
Estimación Bayesiana para el cálculo del Valor en Riesgo (VaR) en modelos de series financieras con relaciones de dependencia no lineal en Colombia

Bayesian estimation for the calculation of Value at Risk (VaR) in models of financial series with relations of non-linear dependence in Colombia

Daniel Triana^a
juan.triana@usantotomas.edu.co

Luis Miguel Torres Aponte^b
luis.torres@usantotomas.edu.co

Miguel Ángel Alba^c
miguel.alba@usantotomas.edu.co

Wilmer Pineda-Ríos^d
wilmerpineda@usantotomas.edu.co

Resumen

El Valor en Riesgo (VaR), se define como la máxima pérdida que se puede tener en la inversión de un portafolio con un determinado nivel de confianza, en un periodo determinado, en condiciones normales del mercado. Para calcularlo, existen diversas herramientas paramétricas y no paramétricas que se fundamentan en el supuesto de que los rendimientos de los activos siguen una distribución de probabilidad (con frecuencia la distribución Normal). Por su parte las cópulas, vistas como funciones de distribución multivariadas, capturan las relaciones de dependencia diferentes a las lineales, ayudando a condensar la volatilidad de los activos multivariados que usualmente presentan comportamientos complejos de dependencia. Este trabajo presenta la implementación de la teoría de cópulas bajo estándares del paradigma bayesiano con el objetivo de obtener la distribución de probabilidad que permita una correcta estimación del VaR de un portafolio comprendido por las acciones de Bancolombia y Ecopetrol.

Palabras clave: Valor en riesgo, cópulas, dependencia no lineal.

Abstract

The Value at Risk (VaR) is defined as the maximum loss that could have a portfolio with a certain confidence level, in a given period and under normal conditions

^aProfesional en Estadística Universidad Santo Tomás

^bProfesional en Estadística Universidad Santo Tomás

^cProfesional en Estadística Universidad Santo Tomás

^dDocente Facultad de Estadística Universidad Santo Tomás

market. The value at risk is calculated by several parametric and non parametric tools that are based on assumptions on yields of the assets following a probability distribution (often the normal distribution). In another hand, the copulas seen as multivariate distribution function, captures non linear dependency relationship, helping to condense volatility on the multivariate assets that usually presents complex dependency behaviors. This paper presents the implementation of the copulas theory by Bayesian methods to obtain the probability distribution which is subsequently used to estimate the VaR for a portfolio comprised by the shares of Bancolombia and Ecopetrol. .

Keywords: Value at risk, copulas, non linear dependence.

1. Introducción

El mercado bursátil, se define como el espacio en que principalmente las empresas transan títulos de renta variable sin preestablecer un rendimiento, siendo las acciones el título característico de este mercado. La rentabilidad está ligada a las utilidades obtenidas por la empresa en la que se invierte, así como a las posibles variaciones en los precios de los valores dadas por las condiciones existentes en el mercado.

En Colombia, las únicas compañías autorizadas para la negociación de títulos de renta variable son las Sociedades Comisionistas de Bolsa y están obligadas a proporcionar información periódica y relevante, para que a su vez el público inversionista cuente con datos suficientes y oportunos para la toma de decisiones (Bolsa de Valores de Colombia, 2017). Un inversionista busca tomar decisiones que maximicen la rentabilidad esperada de la cartera y minimice simultáneamente el riesgo de la inversión al que siempre se encuentra expuesto, lo anterior, cuesta bastante trabajo debido a que se pueden presentar pérdidas en el valor de los activos de un portafolio causado por movimientos adversos realizados (González & Chacón, 2014). En la actualidad, los agentes económicos enfrentan diferentes riesgos tales como: riesgo de mercado, riesgo de crédito, riesgo operativo, riesgo legal, riesgo asociado a la reputación , entre otros (Banco BBVA, 2015).

Valor en Riesgo (*Value at Risk (VaR)* en inglés), es un método estadístico desarrollado por JP Morgan en los años 90 para cuantificar la exposición al riesgo de mercado. Fue adoptado rápidamente por las firmas financieras de todo el mundo y hace parte de las normas de estabilidad de mercados de los diferentes entes reguladores. Se caracteriza por ser una herramienta que principalmente adopta el supuesto de normalidad; supuesto que difícilmente se cumple en series financieras. Así mismo, al tener una amplia aplicación , existen diversas metodologías para ser estimado, las cuales se enfocan en obtener una distribución para la estimación de la varianza no constante de los rendimientos financieros, lo que no permite evidenciar la no correlación serial, pero si la dependencia entre los activos que componen un portafolio mediante los modelos GARCH-ARCH.

Teniendo en cuenta la volatilidad que presentan los activos financieros, asumir normalidad sobre los retornos parece lo más apropiado, dado esto y resaltando la dependencia existente entre los retornos utilizando modelos GARCH-ARCH, se propone implementar la teoría de cópulas. Una cópula se entiende como la relación entre dos o más funciones, la cual denota el enlace entre la distribución conjunta de un vector de probabilidad y las distribuciones marginales unidimensionales asociadas a esta.

El objetivo general de este trabajo es proponer un modelo que considere las relaciones de dependencia no lineal para estimar el Valor en Riesgo (VaR) en un portafolio de dos activos financieros. De igual manera, es de nuestro interés:

1. Identificar los modelos en series de tiempo que mejor expliquen el comportamiento de los rendimientos de los activos financieros.
2. Implementar un algoritmo que permita generar las funciones a-posteriori propias de la Estadística bayesiana con el fin de estimar los parámetros del modelo propuesto.
3. Estimar el Valor en Riesgo para el portafolio.

