

MEDIDAS DE RIESGO FINANCIERO USANDO CÓPULAS: TEORÍA Y APLICACIONES

ÓSCAR BECERRA

Fedesarrollo
(obecerra@fedesarrollo.org)

LUIS F. MELO

Banco de la República
(lmelovel@banrep.gov.co)

RESUMEN. Este documento realiza una descripción de las medidas de dependencia con sus principales ventajas y desventajas y presenta a la cópula como una estructura flexible que permite caracterizar diferentes tipos de dependencia. Adicionalmente, introduce el uso de la cópula en la medición de riesgo financiero, tomando como ejemplo un portafolio compuesto por tres activos representativos del mercado colombiano.

Las pruebas de desempeño o de *backtesting* del valor en riesgo calculado por diferentes metodologías en los años 2006 y 2007 muestran que las mejores son aquellas que modelan la dependencia en media y varianza, tales como modelos VAR-GARCH-Cópula(t) y VAR-GARCH-Cópula(*normal*). Las técnicas con el peor desempeño son Riskmetrics® y la basada en el supuesto de normalidad.

Palabras Clave. Dependencia, cópula, riesgo de mercado, riesgo de crédito, métodos de simulación de Monte Carlo.

Clasificación JEL. C32, C52, G10.

Esta versión: 14 de febrero de 2008.

Los resultados y opiniones son responsabilidad exclusiva de los autores y su contenido no compromete al Banco de la República ni a su junta directiva.

ÍNDICE

1. Introducción	1
2. Dependencia en finanzas	1
3. Medidas de dependencia	2
3.1. Coeficiente de correlación de Pearson	3
3.2. Cópula	6
3.3. Otras medidas de dependencia.	12
3.4. Conceptos de dependencia.	14
3.5. Familias de cópulas	19
3.6. Simulación	27
4. Estimación.	32
4.1. Métodos de estimación.	33
4.2. Verificación de supuestos.	37
4.3. Selección de la cópula.	38
4.4. Estimación de cópulas en presencia de dependencia temporal.	43
4.5. Modelación de valores extremos a través de cópulas.	48
5. Aplicación de la cópula en modelos de riesgo.	56
5.1. Modelos de riesgo de mercado.	56
5.2. Modelos de riesgo de crédito.	63
6. Una aplicación de las cópulas a la valoración del riesgo de mercado	66
6.1. Pruebas de desempeño (<i>Backtesting</i>) del VaR_α y el ES_α .	71
6.2. Otras aplicaciones del VaR_α	76
7. Comentarios finales	77
Referencias	79
Anexo A. Distribuciones esféricas y elípticas	84
A.1. Distribuciones esféricas	84
A.2. Distribuciones elípticas	85
A.3. Pruebas para distribuciones elípticas	86
Anexo B. Pruebas de especificación sobre el modelo VAR-GARCH	89
Anexo C. Cópulas comúnmente mencionadas en la literatura	90

1. INTRODUCCIÓN

En cualquier economía, la adopción de métodos más exactos en la medición de riesgo genera garantías para la estabilidad del sistema financiero. Una de las medidas de mayor uso en este contexto, el VaR_{α} , corresponde al α -ésimo cuantíl de la distribución de pérdidas y ganancias de un activo. Esta medida representa la máxima pérdida en que incurre un activo en el $\alpha \times 100\%$ mejor de los casos¹. Para el caso colombiano, el VaR tiene implicaciones directas en el cálculo de la relación de solvencia de las entidades financieras, un VaR más alto representa una relación de solvencia más baja, lo cual obliga a las entidades a reasignar sus activos riesgosos o a realizar aportes adicionales de capital.

Sin embargo, no existe una metodología única para el cálculo del VaR , más aún cuando se consideran las interacciones existentes entre diferentes factores de riesgo (activos) al interior de un portafolio. Este documento generaliza las metodologías analizadas en Melo y Becerra [2006] para que puedan ser evaluadas en un portafolio.

En este contexto, existen dos clases de dependencia: longitudinal (a través del tiempo) y transversal (entre activos). La longitudinal puede ser modelada por técnicas tradicionales de series de tiempo, mientras que la transversal o contemporánea se puede analizar mediante modelos de *dependencia* como los de **cópula**.

La cópula, en líneas generales, es una función que aproxima el comportamiento conjunto de variables aleatorias a partir de sus comportamientos individuales. Este documento se centra en la definición, implicaciones, técnicas para evaluar desempeño y en general, metodologías utilizadas en el desarrollo de modelos de medición de riesgo asociados con cópulas.

El resto del documento se compone de la siguiente manera: en la segunda sección se comenta el concepto de dependencia en finanzas. En la tercera sección, se definen algunas medidas y conceptos de dependencia y se introduce el concepto de cópula. En la cuarta sección se presentan metodologías de estimación, verificación de supuestos y métodos de selección de cópulas. En la quinta sección se muestra la aplicación de las cópulas a modelos de riesgo de mercado y de crédito. Posteriormente, en la sexta sección se presentan los resultados de la estimación y se evalúa el desempeño para diferentes modelos de valor en riesgo y *Expected Shortfall* sobre un portafolio compuesto por el IGBC, la TRM y un título genérico con vencimiento de 10 años (TES). Finalmente, en la séptima sección se concluye.

2. DEPENDENCIA EN FINANZAS

En finanzas, existen gran cantidad de factores de riesgo que interactúan constantemente dentro de un portafolio. En consecuencia, se han desarrollado herramientas apropiadas para modelar dicha interacción, en este plano el concepto de dependencia juega un papel fundamental.

¹Bajo este contexto, los mejores casos indican las pérdidas más pequeñas.

Dado un conjunto de factores de riesgo, un administrador puede identificar los comportamientos marginales a través de técnicas univariadas apropiadas, sin embargo, es importante tener en cuenta que el comportamiento de un factor de riesgo puede estar influenciado por comportamientos de otros factores de riesgo. Por lo tanto, las técnicas que modelan la dependencia entre los factores de riesgo constituyen una fuente valiosa de información.

Existen diferentes medidas de dependencia, algunas más conocidas que otras. Una de las más conocidas en el área financiera es el *coeficiente de correlación de Pearson*, el cual ha desempeñado un papel importante en modelos clásicos de asignación de portafolio tales como el modelo *CAPM*, *Capital Asset Pricing Model*, y el *APT*, *Arbitrage Pricing Theory*, donde se usa como medida de dependencia entre diferentes instrumentos financieros. No obstante, el coeficiente de correlación es una medida apropiada únicamente bajo ciertos supuestos sobre la función de distribución multivariada de los retornos. Un ejemplo de ello es el problema de optimización de portafolio de Markowitz [1952]:

$$\begin{aligned} & \underset{\omega}{\text{máx}} \quad \omega' \mu \\ & \text{s.a.} \quad \begin{cases} \omega' \mathbf{1} = 1 \\ \omega' \Sigma \omega \leq s \\ \omega_i \geq 0 \quad i = 1, \dots, d \end{cases} \end{aligned} \quad (2.1)$$

donde $\omega = (\omega_1, \dots, \omega_d)'$ es el vector de participaciones de cada uno de los d activos dentro del portafolio, μ es el vector de retornos esperados de cada uno de los activos y Σ es la matriz de varianzas y covarianzas. En este caso la varianza corresponde a la medida de riesgo, sin embargo, es posible mostrar que esta medida es apropiada únicamente cuando los activos siguen una distribución normal multivariada [Bouyé et al., 2000] o de forma más general, cuando su función de distribución es elíptica²[Embrechts et al., 1999a].

Una función que representa de manera más general las relaciones de dependencia existentes entre vectores de variables aleatorias es la cópula, la cual relaciona el comportamiento de las distribuciones univariadas de cada uno de los factores de riesgo a su distribución multivariada. A partir de la cópula es posible construir otras medidas de dependencia, tales como los coeficientes de correlación de rango y el coeficiente de dependencia en las colas, cuyas definiciones se presentarán más adelante, e identificar características propias de cada una de ellas.

