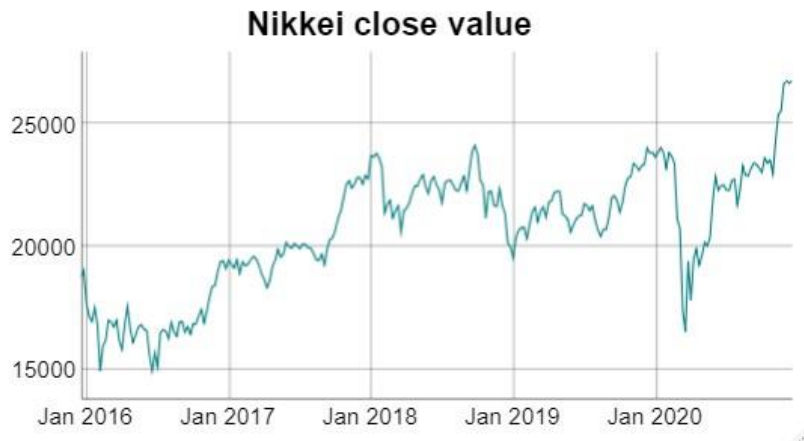


Figura 4: Período 2015-2020. Fuente. cálculos propios en R.

El código usado fue el siguiente

### 10.3. Código utilizado y resultados del código

#### 10.3.1. Código

Para la estimación del modelo GARCH, se tomó como referencia a la versión inicial encontrada en [13].

```
library ( xts ) library (
qrmtools ) library (
rugarch ) library (
forecast ) library (
tseries ) library (
copula ) library (
dygraphs )
```

```
#Cargue del df para el modelo y tratamiento
```

```
EUR[ c ( 2 , 3 , 4 , 6 , 7 ) ] <- list ( NULL )
```

```
USD[ c ( 2 , 3 , 4 , 6 , 7 ) ] <- list ( NULL )
```

```
JPY[ c ( 2 , 3 , 4 , 6 , 7 ) ] <- list ( NULL )
```

```

time_df <- as.Date(EUR$Date) df <- data.frame(matrix(ncol = 4, nrow = length
(EUR$Date))) df_names <- c("Date", "EUR", "USD", "JPY") colnames(df) <- df
names df$Date <- time_df
df$EUR <- as.numeric(EUR$Close)
df$USD <- as.numeric(USD$Close) df$JPY
<- as.numeric(JPY$Close)

```

```

Rport <- returns(df[, -1])
Xport <- xts(Rport, order.by = df[-1, 1]) Inspeccin

```

de ARIMA y estacionariedad

```

adf.test(Xport$EUR) adf.test
(Xport$USD) adf.test
(Xport$JPY)

```

```

auto.arima(Xport$EUR)
auto.arima(Xport$USD)
auto.arima(Xport$JPY)

```

#Specs de los univariados

```

meanModel_EUR <- list(armaOrder = c(1,1)) varModel_EUR <- list(model = "sGARCH"
, garchOrder = c(1,1)) uspec_EUR <- ugarchspec(varModel_EUR, mean.model =
meanModel_EUR, distribution.model = "std")

```

```

fit_EUR <- apply(Xport$EUR, 2, function(x) ugarchfit(uspec_EUR, data = x))
epsEUR <- sapply(fit_EUR, residuals, standardize = TRUE) U_EUR <- pobs(eps
EUR)

```

```

meanModel_USD <- list(armaOrder = c(2,3)) varModel_USD <- list(model = "sGARCH"
, garchOrder = c(1,1)) uspec_USD <- ugarchspec(varModel_USD, mean.model =
meanModel_USD, distribution.model = "std")

```

```

fit_USD <- apply(Xport$USD, 2, function(x) ugarchfit(uspec_USD, data = x))