Este trabajo está organizado de la siguiente manera: en la sección 2, se complementa la introducción ofrecida acerca del Valor en Riesgo donde se mencionan los antecedentes y el objetivo de la implementación de esta teoría en el mundo financiero y de la economía global. En la sección 3, se profundiza la teoría de modelos de cópulas, donde se realiza un acercamiento al componente estadístico y matemático que las sustentan y algunas aplicaciones al campo financiero. En la sección 4, se dedica un espacio para introducir a los lectores sobre el paradigma bayesiano, donde podrán encontrar las bases estadísticas que soportan esta corriente del conocimiento. En la sección 5, se da un resumen que contiene los procedimientos que se implementan para el desarrollo del trabajo. En la sección 6, se ejecuta la aplicación de los modelos cópulas teniendo como insumo la información de un portafolio de inversión comprendido por los activos financieros de Bancolombia y Ecopetrol para el periodo comprendido entre el 30 de octubre de 2015 y el 31 de marzo de 2017. Finalmente, se presentan algunas conclusiones y recomendaciones con las que se abre un espacio de discusión en donde se validan los resultados obtenidos y se argumenta la relación de dependencia entre los rendimientos de activos financieros mediante el uso de modelos cópula con argumentos bayesianos; de igual manera, se establece la dependencia no lineal entre los retornos de las series financieras y se exponen los resultados de la estimación del Valor en Riesgo del portafolio.

2. Valor en Riesgo (VaR)

Para finales del siglo XX, los bancos comerciales de Estados Unidos según el Acuerdo de Capital de Basilea I (1988) iniciaron la implementación de los requerimientos

regulatorios sobre la función del riesgo de crédito o incumplimiento en el pago de sus activos, dichos requerimientos, se cumplieron a partir de la creación de la herramienta denominada modelo VaR. El VaR es creado por JP Morgan con el objetivo de estimar los rendimientos futuros frente a los movimientos del mercado teniendo en cuenta funciones de dependencia que permitan capturar la volatilidad del mercado y con base en dicho pronóstico, estimar la máxima pérdida posible en una inversión realizada (Lopez, 2006).

En términos formales, el VaR es una técnica Estadística que permite medir y cuantificar la exposición al riesgo de mercado, a un nivel de confianza de $1 - \alpha$. Se tiene que la distribución de pérdidas y ganancias tienen como variable aleatoria Q , donde una realización particular de esta sería q con su respectiva densidad $f_q(\cdot)$, entonces el nivel de significancia se puede definir como:

$$Pr [Q \leq VaR(\alpha)] = \alpha, \quad (1)$$

de igual forma, puede ser escrita mediante la siguiente integral (Danielsson, 2011).

$$\alpha = \int_{-\infty}^{VaR(\alpha)} f_q(x) dx = F_q(VaR(\alpha)). \quad (2)$$

Por lo tanto, el VaR puede escribirse como la inversa de la distribución de pérdidas y ganancias con un nivel de significancia dado, como se muestra en la siguiente ecuación .

$$VaR_\alpha = F_q^{-1}(\alpha), \quad (3)$$

se dice que el VaR es el primer valor de la variable de interés $q > -\infty$ hasta donde se acumula el α % del área bajo la curva.

Para calcular el VaR de un portafolio, se debe tener en cuenta las ponderaciones w_1, w_2 para cada activo, cumpliendo que $w_1 + w_2 = 1$. Luego, como se muestra en (Danielsson, 2011) la varianza σ_p^2 del portafolio será:

$$\sigma_p^2 = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \rho_{12} \sigma_1 \sigma_2, \quad (4)$$

donde σ_1^2 y σ_2^2 son la varianza de los rendimientos de los dos activos, w_1 y w_2 los pesos específicos del portafolio y ρ_{12} el coeficiente de correlación de los rendimientos; por lo tanto, se puede calcular el VaR con la siguiente ecuación :

$$VaR_P = S \sigma_p Z_\alpha \sqrt{t}, \quad (5)$$

donde S corresponde al valor invertido en el portafolio, σ_p es la desviación estándar del portafolio, Z_α es el percentil de la distribución normal a un grado α de significancia y t corresponde a la variable tiempo.

Sin embargo, el VaR necesita de otras herramientas para optimizar sus resultados relacionados con la inversión es por ello que Markowitz propone una metodología denominada *Frontera Eficiente*, este surge con el fin de apoyar el proceso de toma de decisiones sobre el portafolio, permitiendo identificar la distribución óptima de los recursos a invertir con un mínimo riesgo y una rentabilidad dada. Esta herramienta se ha convertido en la mano derecha del cálculo del VaR permitiendo mejorar las decisiones de inversión y facilitando el panorama de inversión tanto para expertos en el finanzas y economía, como para los menos experimentados en el tema (Calderón, J, 2016), de igual manera, la constante evolución de metodologías ha permitido la generación de propuestas basadas en nuevas funciones de dependencia para el cálculo de las variables que rigen al VaR, poniendo a disposición al sector financiero y económico modelos alternativos con mejores medidas de precisión .

3. Modelos cópula

Como se acaba de mencionar en la sección previa, una de las alternativas que se han desarrollado para el cálculo del VaR son los modelos cópula, estos se mencionan por primera vez en el teorema de Sklar (1959), que trata acerca de la relación existente entre una función multidimensional y las marginales asociadas a su dimensión inferior (Ceballos López, 2015). En síntesis, los modelos cópula enfocan sus esfuerzos en encontrar una dependencia no estricta lineal; desde la relación de dependencia de una serie univariada entre una secuencia de observaciones de una serie de tiempo escalar como el resultado de un análisis multivariado de series de tiempo, con el fin de de caracterizar la dependencia transversal o condicional entre series de tiempo individuales.