3. MEDIDAS DE DEPENDENCIA

En la práctica, el coeficiente de correlación lineal de Pearson es la medida de dependencia más usada. Sin embargo, como se muestra posteriormente, esta medida presenta ciertos inconvenientes. A continuación se describen varias definiciones de medidas y conceptos de dependencia y sus implicaciones.

²Una introducción al tema relacionado con las distribuciones elípticas y esféricas es realizada en el Anexo A.

3.1. Coeficiente de correlación de Pearson. Una de las aproximaciones más utilizadas para analizar relaciones de dependencia es el coeficiente de correlación de Pearson, el cual mide la relación lineal existente entre un vector de variables aleatorias. Para la pareja de variables aleatorias (R_i, R_j) el coeficiente de correlación de Pearson, ρ_{ij} , está definido por:

$$\rho_{ij} = \frac{\text{cov}(R_i, R_j)}{\sqrt{V(R_i)V(R_j)}} \quad i, j = 1, 2, \dots, d \quad (3.1)$$

Para un vector aleatorio $R = [R_1, \dots, R_d]'$ de dimensión d , la matriz de correlaciones, \mathbf{P} , está dada por:

$$\mathbf{P} = \begin{bmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1d} \\ \rho_{21} & \rho_{22} & \dots & \rho_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{d1} & \rho_{d2} & \dots & \rho_{dd} \end{bmatrix} \quad (3.2)$$

Algunas propiedades del coeficiente de correlación definido en (3.1) son [Blanco, 2004]:

- i)* $\rho(R_i, R_j) = \rho(R_j, R_i)$.
- ii)* $|\rho(R_i, R_j)| \leq 1$.
- iii)* $\rho(R_i, R_i) = 1$, $\rho(R_i, -R_i) = -1$.
- iv)* $\rho(aR_i + b, R_j) = \rho(R_i, R_j)$, $a > 0$.
- v)* $|\rho(R_i, R_j)| = 1$ si y sólo si existen constantes $a, b, c \in \mathbb{R}$, con $a, b \neq 0$ tal que $P(aR_i + bR_j = c) = 1$.

El coeficiente de correlación de Pearson tiene diferentes ventajas, las cuales lo han hecho una medida de dependencia de amplia aceptación en muchos contextos incluido el financiero, principalmente porque es fácil de calcular, ya que sólo se necesita estimar los dos primeros momentos de los datos observados. Además es invariante ante transformaciones afines positivas (propiedad *iv*). Finalmente, una de las grandes ventajas que tiene el coeficiente de correlación de Pearson es su relación con la función de distribución normal multivariada, en donde resume toda la relación de dependencia existente entre las variables aleatorias [Embrechts et al., 1999a].

Sin embargo, es necesario ser cauteloso cuando se utiliza el coeficiente de correlación como medida de dependencia, ya que puede presentar ciertos inconvenientes. Algunos de estos inconvenientes son [Embrechts et al., 1999]:

- *El coeficiente de correlación es una medida de dependencia escalar.* Este problema no se encuentra asociado únicamente al coeficiente de correlación. Al ser una medida escalar, éste resume toda la relación de dependencia entre los factores de riesgo en un solo número. Aunque esto facilita la interpretación, puede tener inconvenientes en la toma de decisiones.

Por ejemplo, suponga que existen dos activos con los cuales se va a constituir un portafolio, cada uno de los cuales sigue una distribución normal estándar y tienen un coeficiente de correlación de 0,8. Aunque el valor del coeficiente de correlación implica que existe una relación de dependencia lineal entre los dos activos, no indica como es la estructura de dependencia entre estos. El Gráfico 3.1 muestra como es posible tener dos variables aleatorias que satisfacen estas características (comportamientos marginales normales estándar y coeficiente de correlación de 0,8) y aún así llevar a conclusiones completamente diferentes.

En el panel superior izquierdo del Gráfico 3.1 se encuentra una estructura de dependencia normal multivariada, donde la mayoría de las observaciones se concentran en el centro de la distribución y los datos extremos no presentan una relación clara. Los demás gráficos muestran otros tipos de relaciones entre estos dos activos, en las que los retornos extremos poseen relaciones de dependencia diferentes a la de la normal.

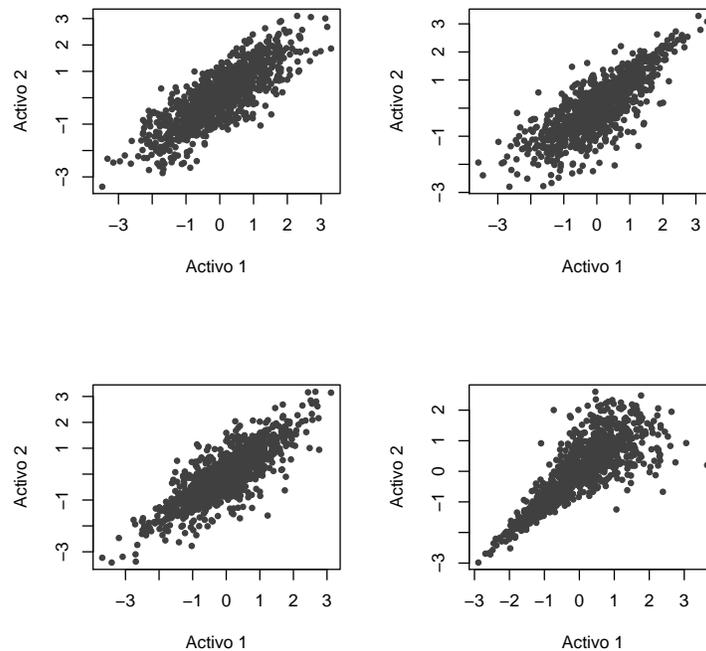


GRÁFICO 3.1. Simulación de 5,000 realizaciones de dos variables aleatorias normal estándar con coeficiente de correlación 0,8 y cuatro estructuras de dependencia diferentes.

- Riesgos que sean perfectamente dependientes no implican que $|\rho(R_i, R_j)| = 1$. Tal como lo muestran McNeil et al. [2005], dependiendo del comportamiento marginal que tengan

los factores de riesgo, los valores del coeficiente de correlación se encontrarán en el intervalo $[\rho_{\min}, \rho_{\max}]$ donde $-1 \leq \rho_{\min} \leq 0 \leq \rho_{\max} \leq 1$. Esto tiene particular importancia en la toma de decisiones, ya que es posible que coeficientes de correlación lejanos a uno, impliquen un alto grado de asociación entre dos factores de riesgo.

- *Que el coeficiente de correlación sea cero no implica independencia entre los factores de riesgo.* Esto únicamente es cierto para el caso normal multivariado. Para cualquier otro tipo función de distribución multivariada la correlación cero no necesariamente indica independencia entre los factores de riesgo. Un claro ejemplo de esto se muestra en el Gráfico 3.2, donde se presentan los valores de una variable aleatoria normal estándar y estos mismos valores al cuadrado. A simple vista se ve una clara relación entre estas dos variables, sin embargo, el coeficiente de correlación es cero ya que $Cov(R_i, R_i^2) = E(R_i^3) = 0$, el tercer momento de una variable aleatoria normal.

El inconveniente del coeficiente de correlación reseñado anteriormente se encuentra ligado al tipo de dependencia que éste puede describir. El coeficiente de correlación indica el grado de asociación lineal que tienen dos variables aleatorias, y en general, éste se encuentra limitado en situaciones donde el grado de asociación no es lineal.

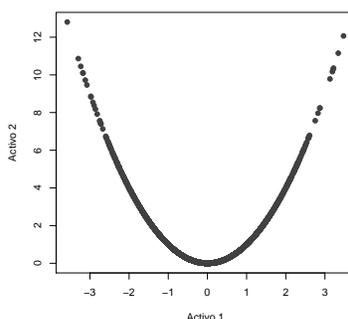


GRÁFICO 3.2. Simulación de 5,000 realizaciones de una variable normal aleatoria estándar y su cuadrado.

- *El coeficiente de correlación es invariante únicamente ante transformaciones afines positivas³.* Esta característica es de gran importancia, ya que por ejemplo no es lo mismo calcular el coeficiente de correlación sobre los retornos logarítmicos (compuestos) que sobre los precios de dos activos. Esto implica que la toma de decisiones asociada a este tipo de análisis puede estar influenciada por la metodología a través de la cual se maneja la información.