```

```
epsUSD <- sapply ( fit USD ,residuals , standardize = TRUE) U USD <- pobs(eps
USD)
```

```
meanModel JPY <- l i s t (armaOrder = c (1 ,1)) varModel JPY <- l i s t (model = "sGARCH"
, garchOrder = c (1 ,1)) uspec JPY <- ugarchspec (varModel _JPY , mean. model =
meanModel _JPY , distribution . model = "std ")
```

```
fit JPY <- apply (Xport$JPY , 2 , function (x) ugarchfit (uspec _JPY , data = x))
epsJPY <- sapply ( fit JPY , residuals , standardize = TRUE) U JPY <- pobs(eps
JPY)
```

```
#Copula EUR USD
```

```
U1 <- cbind (U EUR, U USD) fitcop <- fitCopula (tCopula () , data = U1, method
= "mpl") coef ( fitcop ) cop <- fitcop@copula
```

```
cor (Xport$EUR,Xport$USD)
```

```
#Copula JPY USD U2 <- cbind (U JPY, U USD) fitcop <- fitCopula (tCopula () ,
data = U2, method = "mpl") coef ( fitcop )
```

```
cop <- fitcop@copula cor
(Xport$JPY ,Xport$USD)
```

```
#Copula JPY EUR U3 <- cbind (U JPY, U EUR) fitcop <- fitCopula (tCopula () ,
data = U3, method = "mpl") coef ( fitcop )
```

```
cop <- fitcop@copula cor
(Xport$JPY ,Xport$EUR)
```

Algunos de los resultados del código son los siguientes:

### 10.3.2. Inspección del modelo ARMA

“USA” se refiere al índice SP500, “EUR” al índice DAX y “JPY” al Nikkei.



auto . arima(Xport\$EUR) Series :  
Xport\$EUR ARIMA(0,0,0) with zero  
mean

sigma<sup>2</sup> estimated as 0.000933: log likelihood =538.1 AIC=-1074.19  
AICc=-1074.17 BIC=-1070.63

auto . arima(Xport\$USD) Series :  
Xport\$USD  
ARIMA(2,0,3) with non-zero mean

Coefficients :

ar1	ar2	ma1	ma2	ma3	mean
0.9698	-0.8799	-1.0503	1.0470	-0.1588	0.0023
s . e . 0.1503 0.1552					
0.1539	0.1158	0.0724	0.0014		

sigma<sup>2</sup> estimated as 0.0006266: log likelihood =592.64  
AIC=-1171.28 AICc=-1170.84 BIC=-1146.36

auto . arima(Xport\$JPY) Series :  
Xport\$JPY  
ARIMA(0,0,0) with zero mean

sigma<sup>2</sup> estimated as 0.00096: log likelihood =534.39  
AIC=-1066.78 AICc=-1066.77 BIC=-1063.22

### 10.3.3. Prueba de estacionariedad adf . test (Xport\$EUR)

Augmented Dickey-Fuller Test

data : Xport\$EUR  
Dickey-Fuller = -6.7708, Lag order = 6 , p-value = 0.01 alternative  
hypothesis : stationary adf . test (Xport\$USD)

#### Augmented Dickey-Fuller Test

data : Xport\$USD

Dickey-Fuller = -6.8986, Lag order = 6, p-value = 0.01 alternative hypothesis : stationary

#### Augmented Dickey-Fuller Test

data : Xport\$JPY

Dickey-Fuller = -6.5636, Lag order = 6, p-value = 0.01 alternative hypothesis : stationary

### 10.3.4. Resultados del la c´opula acorde con el orden ARMA-GARCH

Se entiende por “rho” como el coeficiente  $\rho$  de Spearman y “df” como los grados de libertad de la co´pula.