Andrew J. Patton en 2009, realiza una revisión detallada de las más relevantes e influyentes aplicaciones de cópulas en finanzas y economía, dando como resultado que el documento de Embrechts, McNeil y Straumann (2002) se categoriza por ser de los primeros en explorar los usos de la cópulas, desarrollando la teoría de la estimación y evaluación . Schweizer & Wolff publicaron un artículo acerca de la dependencia entre variables aleatorias con aplicación de modelos cópula para capturar dicha relación . Luego, hacia los años 90 se realizaron investigaciones acerca de la utilidad que pueden tener las cópulas en el área de riesgo financiero, a través de técnicas y herramientas que se pueden obtener a partir de la relación entre variables aleatorias. No obstante, es posible encontrar documentación relacionada con la introducción a las cópulas, Joe (1997) proporciona detalles los fundamentos estadísticos y matemáticos que son respaldados por Nelsen (2006), sin dejar a un lado, las aplicaciones y prácticas evaluadas por Cherubini, et al. (2004) sobre los precios de los derivados.

En el área de finanzas, la aplicación de las cópulas se ha incrementado en los últimos años para realizar procesos matemáticos de gran complejidad, entre ellas se resaltan los trabajos de Parra y Koodi (2006) que realizan una aplicación de la teoría de cópulas para estimar el VaR de un portafolio con los índices Nasdaq y S&P500, y otros ejercicios para medir desde el riesgo operacional hasta el riesgo financiero de portafolios con diversos activos; además, se pueden encontrar métodos óptimos para la clasificación y análisis de la información. Moróe en 2016 realizó un trabajo para predecir la volatilidad y el valor en riesgo en un portafolio de commodities, empleando modelos EGARCH y el método de simulación Monte Carlo; además compararon las cópulas t y normal para encontrar las relaciones de dependencia en el portafolio de estudio.

La comparación de cópulas suele ser significativa en el ajuste del comportamiento de los datos, como es el caso de Bolancé, Guillen & Padilla en 2015, en el que estimaron el valor en riesgo utilizando distintas cópulas, en este trabajo, los investigadores concluyen que la modelización de la dependencia se ve afectada significativamente por la elección de la cópula y la marginal utilizada, además, encontraron que la cópula t es la distribución que mejor modela el comportamiento de los activos.

Finalmente, una cópula es una función que enlaza las funciones de distribución univariada para formar una función de distribución multivariada. Si se cumple que todas las variables son distribuidas continuas, entonces la cópula será una función de distribución multivariada con distribuciones marginales univariadas uniformes de cero a uno $U(0,1)$.

$$F(x) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n)), \quad \forall x \in \mathbb{R}_n. \quad (6)$$

A partir de cualquier distribución multivariada F , es posible calcular las distribuciones marginales F_i y la relación cópula C . De igual manera, para un modelo de series de tiempo, cualquier conjunto de distribuciones marginales (F_1, F_2, \dots, F_n) y cualquier cópula C , la ecuación $F(x)$ se puede utilizar para obtener una distribución conjunta a partir de las distribuciones marginales, sin importar que dichas marginales sean similares entre sí o que se limite la elección de la cópula por la elección de las distribuciones marginales; esta característica convierte a las cópulas en una herramienta potencialmente útil para construir modelos econométricos.

Formalmente, considere una función 2-dimensional cópula C entendida como la relación entre las funciones $\int_{-Inf}^x f(t)dt = F_X$ y $\int_{-Inf}^y f(t)dt = F_Y$ donde es posible observar el vínculo entre la distribución conjunta de un vector de probabilidad y las marginales unidimensionales asociadas a esta, bajo las siguientes propiedades:

1. $C(X) = 0 \quad \forall X \in [0, 1]^2$
2. $C(x, 1) = C(1, y) = 1$ para todo (x, y) en el cuadrado $[0, 1]^2$
3. Para todo u_1, v_1, u_2, v_2 en $\mathbf{I} = [0, 1]^2$ tal que $u_1 \leq u_2$ y $v_1 \leq v_2$ se cumple

que

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0 \quad (7)$$

Tras el previo enunciado, es sencillo identificar que cada distribución marginal F_i contiene toda la información univariada sobre la variable individual X_i , mientras que la distribución conjunta F contiene tanto la información univariada como multivariada, por lo tanto, la cópula C deberá contener toda la información de dependencia entre las variables $X_{i's}$. Por tal motivo, las cópulas se conocen como funciones de dependencia (Galambos, 1978).

3.1. Teorema de Sklar

Según el teorema de Sklar, se puede dar una apertura al desarrollo aplicativo de los modelos cópula. Sea H una función de distribución conjunta de 2 dimensiones con sus distribuciones marginales F_1, F_2 . Además, existe una cópula C tal que para todo $x \in \mathbb{R}^n$,

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \quad (8)$$

Si las distribuciones marginales F_1, F_2 son continuas, entonces C es única, en otro caso, C es determinado por $Rank(F_1)$ y $Rank(F_2)$. En un caso inverso, si C es una cópula y F_1, F_2 son funciones de distribución, entonces la función H definida por la ecuación anterior es una función de distribución conjunta con marginales F_1 y F_2 . (Nelsen, 2006)

En resumen, se pueden vincular cualquier grupo de n distribuciones univariadas con cualquier cópula, obteniendo como resultado una función de distribución multivariada válida.