Por ejemplo, para el caso de las variables aleatorias que siguen la distribución normal multivariada del Gráfico 3.1, los valores del coeficiente de correlación para algunas transformaciones monótonas crecientes se presentan en la Tabla 3.1. En ésta se observa

³Una transformación afín positiva T sobre R_i está definida de la siguiente forma: $T(R_i) = aR_i + b$, $a > 0$.

que el coeficiente de correlación cambia para todos los casos en que la transformación es una transformación no afín positiva.

Transformación	$\hat{\rho}$
$\rho(R_1, R_2)$	0,80
$\rho(e^{R_1}, e^{R_2})$	0,73
$\rho(R_1^3, R_2^3)$	0,70
$\rho(R_1, R_2^5)$	0,39
$\rho(3R_1, R_2)$	0,80

TABLA 3.1. Valores del coeficiente de correlación ante diferentes transformaciones de 5.000 realizaciones de dos variables aleatorias normal estándar que siguen una distribución normal multivariada.

- *El coeficiente de correlación únicamente se encuentra definido para parejas de variables aleatorias que tengan varianzas finitas.* Esta desventaja es de especial importancia cuando se trata de analizar factores de riesgo que tengan funciones de distribución con colas pesadas, ya que en estos la varianza puede no ser finita [Embrechts et al., 1997]. Por lo tanto, el coeficiente de correlación no es una medida adecuada para estos casos.

3.2. Cópula. Para evitar los problemas anteriormente mencionados, es posible modelar la dependencia entre d factores de riesgo a través de una función que satisface algunas propiedades particulares que la hacen idónea para cumplir esta tarea. Esta aproximación tiene la ventaja de presentar el concepto de dependencia como una estructura que describa completamente la relación entre los factores de riesgo, en lugar de tratar de resumirla en un sólo número. No obstante, pensar en una estructura de este tipo implica una mayor complejidad en su interpretación.

La función mencionada en el párrafo anterior es conocida como cópula, la cual es una función de distribución multivariada $C : [0, 1]^d \rightarrow [0, 1]$ donde sus distribuciones marginales (individuales) son uniformes estándar $U_i \sim U(0, 1)$ $i = 1, 2, \dots, d$. La cópula posee las siguientes propiedades:

- i) $C(u_1, \dots, u_d)$ es creciente en cada componente u_i .
- ii) $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$ para todo $i \in \{1, 2, \dots, d\}$, $u_i \in [0, 1]$.
- iii) Para todo $(a_1, \dots, a_d), (b_1, \dots, b_d) \in [0, 1]^d$ con $a_i \leq b_i$ se tiene que:

$$\sum_{i_1=1}^2 \dots \sum_{i_d=1}^2 (-1)^{i_1+\dots+i_d} C(u_{1i_1}, \dots, u_{di_d}) \geq 0$$

donde $u_{j1} = a_j$ y $u_{j2} = b_j$ para todo $j \in \{1, 2, \dots, d\}$.

iv) Para cada cópula $C(u_1, \dots, u_d)$ se tienen los siguientes límites:

$$\max \left\{ \sum_{i=1}^d u_i + 1 - d, 0 \right\} \leq C(\mathbf{u}) \leq \min\{u_1, \dots, u_d\}$$

Las propiedades *i)*, *iii)* y *iv)* son propias de cualquier función de distribución multivariada, mientras que la propiedad *ii)* esta asociada con el hecho que las funciones de distribución marginales son uniformes estándar [Joe, 1997; McNeil et al., 2005]. Adicionalmente, los límites establecidos en la propiedad número *iv)* son conocidos como los límites de Fréchet, los cuales se tratarán en la Sección 3.4.

Sklar [1959] mostró que es posible comprender la estructura de dependencia de un vector aleatorio a través de la cópula a partir de dos observaciones: la primera, que toda función de distribución multivariada de variables aleatorias continuas tiene una cópula asociada y la segunda, que la función de distribución multivariada puede ser construida a partir de la cópula y las funciones de distribución marginales.

Teorema 3.1 (Teorema de Sklar). *Sea $F(r_1, \dots, r_d)$ la función de distribución conjunta de las variables aleatorias R_1, \dots, R_d , con funciones de distribución marginales $F_1(r_1), \dots, F_d(r_d)$. Existe una función $C : [0, 1]^d \rightarrow [0, 1]$ tal que:*

$$F(r_1, \dots, r_d) = C(F_1(r_1), \dots, F_d(r_d)) \quad (3.3)$$

para todo r_1, \dots, r_d . La función C es conocida como la "cópula" de variables aleatorias. Si $F_1(r_1), \dots, F_d(r_d)$ son continuas entonces C es única. De otra forma, C está definida únicamente en el $\text{Ran}(F_1) \times \dots \times \text{Ran}(F_d)$ donde $\text{Ran}(F_i)$ denota el rango de la función F_i . Análogamente, si C es una cópula y $F_1(r_1), \dots, F_d(r_d)$ son funciones de distribución univariadas, entonces $F(r_1, \dots, r_d)$ definida en (3.3) es una función de distribución conjunta con marginales $F_1(r_1), \dots, F_d(r_d)$.

Para la demostración del Teorema 3.1 es necesario tener presentes varias características de las funciones de distribución⁴. La primera de ellas se relaciona con el hecho que si la variable aleatoria R_i tiene una función de distribución F_i continua, entonces se tiene que:

$$F_i(R_i) \sim U(0, 1) \quad (3.4)$$

Análogamente, si $U \sim U(0, 1)$ entonces $F_i^{-1}(U) \sim F_i$, donde F_i^{-1} es la función de distribución inversa de F_i , la cual esta definida como:

$$F_i^{-1}(\alpha) = \inf \{r \in \mathbb{R} : F_i(r) \geq \alpha\} \quad (3.5)$$

El resultado presentado en el párrafo anterior se encuentra en la mayoría de los libros de probabilidad básica y es el fundamento para la generación de números aleatorios.

A partir de (3.4) es posible demostrar la existencia y unicidad de la cópula para el caso de variables aleatorias continuas. Expresando (3.3) en términos de $r_i = F_i^{-1}(u_i)$ para

⁴La demostración completa del teorema de Sklar se encuentra en Nelsen [2006]

$i = 1, \dots, d$ se obtiene:

$$C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)) \quad (3.6)$$

La expresión (3.6) implica que una cópula C puede ser representada en términos de la función de distribución multivariada F y sus distribuciones marginales F_1, \dots, F_d . Por otro lado, partiendo de F, F_1, \dots, F_d y $u_i = F(r_i)$ para $i = 1, \dots, d$, se tiene que:

$$\begin{aligned} F(r_1, \dots, r_d) &= P(R_1 \leq r_1, \dots, R_d \leq r_d) \\ &= P(F_1^{-1}(R_1) \leq F_1^{-1}(r_1), \dots, F_d^{-1}(R_d) \leq F_d^{-1}(r_d)) \\ &= P(U_1 \leq u_1, \dots, U_d \leq u_d) \\ &= C(u_1, \dots, u_d) \end{aligned} \quad (3.7)$$

Por lo tanto, la definición (3.6) implica que la cópula C asociada a la distribución multivariada F si existe, mientras el resultado (3.7) implica que esta cópula es única [McNeil et al., 2005].

Las ecuaciones (3.3) y (3.7) muestran la relación entre la cópula y la función de distribución conjunta. El teorema de Sklar indica, por un lado, que la función de distribución conjunta puede ser entendida como la combinación de distribuciones marginales a través de la cópula; mientras que (3.7) muestra como se puede expresar la cópula de la función de distribución multivariada [McNeil et al., 2005].

Tal como lo indican Embrechts et al. [1999], el Teorema 3.1 también se puede interpretar como la descomposición de la función de distribución multivariada en dos componentes: la estructura de dependencia y el comportamiento univariado. El primer componente corresponde a la cópula, mientras que el segundo esta asociado a las funciones de distribución marginales.