Orden ARMA

A1 <- c (1 ,1) EUR

A2 <- c (2 ,3) USA

A3 <- c (1 ,1) JPY Orden

GARCH

G1 <- c (1 ,1) EUR G2 <-

c (1 ,1) USA

G3 <- c (1 ,1) JPY

coef (fitcop EUR USD) rho .1

df

0.6821824 6.1783163 coef

(fitcop JPY USD) rho .1 df

0.6173971 9.8679214 coef

(fitcop JPY EUR) rho .1 df

0.7062449 39.2632678

### 10.3.5. Coeficiente de Pearson

USD  
EUR 0.750504

USD  
JPY 0.6873284

EUR  
JPY 0.7714999

### 10.3.6. Sobre el modelo ARMA-GARCH: vector de innovaciones

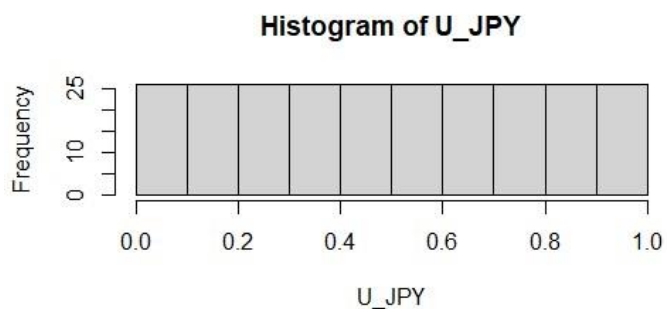


Figura 5: Vector de innovaciones estandarizadas del índice Nikkei, luego de ser transformado en un vector uniformemente distribuido en (0,1)

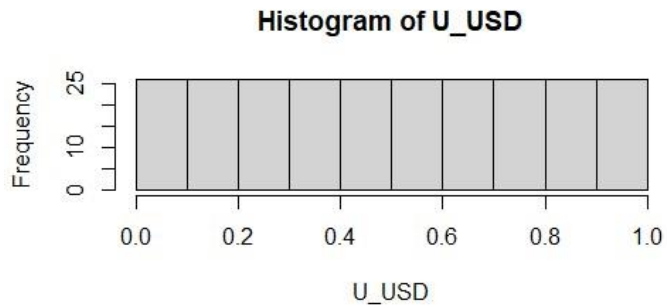


Figura 6: Vector de innovaciones estandarizadas del índice SP500, luego de ser transformado en un vector uniformemente distribuido en (0,1)

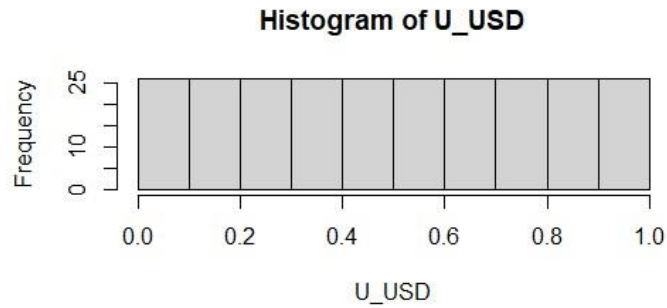


Figura 7: Vector de innovaciones estandarizadas del índice DAX, luego de ser transformado en un vector uniformemente distribuido en (0,1)

A continuación se muestra la PACF respecto al vector de innovaciones resultantes del modelo GARCH.

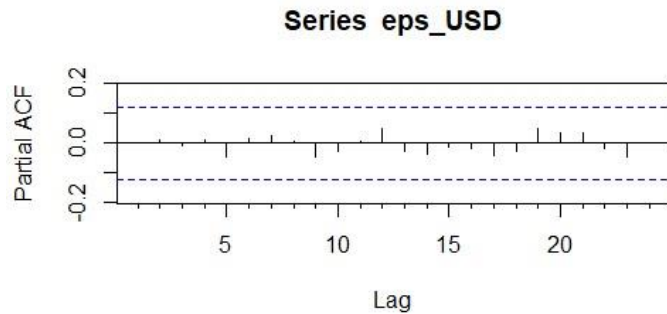


Figura 8: Índice SP500.