(Nelsen, 2006)

3.2. Familia de cópulas

A continuación se presentan algunas de las posibles cópulas para el moldeamiento de los activos financieros:

cópula Gaussiana: Se define a la cópula gaussiana como la distribución sobre el espacio medible en el cubo $[0, 1]^d$ construida a partir de la Normal multivariada y dada una matriz de correlación $\Sigma \in \mathbb{R}^{d \times d}$ donde la cópula se escribe de la siguiente manera:

$$C(u) = \Phi_{\Sigma}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)), \quad (9)$$

donde Φ^{-1} es la inversa de la distribución normal estándar Φ_{Σ} ; en matemáticos la densidad de la cópula esta dada por:

$$C(u) = \frac{1}{\sqrt{\det \Sigma}} \left(-\frac{1}{2} \begin{pmatrix} \Phi^{-1}(u_1) \\ \vdots \\ \Phi^{-1}(u_d) \end{pmatrix}^T (\Sigma^{-1} - \mathbf{I}) \begin{pmatrix} \Phi^{-1}(u_1) \\ \vdots \\ \Phi^{-1}(u_d) \end{pmatrix} \right).$$

cópula t: En este caso muy particular, la cópula bivariada 2-dimensional para la distribución t-student multivariada es la siguiente:

$$C_{v,\rho}^t(u, v) = \int_{-\infty}^{t_v^{-1}(u)} \int_{-\infty}^{t_v^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \left\{ 1 + \frac{a^2 - 2\rho ab + b^2}{v(1-\rho^2)} \right\}^{-\frac{(v+2)}{2}} da db, \quad (10)$$

donde t_v^{-1} corresponde a la inversa de la distribución t con v grados de libertad. Si las distribuciones marginales son t de Student con los mismos grados de libertad entonces C es una cópula t con v grados de libertad y correlación ρ por lo que la función $H(x, y)$ es la distribución estándar bivariada centrada, con el coeficiente de correlación lineal ρ y con v grados de libertad.

En términos generales la cópula t puede escribirse como:

$$C_{v,\rho}^t(\mathbf{u}) = \frac{f_{v,P}(t_v^{-1}(u_1), \dots, t_v^{-1}(u_d))}{\prod_{i=1}^d f_v(t_v^{-1}(u_i))}, \quad u \in (0, 1)^d, \quad (11)$$

donde $f_{v,P}$ es la densidad de un vector distribuido $t_d(v, 0, P)$ con v grados de libertad y P como matriz de correlaciones, y f_v como densidad univariada estándar t con v grados de libertad (Nelsen, 2006).

cópolas arquimedianas: Las cópolas arquimedianas son una clase asociativa de cópolas cuya expresión admite una fórmula explícita; estas son populares por permitir modelar la dependencia en dimensiones altas con un solo parámetro. En general, una cópula se dice arquimediana si se expresa como:

$$C(u_1, \dots, u_d; \theta) = \psi^{[-1]}(\psi(u_1; \theta) + \dots + \psi(u_d; \theta); \theta), \quad (12)$$

donde ψ es una función decreciente, continua y convexa perteneciente al intervalo $[0, 1] \times \Theta \rightarrow [0, \infty)$, θ es el parámetro de la cópula con un espacio de parámetros Θ y $\psi^{[-1]}$ es considerado como pseudo inversa de ψ tal que:

$$\psi^{[-1]}(t; \theta) = \begin{cases} \psi^{-1}(t; \theta) & \text{si } 0 \leq t \leq \psi(0; \theta) \\ 0 & \text{si } \psi(0; \theta) \leq t \leq \infty \end{cases} \quad (13)$$

En general las cópulas arquimedianas mas importantes son las siguientes:

Tabla 1: cópulas arquimedianas. Fuente: *Nelsen(2006)*

cópula	Función de distribución	Parámetro
Ali-Mikhail-Haq	$\frac{uv}{1-\theta(1-u)(1-v)}$	$\theta \in [-1, 1]$
Clayton	$[\max\{u^{-\theta} + v^{-\theta} - 1; 0\}]^{-1/\theta}$	$\theta \in [-1, \infty)/0$
Frank	$-\frac{1}{\theta} \log \left[1 + \frac{(\exp(-\theta u)-1)(\exp(-\theta v)-1)}{\exp(-\theta)-1} \right]$	$\theta \in (R)/0$
Gumbel	$\exp \left[-\left((-\log(u))^\theta + (-\log(v))^{1/\theta} \right) \right]$	$\theta \in [1, \infty)$
Joe	$1 - \left[(1-u)^\theta + (1-v)^\theta - (1-u)^\theta(1-v)^\theta \right]^{1/\theta}$	$\theta \in [1, \infty)$

4. Paradigma Bayesiano

El teorema de Bayes que surgió en 1763 como resultado de un trabajo póstumo en una publicación de Richard Price (Peña, 2013), después de 50 de la publicación del tratado de Jacob Bernoulli en donde introduce la ley de los grandes números. El teorema de Bayes se denota con A y B eventos con $P(B) \neq 0$ de la siguiente forma:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \quad (14)$$

Donde $P(A)$ y $P(B)$ son las probabilidades de las observaciones independientes de A y B, conocido como probabilidad marginal, $P(B|A)$ es la probabilidad condicional del evento B dado que A es verdadero y $P(A|B)$ es la probabilidad condicional del evento A dado que B es verdadero.

El teorema de Bayes ha sido empleado para encontrar probabilidades de cierto evento de interés (D'Alessio, 2018), por ejemplo, en encontrar la probabilidad de que una enfermedad sea encontrada en un grupo de personas con una característica dada, tomando las tasas globales de la enfermedad y el predominio de dichas características en personas tanto sanas como enfermas; además, se pueden encontrar aplicaciones de este teorema en el campo de la tecnología denominado *Software basado en conocimiento*, donde esta metodología permite calcular probabilidades que permiten la evaluación de posibles eventos en los software que presenta el usuario y determinar qué consejo proporcionarle para ofrecerle un mejor servicio al mismo

(D'Alessio, 2018). Con lo anterior se puede decir que el teorema puede actuar en diversos campos ayudando así a generar decisiones ante eventuales problemas que se puedan presentar.