Uno de los ejemplos más intuitivos acerca de la cópula se encuentra cuando hay independencia entre variables aleatorias. Si R_1 y R_2 son dos variables aleatorias independientes, se tiene que su función de distribución conjunta es igual a:

$$F(r_1, r_2) = F_1(r_1) F_2(r_2) \quad (3.8)$$

Es decir, en el caso de independencia la función de distribución conjunta es igual al producto de las funciones de distribución marginales. A partir de las definiciones presentadas en (3.4) y (3.8) se tiene que:

$$\begin{aligned} F(r_1, r_2) &= P(R_1 \leq r_1, R_2 \leq r_2) \\ &= P(F_1^{-1}(R_1) \leq F_1^{-1}(r_1), F_2^{-1}(R_2) \leq F_2^{-1}(r_2)) \\ &= P(U_1 \leq u_1, U_2 \leq u_2) \\ &= u_1 u_2 \\ &= C^I(u_1, u_2) \end{aligned} \quad (3.9)$$

Donde U_1 y U_2 son dos variables aleatorias que siguen una distribución uniforme entre cero y uno. La función presentada en (3.9) es conocida como la cópula de independencia. De forma general, para d variables aleatorias la cópula de independencia es:

$$C^I(u_1, \dots, u_d) = \prod_{j=1}^d u_j \quad (3.10)$$

La cópula es una función creciente que se encuentra en el espacio $[0, 1]^{d+1}$. El Gráfico 3.3 muestra la cópula de independencia y sus curvas de nivel para el caso bivariado, ecuación (3.9), mientras el Gráfico 3.4 muestra las cópulas utilizadas para las simulaciones del ejemplo mostrado en el Gráfico 3.1.

Aunque la diferencia entre estas cópulas no es tan evidente en los gráficos en tres dimensiones, los gráficos de curvas de nivel ofrecen cierta intuición acerca de la dependencia en cada uno de los casos. Dado que todas las variables aleatorias simuladas en el ejemplo tienen dependencia positiva, las curvas de nivel de las cópulas se encuentran más a la izquierda con respecto de las curvas de nivel de la cópula de independencia (ver Sección 3.4.3).

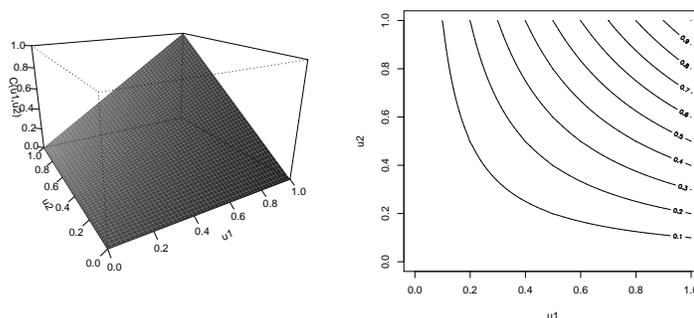


GRÁFICO 3.3. Cópula de independencia

3.2.1. Ventajas de la Cópula. Tal como se mostró en la Sección 3.1, el coeficiente de correlación de Pearson presenta algunos inconvenientes cuando las variables aleatorias no satisfacen el supuesto de normalidad multivariada, o como se discutirá más adelante, cuando las variables no siguen una distribución elíptica. La cópula supera algunos de estos inconvenientes tal como se describe a continuación.

- Dado que la cópula extrae la estructura de dependencia de la función de distribución multivariada, ésta contiene mucha más información acerca de la dependencia entre d variables aleatorias que la que puede contener un sólo número.

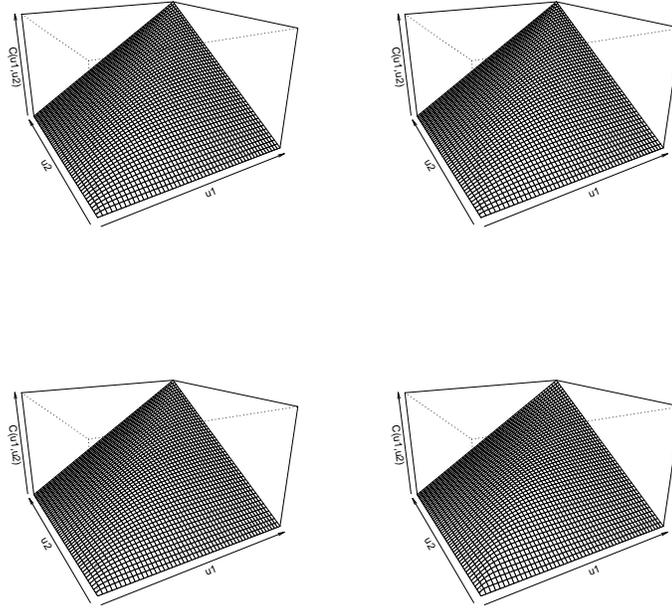


GRÁFICO 3.4. Cópulas usadas para las simulaciones del Gráfico 3.1

- La cópula tiene en cuenta todos los posibles casos de dependencia. Si existe dependencia perfecta positiva entre las variables aleatorias de interés, se dice que las variables aleatorias son “comonotónicas”, por su parte, cuando la dependencia es perfecta negativa, se dice que las variables son “contramonotónicas” (Sección 3.4). En ambos casos, estas situaciones pueden ser descritas por una cópula específica. Adicionalmente, cuando las variables aleatorias son independientes, su relación se resume en la cópula de independencia.
- La cópula es invariante ante transformaciones monótonas crecientes, incluyendo las transformaciones afines positivas. Sean R_1, \dots, R_d variables aleatorias con funciones de distribución marginales F_1, \dots, F_d y cópula C . Si T_1, \dots, T_d son transformaciones monótonas crecientes, esta propiedad implica que R_1, \dots, R_d y $T_1(R_1), \dots, T_d(R_d)$ tienen la misma cópula C . En efecto, si a R_i se le aplica una transformación monótona creciente T_i , entonces la función de distribución de $T_i(R_i)$ es:

$$\begin{aligned}
 P(T_i(R_i) \leq \tilde{r}_i) &= P(R_i \leq T^{-1}(\tilde{r}_i)) \\
 F_i \circ T^{-1}(\tilde{r}_i) &= \tilde{F}_i(\tilde{r}_i)
 \end{aligned}
 \tag{3.11}$$

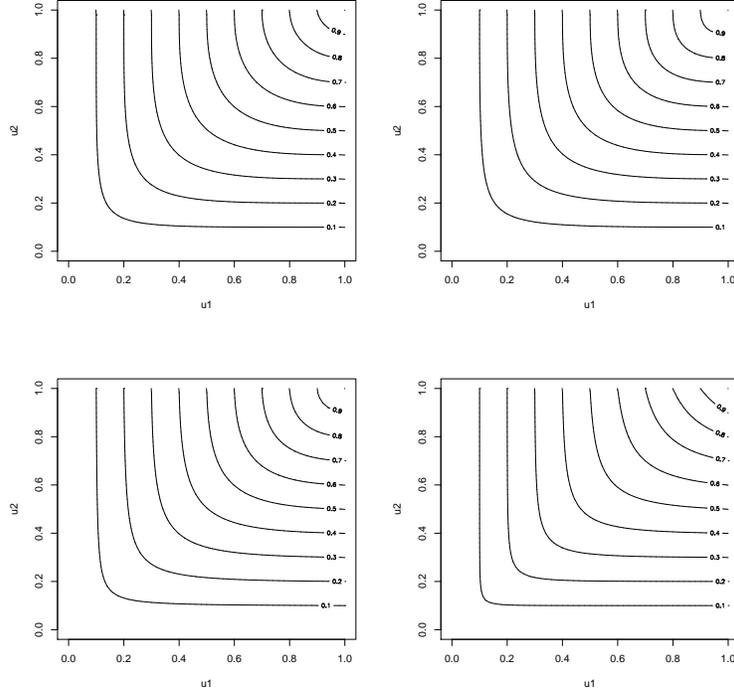


GRÁFICO 3.5. Curvas de nivel de las cópulas usadas para las simulaciones del Gráfico 3.1