Finalmente, los resultados del test Ljung-Box son los siguientes:

data : epsUSD  
 X-squared = 0.00085437 , df = 1 , p-value = 0.9767 data : epsEUR  
 X-squared = 0.38665 , df = 1 , p-value = 0.5341 data : epsJPY  
 X-squared = 0.16248 , df = 1 , p-value = 0.6869

En general, los resultados soportan la hipótesis de que las perturbaciones del modelo ARMA-GARCH se comportan como ruido blanco.

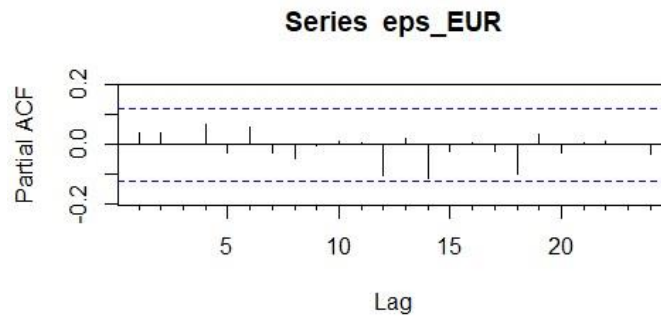


Figura 9: Indice DAX.

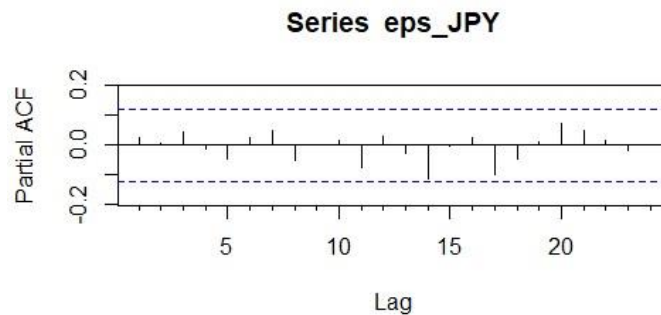


Figura 10: Indice Nikkei.

## Referencias

- [1] Riadh Aloui, Mohamed Safouane Ben Aissa y Duc Khuong Nguyen. "Conditional dependence structure between oil prices and exchange rates: a copula-GARCH approach". En: *Journal of International Money and Finance* 32 (2013), págs. 719-738.
- [2] Lu Chen y Shenglian Guo. *Copulas and Its application in hydrology and water resources*. Springer, 2019.
- [3] Ray Y Chou, Victor Ng y Lynn K Pi. "Cointegration of international stock market indices". En: *IMF Working Papers* 94 (1994), pág. 94.
- [4] Jeffrey Chu y col. "GARCH modelling of cryptocurrencies". En: *Journal of Risk and Financial Management* 10.4 (2017), pág. 17.
- [5] Arnaud Costinot, Thierry Roncalli y Jerome Teiletche. "Revisiting the dependence between financial markets with copulas". En: *Available at SSRN 1032535* (2000).