La inferencia bayesiana considera que un parámetro θ es una variable aleatoria y se le asigna una distribución inicial o a priori de probabilidades $p(\theta)$; además, se tiene un función de verosimilitud $p(y|\theta)$ que se obtienen del conjunto de datos $y = (y_1, \dots, y_n)$ y dependen del parámetro θ (O'Hagan & Forster, 2004), de esta forma, se tiene que la distribución condicional de θ dados los datos con el teorema de bayes, se denota de la siguiente manera:

$$P(\theta|y) = \frac{P(y|\theta)P(\theta)}{P(y)}. \quad (15)$$

Este teorema es considerado como el método principal de la inferencia bayesiana, con el objetivo de actualizar la distribución a priori con los parámetros, obteniendo así la distribución a posteriori, la cual se considera como parte fundamental de la estimación bayesiana. Después de observar los datos, se tiene que la distribución a posteriori sintetiza toda la información del parámetro, realizando inferencias sobre los datos (O'Hagan & Forster, 2004). La distribución a posteriori permite calcular probabilidades de que los parámetros se encuentren en intervalos de valores dados, denominados intervalos de credibilidad.

Díaz & Batanero (2006) mencionan que una diferencia fundamental de la inferencia bayesiana respecto a la clásica es el carácter subjetivo de las probabilidades, debido a que no se plantea el problema de muestreo ni tampoco el concepto de distribución muestral, con esto, la inferencia bayesiana hace uso de toda la información previa disponible, mientras que la inferencia clásica esta información no la tiene en cuenta.

En este artículo se menciona que el enfoque bayesiano tiene en cuenta la perspectiva del investigador, debido a que este postula una distribución a priori teniendo en cuenta el conocimiento del problema; la elección de la distribución a priori puede condicionar los resultados de la inferencia que distintos investigadores puedan encontrar en sus resultados, sin embargo, se ha sugerido utilizar distribuciones a priori no informativas, actualizando estas a nuevas aplicaciones con los resultados del paso anterior (Díaz & Batanero, 2006).

5. Metodología

La presente investigación se caracteriza por ser un estudio cuantitativo que calcula el Valor en Riesgo de un portafolio comprendido por los activos financieros de Ecopetrol y Bancolombia mediante la estimación bayesiana de las relaciones de dependencia no lineales existentes entre las series financieras del ejercicio.

En primer lugar, se desarrollan los distintos estadísticos descriptivos junto con los gráficos de línea para identificar las características particulares de las series financieras entre las cuales se resalta la volatilidad y el comportamiento leptocúrtico de

los datos. Se realizan las funciones de autocorrelación (FAC - FACP) y las pruebas de hipótesis de raíz unitaria. Seguido, con base en las FAC-FACP se extraen los rezagos significativos para postular los posibles modelos ARMA (Autoregresivos con Media Movil).

después de seleccionar los mejores modelos mediante el criterio de selección AIC y realizar las pruebas de homocedasticidad e independencia sobre los residuales con un $\alpha = 10\%$, se prosigue a modelar la volatilidad esperada (que se evidenciará gráficamente en las *figuras 1 y 2*) mediante la teoría de los modelos GARCH-ARCH para cada uno de los rendimientos de los activos financieros. Tras elegir los modelos que capturan la volatilidad, se realiza la prueba de normalidad sobre los residuales; en particular, si los errores no presentan distribución normal, se postulan las posibles distribuciones de los residuales junto con la estimación de sus correspondientes parámetros calculados mediante la aplicación de la teoría bayesiana, para posteriormente ser convertidos a datos cópula (valores entre 0 y 1).

Se determina la cópula asociada a la estructura de dependencia de los activos financieros por medio del diagrama de dispersión, que contrasta los residuales de los modelos GARCH-ARCH. Se estiman los parámetros de la distribución multivariada cópula por medio de la metodología Markov Chains Montecarlo (MCMC) para generar así un modelo con relaciones de dependencia no lineal que finalmente, se utiliza para estimar el Valor en Riesgo del portafolio teniendo como herramienta principal la varianza del modelo con dependencia lineal, el modelo de volatilidad y el ajuste de la cópula.

6. Aplicación

Descripción de los datos

La teoría presentada se aplica a un portafolio compuesto por las acciones de Bancolombia y Ecopetrol. La base de datos contiene el precio de cierre para cada acción durante 519 días seguidos, teniendo en cuenta fines de semana y días festivos, desde el 30 de octubre de 2015 hasta el 31 de marzo de 2017.

Se procede a realizar las pruebas de raíz unitaria para los rendimientos de Bancolombia y Ecopetrol, las cuales concluyen rechazar la hipótesis nula a un $\alpha = 10\%$ es decir, los retornos de ambas acciones no tienen raíz unitaria, por lo tanto, los rendimientos de Bancolombia y Ecopetrol se consideran series estacionarias, después de realizar las respectivas pruebas; dado esto, se considera la construcción de un modelo para evaluar el comportamiento de las series.

Estimación del modelo

Como se ha comprobado que las series son estacionarias, se procede a calcular el mejor modelo bajo la teoría clásica para los rendimientos de cada activo. Se analizan los correspondientes correlogramas. Se eligen los siguientes modelos bajo criterio AIC para los retornos de los activos:

$$\begin{aligned} Y_t &= \phi_1 Y_{t-1} + Z_t \\ X_t &= \eta_1 X_{t-2} + Z_t, \end{aligned}$$

donde, Y y X corresponde a los rendimientos de Bancolombia y Ecopetrol respectivamente, y ϕ_1, η_1 equivalentes a $0.0554 - 0.081070$. Se concluye que los modelos para los rendimientos son correspondientes a: un $AR(1)$ para Bancolombia y un $AR(2)$ para Ecopetrol.

Validación de supuestos

Se realiza el diagnóstico sobre los residuales de los modelos AR. Se rechazan las hipótesis nulas para las pruebas de normalidad y homocedasticidad, por ende, se confirma la volatilidad de los retornos de los activos, y se concluye que estos no siguen una distribución Normal. Estas pruebas están contenidas en el *Anexo x*.