donde $\tilde{r}_i \in \text{Rango}(T_i)$. Es posible mostrar que la función de distribución inversa de \tilde{F}_i es:

$$\tilde{F}_i^{-1} = T_i \circ F_i^{-1} \quad (3.12)$$

aplicando la definición (3.6) se tiene que:

$$\begin{aligned} C(u_1, \dots, u_d) &= P(R_1 \leq F_1^{-1}(u_1), \dots, R_d \leq F_d^{-1}(u_d)) \\ &= P(T_1(R_1) \leq T_1 \circ F_1^{-1}(u_1), \dots, T_d(R_d) \leq T_d \circ F_d^{-1}(u_d)) \end{aligned} \quad (3.13)$$

ya que T_1, \dots, T_d son transformaciones monótonas crecientes. Usando los resultados (3.11) y (3.12) sobre (3.13) se obtiene:

$$\begin{aligned} C(u_1, \dots, u_d) &= P\left(T_1(R_1) \leq \tilde{F}_1^{-1}(u_1), \dots, T_d(R_d) \leq \tilde{F}_d^{-1}(u_d)\right) \\ &= P\left(\tilde{F}_1(T_1(R_1)) \leq \tilde{F}_1\left(\tilde{F}_1^{-1}(u_1)\right), \dots, \tilde{F}_d(T_d(R_d)) \leq \tilde{F}_d\left(\tilde{F}_d^{-1}(u_d)\right)\right) \\ &= P\left(\tilde{F}_1(T_1(R_1)) \leq u_1, \dots, \tilde{F}_d(T_d(R_d)) \leq u_d\right) \\ &= C(u_1, \dots, u_d) \end{aligned} \quad (3.14)$$

Con lo cual se concluye que R_1, \dots, R_d y $T_1(R_1), \dots, T_d(R_d)$ tienen la misma cópula C . Esta propiedad muestra que la estructura de dependencia es invariante ante traslaciones, ya que las que cambian en estos casos son las distribuciones marginales [McNeil et al., 2005].

- Finalmente, ya que la cópula representa la estructura de dependencia independientemente de las funciones de distribución marginales, ésta puede tratar con múltiples especificaciones de funciones de distribución, inclusive aquellas que no tienen definidos algunos de sus momentos.

3.3. Otras medidas de dependencia. Teniendo en cuenta los problemas de la correlación de Pearson, es útil definir otras medidas de dependencia que no tengan algunas de estas limitaciones. Una clase importante de estas estadísticas son las medidas de correlación de rango y los coeficientes de dependencia en las colas, los cuales pueden ser expresados en términos de cópulas. Es así como en la Sección 4.1.2 se muestra que las medidas de correlación de rango pueden ser utilizados en la parametrización de las cópulas.

3.3.1. Medidas de correlación de rango. Estas estadísticas miden el grado de dependencia entre dos variables aleatorias sin suponer que la relación entre las variables es lineal, además, estas correlaciones requieren solamente que las variables sean medidas a nivel ordinal. Como su nombre lo indica, la estimación empírica de las medidas de correlación de rango puede ser calculada de forma no paramétrica a partir de los rangos de los datos. Sin embargo, éstas pueden ser obtenidas en función de la cópula asociada a la función de distribución conjunta de las variables aleatorias. Este resultado implica que las medidas de correlación de rango son invariantes bajo transformaciones monótonas crecientes.

Sean R_1 y R_2 dos variables aleatorias con funciones de distribución marginales F_1 y F_2 y función de distribución conjunta F . Entonces, el coeficiente de correlación de rango de Spearman está dado por:

$$\rho_S(R_1, R_2) = \rho(F_1(R_1), F_2(R_2)) \quad (3.15)$$

donde ρ es el coeficiente de correlación de Pearson. Por lo tanto, ρ_S es simplemente la correlación lineal entre las variables aleatorias R_1 y R_2 transformadas a través de sus funciones de distribución.

Sean (R_{11}, R_{21}) y (R_{12}, R_{22}) dos parejas de variables aleatorias independientes con una función de distribución bivariada F , el coeficiente de correlación de rango de Kendall (o τ de Kendall) es el siguiente:

$$\rho_\tau(R_1, R_2) = P((R_{11} - R_{12})(R_{21} - R_{22}) > 0) - P((R_{11} - R_{12})(R_{21} - R_{22}) < 0) \quad (3.16)$$

De acuerdo con la definición (3.16), el τ de Kendall puede entenderse como la diferencia entre la probabilidad de concordancia menos la probabilidad de discordancia. En este contexto, dos variables aleatorias son concordantes si el sentido de los cambios en una variable tiende a estar asociado con el sentido de los cambios de la otra. Es decir, cambios positivos (negativos) de una variable se encuentran asociados con cambios positivos

(negativos) de la otra. De forma análoga, dos variables aleatorias son discordantes si cambios positivos (negativos) de una de ellas tienden a estar asociados con cambios negativos (positivos) de la otra.

A pesar que las propiedades estadísticas de las medidas de correlación de rango son más difíciles de obtener que las de la correlación de Pearson, se ha demostrado que tienen varias propiedades deseables para una medida de dependencia. Son simétricas, definidas en el intervalo $[-1, 1]$ y toman el valor de cero bajo independencia⁵.

Las principales ventajas de las medidas de correlación de rango con respecto a la correlación de Pearson son la invarianza bajo transformaciones monótonas crecientes y su capacidad para identificar los casos en que existe dependencia perfecta. Esto último se encuentra asociado al concepto de comonotonidad, el cual será discutido en la Sección 3.4.1. La propiedad de invarianza se puede apreciar en el ejercicio presentado en la Tabla 3.2, donde se muestran los coeficientes de correlación de rango y de Pearson para 5.000 realizaciones de dos variables aleatorias normales estándar.

Transformación	$\hat{\rho}$ (Pearson)	$\hat{\rho}_\tau$ (Kendall)	$\hat{\rho}_S$ (Spearman)
$\text{corr}(R_1, R_2)$	0,80	0,59	0,78
$\text{corr}(e^{R_1}, e^{R_2})$	0,73	0,59	0,78
$\text{corr}(R_1^3, R_2^3)$	0,70	0,59	0,78
$\text{corr}(R_1, R_2^5)$	0,39	0,59	0,78
$\text{corr}(3R_1, R_2)$	0,80	0,59	0,78

TABLA 3.2. Valores de diferentes coeficientes de correlación ante varias transformaciones de 5.000 realizaciones de dos variables aleatorias normales estándar que siguen una distribución normal multivariada

En el caso de funciones marginales continuas, las medidas de correlación de rango se pueden definir en términos de la cópula C asociada a F , F_1 y F_2 . Al respecto Nelsen [2006] demuestra los siguientes resultados:

$$\rho_S(R_1, R_2) = 12 \int_0^1 \int_0^1 (C(u_1, u_2) - u_1 u_2) du_1 du_2 = 12 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2 - 3 \quad (3.17)$$

$$\rho_\tau(R_1, R_2) = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1 \quad (3.18)$$

3.3.2. Coeficiente de dependencia en las colas. Una de las principales ventajas que ofrece el estudio de las funciones de distribución multivariadas a partir de la cópula, es la posibilidad de identificar los patrones de dependencia existentes entre los valores extremos

⁵Sin embargo, una correlación de rango de 0 no necesariamente implica independencia.

de las variables aleatorias. Un ejemplo de la importancia de este tema se encuentra en las relaciones de dependencia de retornos extremos de diferentes factores de riesgo, lo cual tiene implicaciones, por ejemplo, en medición de riesgo financiero y administración de portafolio.

Si las funciones de distribución marginales de las variables aleatorias de interés son continuas, la dependencia en las colas puede ser expresada como una medida basada en la cópula asociada a la función de distribución conjunta. Por consiguiente, esta medida de dependencia en las colas también es invariante bajo transformaciones monótonas crecientes.