- [6] MH DeGroot y MJ Schervish. *Probability and Statistics, Pearson Education*. 2012. [7] Majid Dehghani, Bahram Saghafian y Mansoor Zargar. "Probabilistic hydrological drought index forecasting based on meteorological drought index using Archimedean copulas". En: *Hydrology Research* 50.5 (2019), págs. 1230-1250.
- [8] Fabrizio Durante, Juan Fernandez-Sanchez y Carlo Sempi. "A topological proof of Sklar's theorem". En: *Applied Mathematics Letters* 26.9 (2013), págs. 945-948.
- [9] Maurice Fréchet. "Sur les tableaux de corrélation dont les marges sont données". En: *Ann. Univ. Lyon, 3<sup>e</sup> serie, Sciences, Sect. A* 14 (1951), págs. 53-77.
- [10] Damodar N Gujarati. "Econometrics by example". En: (2011).
- [11] Hui Guo y Christopher J Neely. "Investigating the intertemporal risk–return relation in international stock markets with the component GARCH model". En: *Economics letters* 99.2 (2008), págs. 371-374.
- [12] Zengchao Hao y Vijay P Singh. "Review of dependence modeling in hydrology and water resources". En: *Progress in Physical Geography* 40.4 (2016), págs. 549-578.
- [13] Marius Hofert y col. *Elements of copula modeling with R*. Springer, 2019.
- [14] Wassilij Hořffding. "Masstabinvariante korrelationstheorie". En: *Schriften des Mathematischen Instituts und Instituts für Angewandte Mathematik der Universität Berlin* 5 (1940), págs. 181-233.
- [15] Eric Jondeau y Michael Rockinger. "The copula-garch model of conditional dependencies: An international stock market application". En: *Journal of international money and finance* 25.5 (2006), págs. 827-853.
- [16] Paraskevi Katsiampa. "Volatility estimation for Bitcoin: A comparison of GARCH models". En: *Economics Letters* 158 (2017), págs. 3-6.
- [17] Yan Liu y Richard Luger. "Efficient estimation of copula-GARCH models". En: *Computational Statistics & Data Analysis* 53.6 (2009), págs. 2284-2297.
- [18] Mieszko Mazur, Man Dang y Miguel Vega. "COVID-19 and the march 2020 stock market crash. Evidence from S&P1500". En: *Finance Research Letters* (2020), pág. 101690.
- [19] Beatriz Vaz de Melo Mendes y Rafael Martins de Souza. "Measuring financial risks with copulas". En: *International Review of Financial Analysis* 13.1 (2004), págs. 27-45.
- [20] RB Nelsen. "Kendall tau metric". En: *Encyclopaedia of mathematics* 3 (2001), págs. 226-227.

- [21] Roger B Nelsen. *An introduction to copulas*. Springer Science & Business Media, 2007.
- [22] Aristidis K Nikoloulopoulos, Harry Joe y HaijunV Li. "Vine copulas with asymmetric tail dependence and applications to financial return data". En: *Computational Statistics & Data Analysis* 56.11 (2012), pa'gs. 3659-3673.
- [23] Thabani Nyoni. "Modeling and forecasting inflation in Kenya: Recent insights from ARIMA and GARCH analysis". En: *Dimorian Review* 5.6 (2018), pa'gs. 16-40.
- [24] Yue Peng y Wing Lon Ng. "Analysing financial contagion and asymmetric market dependence with volatility indices via copulas". En: *Annals of Finance* 8.1 (2012), pa'gs. 49-74.
- [25] Frits H Ruymgaart y MCA van Zuijlen. "Asymptotic normality of multivariate linear rank statistics in the non-iid case". En: *The Annals of Statistics* (1978), pa'gs. 588-602.
- [26] Jun Shao. *Mathematical statistics*. 2003.
- [27] M Sklar. "Fonctions de repartition an dimensions et leurs marges". En: *Publ. inst. statist. univ. Paris* 8 (1959), pa'gs. 229-231.
- [28] Michail D Vamvakaris, Athanasios A Pantelous y Konstantin M Zuev. "Time series analysis of S&P 500 index: A horizontal visibility graph approach". En: *Physica A: Statistical Mechanics and Its Applications* 497 (2018), pa'gs. 41-51.
- [29] Yan-hua WEI y Shi-ying ZHANG. "Dependence Analysis of Finance Markets: Copula-GARCH Model and Its Application [J]". En: *Systems Engineering* 4 (2004), pa'gs. 7-12.
- [30] Gregor NF Weiß. "Copula-GARCH versus dynamic conditional correlation: an empirical study on VaR and ES forecasting accuracy". En: *Review of Quantitative Finance and Accounting* 41.2 (2013), pa'gs. 179-202.
- [31] Edward F Wolff. "N-dimensional measures of dependence." En: *Stochastica* 4.3 (1980), p'ags. 175-188.