Con el fin de expresar la varianza condicional o la volatilidad de los residuales de los retornos, se desarrollan los modelos ARCH-GARCH y se obtiene que los modelos que capturan la volatilidad son GARCH(1,1) tanto para Bancolombia como para Ecopetrol. Estos modelos son:

$$\begin{aligned} Y_t &= \phi_1 Y_{t-1} + a + \phi_2 Y_{t-1}^2 + \theta_1 Z_{t-1}^2 \\ Eco_t &= \eta_1 Eco_{t-2} + b + \eta_2 Eco_{t-1}^2 + \theta_2 Z_{t-1}^2, \end{aligned}$$

donde, $\phi_2 = 9.447e - 2$, $\eta_2 = 1.472e - 01$, $\theta_1 = 5.357e - 14$, $\theta_2 = 6.136e - 14$ y a, b son constantes equivalentes a $1.369e - 4$ y $2.653e - 4$.

Finalmente, se aplica la prueba de Ljung-Box con la cual se mide la autocorrelación de los residuales. En esta, no se rechaza la hipótesis nula con $p\text{-value} = 0.1715892$ para Bancolombia y $p\text{-value} = 0.3221554$ para Ecopetrol y se concluye que los residuales son independientes.

Estimación Grados de Libertad

Tras encontrar los modelos que capturan la volatilidad de los retornos, se procede a determinar que distribución siguen los residuales, dado que se rechazó la normalidad. Para ello, se realiza el contraste de los residuales con los cuantiles de la distribución normal, de la siguiente manera:

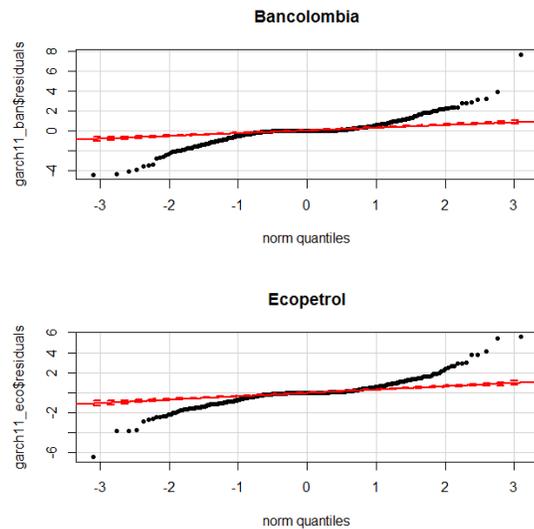


Figura 1: gráficos de Normalidad. Fuente: *Elaboración propia*

En la *Figura 1* se concluye que visualmente los residuales tienen distribución *t-student* dado que se observa colas más pesadas que la distribución normal. Posteriormente, se estiman los grados de libertad de los residuales distribuidos t_v , postulando una distribución prior $y_i|v \sim U(15)$; Haciendo uso del paquete *Jags* de *R*, se realizan diez mil (10000) iteraciones, implementando una (1) cadena de Markov, excluyendo los primeros mil números generados y tomando valores cada 4 saltos. Como resultado se obtiene que la estimación para los grados de libertad en los residuales bajo pérdida cuadrática son aproximadamente 6.880 y 7.108 para Bancolombia y Ecopetrol, respectivamente. A continuación, se muestran los resultados de las estimaciones con sus correspondientes cadenas de Markov, funciones de autocorrelación y densidades aproximadas:

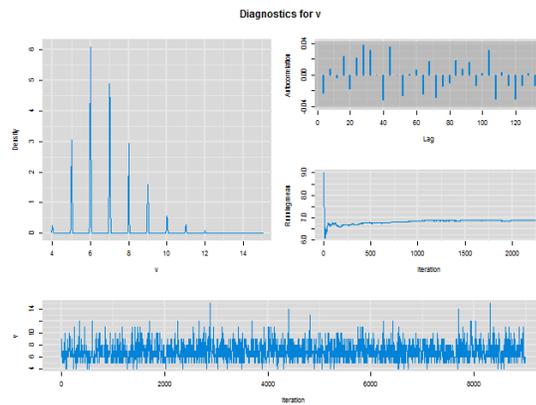


Figura 2: MCMC Bancolombia. Fuente: *Elaboración propia*

En la *Figura 2*, se observa la función de distribución de probabilidad aproximada para los grados de libertad de los residuales de retornos de Bancolombia, de igual manera, se observa que no hay problemas de autocorrelación .

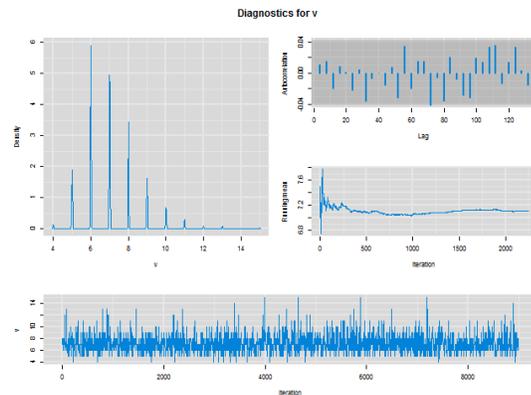


Figura 3: MCMC Ecopetrol. Fuente: *Elaboración propia*

En la *Figura 3*, se evidencia tanto la distribución de probabilidad aproximada, como la independencia de los aleatorios y la convergencia de la cadena para los grados de libertad de los errores de los rendimientos de Ecopetrol.

Como resumen, se presenta la *tabla 2* que muestra los intervalos de credibilidad y las estimaciones para los grados de libertad.