Si R_1 y R_2 son dos variables aleatorias con funciones de distribución F_1 y F_2 , respectivamente, el coeficiente de dependencia en la cola superior entre R_1 y R_2 está definido como:

$$\lambda_u = \lim_{u \rightarrow 1^-} P(R_2 > F_2^{-1}(u) \mid R_1 > F_1^{-1}(u)) \quad (3.19)$$

si este límite existe, entonces $\lambda_u \in [0, 1]$. Se dice que R_1 y R_2 son asintóticamente dependientes en la cola superior si $0 < \lambda_u \leq 1$ y son asintóticamente independientes en el caso que $\lambda_u = 0$.

Análogamente, el coeficiente de dependencia en la cola inferior entre R_1 y R_2 es:

$$\lambda_l = \lim_{u \rightarrow 0^+} P(R_2 \leq F_2^{-1}(u) \mid R_1 \leq F_1^{-1}(u)) \quad (3.20)$$

En el caso de distribuciones continuas, esta medida de dependencia puede ser expresada en términos de la Cópula asociada a la distribución bivariada de estas variables:

$$\lambda_l = \lim_{u \rightarrow 0^+} \frac{P(R_2 \leq F_2^{-1}(u), R_1 \leq F_1^{-1}(u))}{P(R_1 \leq F_1^{-1}(u))} = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u}; \text{ y} \quad (3.21)$$

$$\lambda_u = \lim_{u \rightarrow 1^-} \frac{\bar{C}(u, u)}{1 - u} \quad (3.22)$$

donde $\bar{C}(u_1, u_2) = P(U_1 > u_1, U_2 > u_2) = 1 - u_1 - u_2 + C(u_1, u_2)$.

En la Tabla 3.3 se calculan varias medidas de dependencia sobre las simulaciones presentadas en el Gráfico 3.1, en estos resultados se aprecia que los coeficientes $\hat{\lambda}_l$ y $\hat{\lambda}_u$ capturan el comportamiento estocástico de las colas de las distribuciones.

3.4. Conceptos de dependencia. Como se mencionó, la cópula es una función que permite extraer la estructura de dependencia existente en la función distribución multivariada de d variables aleatorias, las cuales se pueden ver reflejadas en las medidas de dependencia basadas en la cópula ($\rho_S, \rho_\tau, \lambda_L, \lambda_U$). Sin embargo, existen diferentes tipos de dependencia, que van desde el caso de independencia hasta dependencia perfecta entre variables aleatorias.

Tal como el coeficiente de correlación de Pearson indica el grado de dependencia lineal, los coeficientes de correlación de rango muestran el grado de dependencia (lineal o no

Simulación	$\hat{\rho}$ (Pearson)	$\hat{\rho}_\tau$ (Kendall)	$\hat{\rho}_S$ (Spearman)	$\hat{\lambda}_l$	$\hat{\lambda}_u$
Panel 11	0,80	0,59	0,78	0	0
Panel 12	0,80	0,61	0,80	0	0,68
Panel 21	0,80	0,59	0,77	0,48	0,48
Panel 22	0,80	0,63	0,81	0,82	0

TABLA 3.3. Cálculo de medidas de dependencia de las simulaciones presentadas en el Gráfico 3.1.

lineal) entre variables aleatorias. Un ejemplo de la afirmación anterior se encuentra en el panel izquierdo del Gráfico 3.6, donde se muestran 5,000 replicaciones de una variable aleatoria lognormal y su logaritmo natural. Como es sabido, el logaritmo natural de una variable aleatoria lognormal distribuye normal, para este caso normal estándar.

En este ejemplo, la relación lineal entre las variables no refleja completamente su visible grado de asociación: el coeficiente de correlación de Pearson estimado para estas dos variables es igual a 0.7941. Sin embargo, al calcular el coeficiente de correlación de Spearman, el cual mide la correlación existente entre las funciones de distribución de estas dos variables aleatorias, este sí muestra el alto grado de asociación tomando un valor de 1. En el panel derecho del Gráfico 3.6 se encuentra el gráfico de dispersión de las dos funciones de distribución. Aquí, la probabilidad de que ambas variables tomen valores bajos (altos) al mismo tiempo es igual a uno.

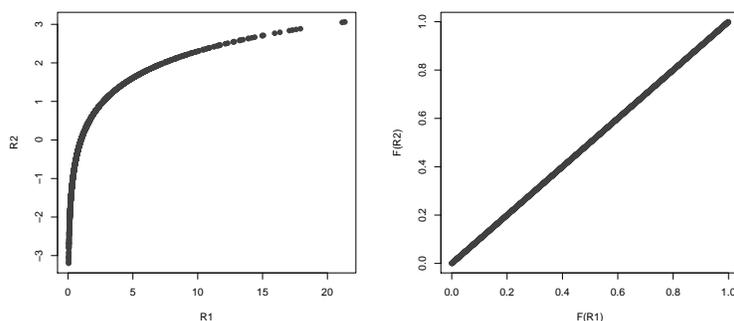


GRÁFICO 3.6. Dependencia perfecta positiva

El Gráfico 3.6 es un ejemplo de dependencia perfecta positiva entre dos variables aleatorias. Este concepto, junto con los de dependencia perfecta negativa y ordenamiento de dependencia se describen a continuación.

3.4.1. Comonotonicidad. En la sección anterior se mostró que existía dependencia perfecta positiva entre dos variables aleatorias cuando una de ellas era una función creciente

de la otra. La dependencia perfecta positiva entre variables aleatorias es conocida como comonotonicidad. La cópula que representa este tipo de dependencia es la cópula asociada al límite superior de Fréchet, el cual se define como:

$$C^u(u_1, \dots, u_d) = \min\{u_1, \dots, u_d\} \tag{3.23}$$

(propiedad *iv*) en la Sección 3.2). Si R_1, \dots, R_d son variables aleatorias comonotónicas, a partir de las definiciones (3.4) y (3.23) se tiene que:

$$\begin{aligned} F(u_1, \dots, u_d) &= \min\{F_1(r_1), \dots, F_d(r_d)\} \\ &= P(U \leq \min\{F_1(r_1), \dots, F_d(r_d)\}) \\ &= P(U \leq F_1(r_1), \dots, U \leq F_d(r_d)) \\ &= P(F_1^{-1}(U) \leq r_1, \dots, F_d^{-1}(U) \leq r_d) \end{aligned} \tag{3.24}$$

de donde se concluye que R_1, \dots, R_d se encuentran en función de la misma variable aleatoria U . Adicionalmente, tal como lo destacan McNeil et al. [2005] en el caso de variables aleatorias continuas, la comonotonicidad se presenta en el caso en que cada una de las variables aleatorias es una función creciente de una sola variable aleatoria Z . Formalmente:

$$(R_1, \dots, R_d) \stackrel{d}{=} (v_1(Z), \dots, v_d(Z)) \tag{3.25}$$

donde v_1, \dots, v_d son funciones monótonas crecientes. El Gráfico 3.7 muestra la cópula de comonotonicidad y las curvas de nivel del ejemplo utilizado para el Gráfico 3.6. En este gráfico es posible ver la forma de "L" que tiene este tipo de cópula.

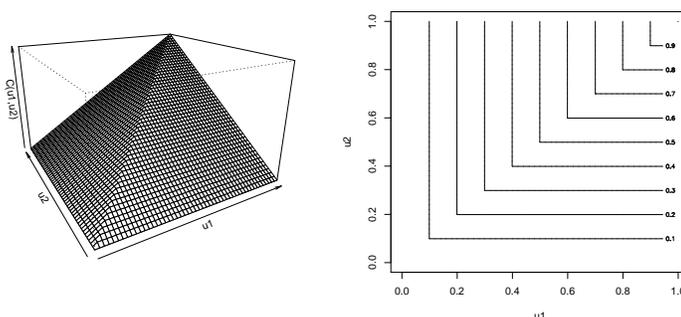


GRÁFICO 3.7. Cópula de comonotonicidad y curvas de nivel

3.4.2. *Contramotonicidad.* Análogo a la forma en que la comonotonicidad describe la dependencia perfecta positiva, la contramotonicidad se refiere al caso en el que existe dependencia perfecta negativa entre variables aleatorias y su cópula esta asociada con el límite inferior de Fréchet:

$$C^l(u_1, \dots, u_d) = \max\{u_1 + \dots + u_d - 1, 0\} \tag{3.26}$$

Sin embargo, tal como lo muestran McNeil et al. [2005] esta cópula es únicamente válida para $d = 2$. Otra forma de entender este concepto, es similar a la definición (3.25), donde en el caso de $d = 2$:

$$(R_1, R_2) \stackrel{d}{=} (v_1(Z), v_2(Z)) \quad (3.27)$$

donde v_1 es una función monótona creciente (decreciente) mientras v_2 es una función monótona decreciente (creciente). Embrechts et al. [1999], muestran que la cópula de contramonotonía corresponde al caso en que dos variables aleatorias R_1 y R_2 dependen del vector $(U, 1 - U)$ donde $U \sim U(0, 1)$.

Un ejemplo de la contramonotonía se muestra en el Gráfico 3.8, donde se presentan 5,000 replicaciones de una variable aleatoria $F_{4,8}$ y una transformación monótona decreciente como lo es $1/F_{4,8}$, la cual sigue una distribución $F_{8,4}$.

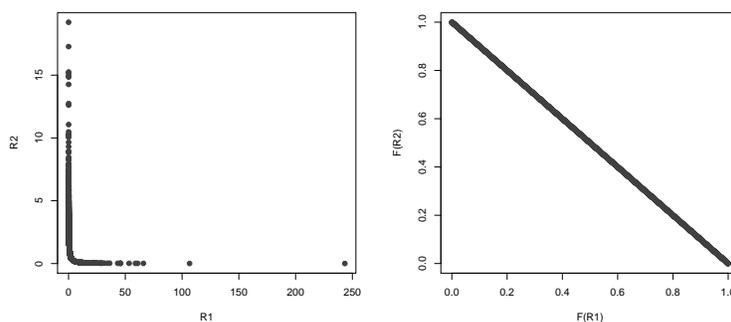


GRÁFICO 3.8. Dependencia perfecta negativa

Nuevamente, la relación no lineal entre las variables (panel izquierdo) y la relación entre sus funciones de distribución (panel derecho) son evidentes. El coeficiente de correlación de Pearson estimado para este caso es $-0,2178$, mientras que el coeficiente de correlación de rango y el τ de Kendall son iguales a -1 .

El Gráfico 3.9 muestra la cópula y las curvas de nivel obtenidas para el ejemplo anterior, en ésta la forma de la cópula es un plano con valores positivos después de la recta $u_1 + u_2 = 1$.

3.4.3. Concordancia. Adicional a los conceptos de comonotonía y contramonotonía se encuentran los conceptos de concordancia y discordancia, los cuales ofrecen información acerca del signo de la dependencia entre dos variables aleatorias [Embrechts et al., 1999]. A partir de este concepto, es posible establecer un “ordenamiento” de las funciones de distribución [Joe, 1997]⁶.

⁶El orden de concordancia es sólo uno de los tipos de ordenamiento que se pueden establecer entre funciones de distribución, otros tipos de ordenamientos se encuentran en Joe [1997].

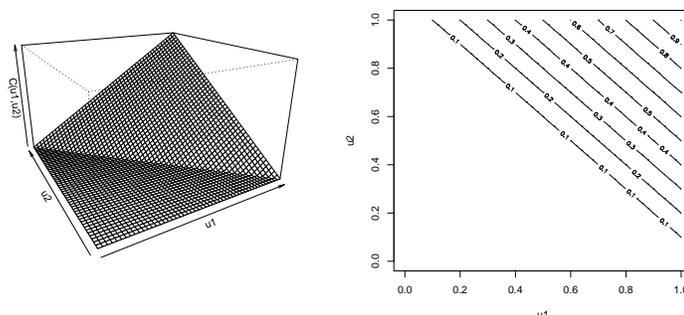


GRÁFICO 3.9. Cópula de contramonotonicidad y curvas de nivel

El caso más simple es aquel en el que se compara las funciones de distribución multivariadas con respecto a la función de distribución de independencia de d variables aleatorias. Si R_1 y R_2 son dos variables aleatorias con función de distribución F , la cual satisface que:

$$F(r_1, r_2) \geq F(r_1) F(r_2) \tag{3.28}$$

para todo $r_1, r_2 \in \mathbb{R}$ se dice que F (o análogamente el vector aleatorio (R_1, R_2)) es dependiente en el cuadrante positivo (PQD, por sus siglas en inglés). De una manera intuitiva, esto implica que las variables R_1 y R_2 tienen una probabilidad más alta de tomar valores grandes o pequeños que en el caso en que son independientes. Para la definición de dependencia en el cuadrante negativo (NQD), basta con cambiar el sentido de la desigualdad en (3.28) [Nelsen, 2006]⁷.

De acuerdo con lo expuesto en las secciones anteriores, es posible establecer una conexión entre la dependencia en los cuadrantes y la cópula. Si (3.28) se mantiene para dos variables aleatorias R_1 y R_2 que tienen una cópula C , se tiene que:

$$C(u_1, u_2) \geq u_1 u_2 \tag{3.29}$$

[Nelsen, 2006]. Es decir, la cópula C es PQD si para todo $u_1, u_2 \in [0, 1]$ se mantiene la condición (3.29). Si se cambia el sentido de la desigualdad, se tiene la definición de NQD.

Basados en estos conceptos, es posible definir un ordenamiento de concordancia entre las funciones de distribución dependiendo de su grado de PQD o NQD. Así, para dos funciones de distribución bivariadas F y F' (análogamente para las parejas de variables aleatorias $(R_1, R_2)'$ y $(R'_1, R'_2)'$ o sus cópulas asociadas C y C') se dice que F es más dependiente en el cuadrante positivo o más concordante que F' , lo cual se escribe como $F' \prec_c F$, si se cumple que:

$$F(r_1, r_2) \geq F'(r_1, r_2) \tag{3.30}$$

⁷La extensión de esta definición para $d \geq 3$ se encuentra en Joe [1997].

para todo $r_1, r_2 \in \mathbb{R}$ [Joe, 1997]. Trasladando este resultado a las cópulas es posible establecer el siguiente ordenamiento:

$$C^l(u_1, u_2) \prec_c C^0(u_1, u_2) \prec_c C^I(u_1, u_2) \prec_c C^1(u_1, u_2) \prec_c C^u(u_1, u_2) \quad (3.31)$$

Donde C^l , C^I y C^u representan las cópulas de contramonotonicidad, independencia y comonotonicidad, respectivamente, C^0 es una cópula NQD y C^1 es una cópula PQD. Nuevamente, las cópulas asociadas a los límites de Fréchet juegan un papel importante en estos conceptos, ya que son los límites de este orden de concordancia.

Gráficamente, el orden de concordancia puede ser entendido como un posicionamiento más hacia la izquierda de las curvas de nivel de la cópula que sea más PQD. El Gráfico 3.10 muestra las curvas de nivel de las cópulas de contramonotonicidad, independencia y comonotonicidad, respectivamente. Es posible ver que las curvas de nivel de la cópula de contramonotonicidad siempre se encuentran a la derecha de la cópula de independencia, mientras las curvas de la cópula de comonotonicidad se encuentran a la izquierda.

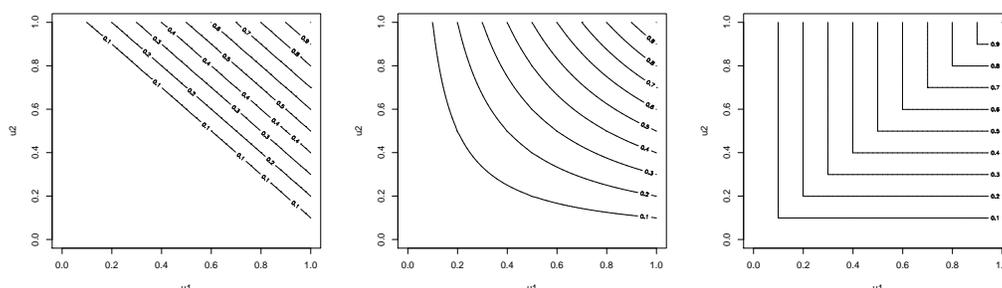


GRÁFICO 3.10. Curvas de nivel de las cópulas de contramonotonicidad, independencia y comonotonicidad, respectivamente.