Tabla 2: Resumen de estimaciones. Fuente: *Elaboración propia*

	Bancolombia	Ecopetrol
Estimación	6,880	7,108
Intervalo de credibilidad	(5;10)	(5;11)

De la *tabla 2* se puede concluir que, dados los datos, los grados de libertad de las distribuciones *t* – *student* para los residuales de los retornos de los activos conformados en el portafolio del ejercicio están en el intervalo (5,10) y (5,11) respectivamente con una probabilidad del 95 %.

Al implementar la metodología MCMC (Markov Chain Monte Carlo), se analizan los diagnósticos de convergencia de Heidelberg & Welch, los cuales pasan todas las pruebas de convergencia.

Selección cópula

Conociendo las distribuciones marginales de los residuales en los retornos, se procede a transformarlos a datos cópula, es decir, dejarlos en términos de probabilidades $[0, 1]$ para así seleccionar la cópula que capture la volatilidad de los residuales, los cuales se denotan como u_1 para Bancolombia y u_2 para Ecopetrol. Para ello, se analiza el gráfico de dispersión de los datos cópula de los residuales correspondientes a los rendimientos de Bancolombia contra los de Ecopetrol.

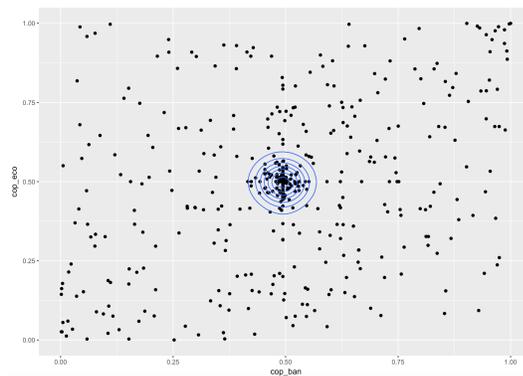


Figura 4: gráfico de Dispersión datos cópula. Fuente: *Elaboración propia*

La postulación de la cópula a implementar se hace a partir del paquete *VineCopulas* de *R* y mediante la función *BiCopSelect* se postula una cópula *t*. Se puede observar en la *figura 4* que la dispersión se ajusta a las características de una cópula

la con distribución t . Por lo tanto, se concluye que la Distribución cópula tiene distribución t bivariada.

Estimación de los parámetros de la Distribución cópula

En primer lugar, se obtiene la función de densidad cópula que corresponde a la segunda derivada con respecto a los parámetros de interés, que se expresa en la siguiente ecuación :

$$f(u_1, u_2, v, \rho) = \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \left(1 + \frac{u_1^2 - 2\rho u_1 u_2 + u_2^2}{v(1-\rho^2)} \right)^{-\frac{v+2}{2}}, \quad (16)$$

donde, u_1 y u_2 son los vectores de datos cópula para los retornos de Bancolombia y Ecopetrol respectivamente, v los grados de libertad de la función de densidad y ρ es la correlación entre u_1 y u_2 .

La ecuación (16), fue tomada de (Parra & Koodi, 2006), expresada como una proporcionalidad de la distribución t bivariada.

De tal manera, se entiende que la función de densidad $f(u_1, u_2, v, \rho)$ corresponde a la verosimilitud utilizada en la metodología Metrópolis-Hastings-Random-Walk (MHRW), con la cual se busca estimar los parámetros de la función de densidad de la cópula. Se postulan las siguientes distribuciones como a priori:

$$v \sim \text{Gamma}_1(10, 100) \quad \rho \sim U(-1, 1). \quad (17)$$

Al calcular la función de densidad posterior, se obtiene:

$$f(v, \rho | u_1, u_2) \propto f(v) f(\rho) f(u_1, u_2 | v, \rho). \quad (18)$$

$$\propto v^{\alpha-1} e^{-\beta v} \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \left(1 + \frac{u_1^2 - 2\rho u_1 u_2 + u_2^2}{v(1-\rho^2)} \right)^{-\frac{v+2}{2}} I_{(-1,1)} u_1, u_2. \quad (19)$$

Implementando la metodología Metropolis Hastings con caminata aleatoria y suponiendo errores normales $e_i \sim N(0, 1)$, se iteran cien mil (100000) datos en donde se obtienen tasas de aceptación de 0.698927 y 0.698927 respectivamente; después, se excluyen los primeros cinco mil (5000) aleatorios y se toman los valores cada cinco saltos. Finalmente, pasan los diagnósticos de convergencia y se tienen 9500 datos aleatorios donde se observan los siguientes resultados en la cadena de Markov, autocorrelación y densidades aproximadas:

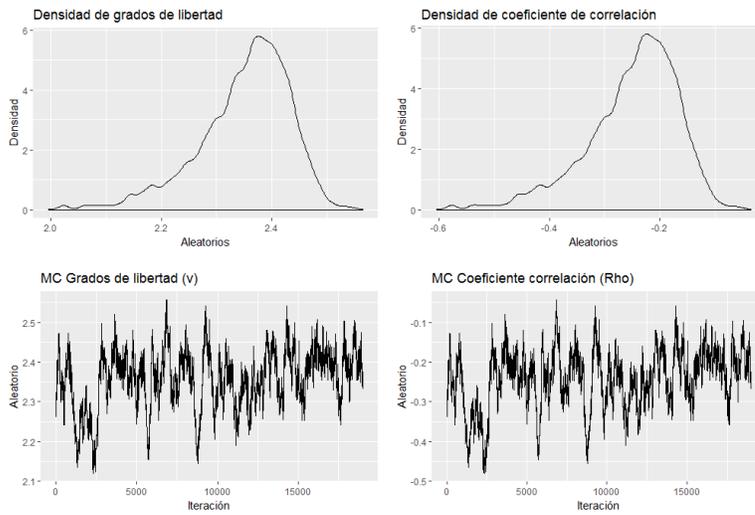


Figura 5: gráficos de densidad y convergencia. Fuente: *Elaboración propia*

En la *figura 5*, se observa la convergencia de la cadena de Markov que se centra en el valor 2.357932 para los grados de libertad v y -0.2420684 para la correlación ρ , al igual que las densidades aproximadas; por otra parte se muestra que no existen problemas de autocorrelación, lo que indica que los aleatorios generados son aproximadamente independientes. El diagnóstico de convergencia de las cadenas generadas por el algoritmo MH-RW también muestran evidencia de convergencia.

Finalmente, al proximar los valores de v y ρ de la t bivariada, se procede a realizar la generación de números aleatorios de la cópula t de dimensión dos con 2.357932 grados de libertad y coeficiente de correlación -0.2420684 ; luego, se generan dos pares de diez mil (10000) valores correspondientes a las posibles pérdidas generadas para el momento $t + 1$ de la cópula t y se convierten en cuantiles de la t con grados de libertad para *bancolombia* 6.879556 y *ecopetrol* 7.091556.

Las posibles pérdidas, se ponderan con la predicción del error ($\sigma_{i,t}$) en el momento $t + 1$ para los modelos Autorregresivos de los retornos de los activos, con las cuales se pondera el error correspondiente a los Z_t de los modelos $X_{1,t}$ y $X_{2,t}$, obteniendo así una distribución de posibles errores sobre la cual se realiza el correspondiente cálculo del VaR.

De forma resumida, se presenta para el cálculo de las predicciones la siguiente forma de los modelos:

$$\begin{aligned}
 X_{1,t} &= \phi_0 + \phi_1 X_{1,t-1} + \epsilon_{2,t} & X_{2,t} &= \phi_0 + \phi_2 X_{1,t-2} + \epsilon_{2,t} \\
 \epsilon_{i,t} &= \sigma_{i,t} \eta_{i,t} & & \\
 \sigma_{i,t}^2 &= \alpha_i + \beta_i \epsilon_{i,t-1}^2 + \gamma_i \sigma_{i,t-1}^2 & &
 \end{aligned}
 \tag{20}$$

donde $\eta_{i,t}$ corresponde a los vectores de simulaciones para ambos activos.

Realizando el cálculo con las expresiones anteriores de la distribución del portafolio luego de realizar las predicciones, se procede a calcular el valor en riesgo al 90 %, 95 % y 99 % de nivel de confianza con dos métodos; el primero, se basa en el cálculo con pesos equilibrados, es decir una ponderación de 0,5 tanto para bancolombia como para ecopetrol. El segundo método se realiza mediante el cálculo de la frontera eficiente del portafolio que resultó ser para bancolombia $w_1 = 0,95$ y $w_2 = 0,5$ para ecopetrol. A continuación, se muestra el cálculo del valor en riesgo:

Tabla 3: Valor en riesgo para el portafolio. Fuente: *Elaboración propia*

$1 - \alpha$	método 1	método 2
90 %	0,05269481	0,02037264
95 %	0,057009117	0,02589244
99 %	0,06873414	0,03882241

7. Discusión

Como resultado de la investigación Estadística presentada, es posible concluir que se encontró la manera óptima para estimar de forma clásica los mejores modelos en series de tiempo, para los retornos de los activos financieros con mayor participación en la bolsa de valores de Colombia correspondientes al portafolio del estudio presentado. De igual manera, fue posible implementar algoritmos con fundamentos de estadísticos bayesiana que permitieron el buen desarrollo y simulaciones de muestras de distribuciones posterior con el cual se obtuvieron las estimaciones de los parámetros correspondientes al modelo propuesto para el posterior cálculo del valor en riesgo del portafolio teniendo en cuenta las relaciones de dependencia no lineales capturadas con un cópula t bivariada y siendo evaluada con diferentes ponderaciones de inversión teniendo en cuenta la frontera eficiente del portafolio.

Recibido: 2017-08-28

Aceptado: 2018-12-11

Referencias

- Albert, J. (2009), *Bayesian computation with R*, Springer Science & Business Media.
- Argáez Sosa, J. (2014), ‘Un paseo por el modelo garch y sus variantes’, *Abstraction & Application* **10**, 35–50.
- Bueno, J. A. G. & Arias, O. P. C. (2017), ‘Comparación de dos portafolios óptimos de renta variable: caso colombia y latinoamérica’, *Puente* **8**(2), 27–41.
- Ceballos López, A. (2015), Implementación de cópulas para la estimación del valor en riesgo, Master’s thesis, Universidad del Rosario.

- Chen, Y.-T. et al. (2002), 'On the robustness of ljung-box and mleod-li q tests: a simulation study', *Economics Bulletin* **3**(17), 1–10.
- Daniélsson, J. (2011), *Financial risk forecasting: the theory and practice of forecasting market risk with implementation in R and Matlab*, Vol. 588, John Wiley & Sons.
- Gallardo Estévez, B. (2010), Teoría de cópulas y aplicaciones en simulación de riesgos financieros y en ingeniería civil, Master's thesis, Universidad de Granada.
- Gamerman, D. & Lopes, H. F. (2006), *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*, Chapman and Hall/CRC.
- Joe, H. (2014), *Dependence modeling with copulas*, Chapman and Hall/CRC.
- Mercado de Renta Variable* (n.d.), <https://www.bvc.com.co/pps/tibco/portalbvc///Home/Mercados/descripciongeneral/acciones>. Accessed: 2017-04-13.
- Nelsen, R. B. (2007), *An introduction to copulas*, Springer Science & Business Media.
- Patton, A. J. (2003), Applications of copula theory in financial econometrics., PhD thesis.
- Qué es el Valor en Riesgo-VaR?* (n.d.), <https://www.bbva.com/es/noticias/economia/ques-el-valor-en-riesgo-var/>. Accessed: 2015-03-23.
- Team, R. C. (2017), 'R: A language and environment for statistical computing. vienna, austria: R foundation for statistical computing; 2016'.