3.5. Familias de cópulas. En secciones anteriores se definió qué es una cópula y se mostraron algunas ecuaciones, gráficos y resultados asociados a cópulas específicas, por ejemplo, las de independencia, comonotonicidad y contramonotonicidad. No obstante, existe una gran cantidad de funciones que satisfacen las condiciones necesarias para ser consideradas como cópulas. Una lista de cópulas bivariadas y multivariadas comúnmente mencionadas en la literatura se encuentra en el Anexo C.

De acuerdo con sus características, las cópulas pueden ser agrupadas en diferentes categorías⁸. Por ejemplo, las cópulas presentadas en la sección anterior son conocidas en la literatura como las cópulas límite o cópulas “fundamentales”, ya que se encuentran asociadas con los casos extremos de dependencia: comonotonicidad, contramonotonicidad e independencia.

⁸No obstante, como se verá más adelante, en muchos casos los criterios de agrupación no son excluyentes.

Un segundo criterio de agrupación comúnmente usado en la literatura se encuentra asociado a la forma funcional de la cópula. Este criterio identifica dos tipos de cópulas, las cópulas “explícitas”, aquellas que pueden ser expresadas a través de una forma funcional cerrada, o “implícitas” las cuales son derivadas de funciones de distribución multivariadas conocidas, aunque sus formas funcionales no son simples o cerradas. Algunos ejemplos de cópulas explícitas se encuentran en las cópulas fundamentales (ecuaciones (3.10), (3.23) y (3.26)), mientras que entre las cópulas implícitas se encuentran las cópulas derivadas de las funciones de distribución normal y t multivariada, las cuales se discuten en la Sección 3.5.1.

Por último, el tercer criterio de agrupación depende directamente de las características particulares de las cópulas. Bajo este criterio, es posible identificar cuatro grandes grupos o “familias” de cópulas, las cuales se presentan a continuación.

3.5.1. Cópulas elípticas. De gran popularidad en los mercados financieros, las cópulas elípticas son aquellas que se encuentran asociadas a variables aleatorias cuya función de distribución multivariada es de la forma $f(\mathbf{x}'\mathbf{x})$, por lo que las curvas de nivel de variables aleatorias que tengan este tipo de cópulas forman elipses [Bouyé et al., 2000]⁹. Las dos cópulas más importantes en esta familia son la cópula normal (o Gaussiana) y la cópula t de student, las cuales se derivan de las funciones de distribución multivariada que poseen estos mismos nombres. La cópula Gaussiana con matriz de correlación \mathbf{P} viene de la forma:

$$\begin{aligned} C_{\mathbf{P}}^{Ga}(\mathbf{u}) &= \Phi_{\mathbf{P}}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)) \\ &= \int_{-\infty}^{\Phi^{-1}(u_1)} \dots \int_{-\infty}^{\Phi^{-1}(u_d)} \frac{1}{(2\pi)^{\frac{d}{2}} |\mathbf{P}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \mathbf{x}' \mathbf{P}^{-1} \mathbf{x}\right) dx_1 \dots dx_d \end{aligned} \quad (3.32)$$

donde $\Phi_{\mathbf{P}}$ es la función de distribución normal multivariada con matriz de correlaciones $\mathbf{P} \in \mathbb{R}^{d \times d}$ y Φ^{-1} es la función de distribución inversa de la distribución normal estándar [Cherubini et al., 2004].

Por su parte, la cópula t de student con matriz de correlación $\mathbf{P} \in \mathbb{R}^{d \times d}$ y $v > 0$ grados de libertad viene de la forma:

$$\begin{aligned} C_{v, \mathbf{P}}^t(\mathbf{u}) &= t_{v, \mathbf{P}}(t_v^{-1}(u_1), \dots, t_v^{-1}(u_d)) \\ &= \int_{-\infty}^{t_v^{-1}(u_1)} \dots \int_{-\infty}^{t_v^{-1}(u_d)} \frac{\Gamma(\frac{v+d}{2}) |\mathbf{P}|^{-\frac{1}{2}}}{\Gamma(\frac{v}{2}) (v\pi)^{\frac{d}{2}}} \left(1 + \frac{1}{v} \mathbf{x}' \mathbf{P}^{-1} \mathbf{x}\right)^{-\frac{v+d}{2}} dx_1 \dots dx_d \end{aligned} \quad (3.33)$$

donde $t_{v, \mathbf{P}}$ es la función de distribución t de student multivariada con matriz de correlaciones \mathbf{P} y v grados de libertad y t_v^{-1} es la función de distribución inversa asociada a la distribución t univariada con v grados de libertad [Cherubini et al., 2004].

Debido a que este tipo de cópulas está caracterizado principalmente por la matriz de correlaciones \mathbf{P} , más el número de grados de libertad en el caso de la cópula t , las cópulas

⁹Para una discusión detallada sobre distribuciones elípticas véase el Anexo A.

elípticas resultan atractivas en el mundo financiero, ya que en portafolios conformados por factores de riesgo que sigan una distribución elíptica se satisfacen las condiciones que garantizan la subaditividad del Valor en Riesgo (VaR_α) y el principio de diversificación. En consecuencia la optimización del portafolio de mínima varianza resulta apropiado.

Implícitamente, en la mayoría de aplicaciones financieras es común encontrar que se supongan relaciones de dependencia descritas por cópulas elípticas. Por ejemplo, cuando se supone una estructura del tipo GARCH multivariado y se usa como función de verosimilitud la función de distribución normal multivariada, implícitamente se está suponiendo que el proceso de innovaciones (estandarizado) de los retornos de los activos sigue esta cópula, con distribuciones marginales normales [Malevergne y Sornette, 2006]. Un razonamiento similar puede ser aplicado a la metodología de *RiskMetrics*® en la cual se supone normalidad multivariada de los retornos de los activos.

No obstante, una ventaja que poseen las cópulas Gaussiana y t de student con respecto a las funciones de distribución de las cuales se derivan, es que a partir de las cópulas es posible utilizar variables aleatorias que sigan distribuciones marginales que no sean Gaussianas o t de student. Nótese que si bien las ecuaciones (3.32) y (3.33) se encuentran definidas en términos de las funciones de distribución inversas Φ^{-1} y t_v^{-1} , éstas se evalúan en las variables $u_i = F_i(r_i)$ con $i = 1, \dots, d$, donde F_i corresponde a una función de distribución arbitraria, por ejemplo, una función de distribución Weibull o Chi cuadrado. De aquí, las variables aleatorias cuya estructura de dependencia este descrita por una cópula Gaussiana o una t de student pueden describir comportamientos marginales diferentes. Las distribuciones derivadas de esta manera son conocidas como “meta-distribuciones” y pueden ser pensadas como una generalización de las distribuciones multivariadas tradicionales. Por ejemplo, una distribución meta-Gaussiana sería aquella cuyo comportamiento multivariado se describa por una cópula Gaussiana, pero al menos una de las distribuciones marginales sea diferente a la distribución normal.

Las densidades de dos distribuciones elípticas bivariadas obtenidas a partir de cópulas elípticas se presentan en el Gráfico 3.11. El panel superior corresponde a dos variables aleatorias con una cópula Gaussiana y funciones de distribución marginales normal estándar y Fréchet con $\theta = 2$, respectivamente. En el panel inferior se muestra la densidad de una distribución meta- t con $v = 4$ grados de libertad y marginales normales estándar; en ambos casos $\rho_{12} = 0,7$.

Las cópulas elípticas ofrecen una estructura flexible para describir comportamientos multivariados diversos. Como lo muestran McNeil et al. [2005], para el caso de la cópula t bivariada, el coeficiente de dependencia en las colas está dado por:

$$\lambda_u = \lambda_l = 2t_{v+1} \left(-\sqrt{\frac{(v+1)(1-\rho)}{1+\rho}} \right) \quad (3.34)$$

el cual es mayor que cero, lo que indica que existe una dependencia significativa entre los valores extremos de las series descritas por esta cópula. Por su parte, el coeficiente de dependencia en las colas para la cópula normal es igual a cero. En la Tabla 3.4 se

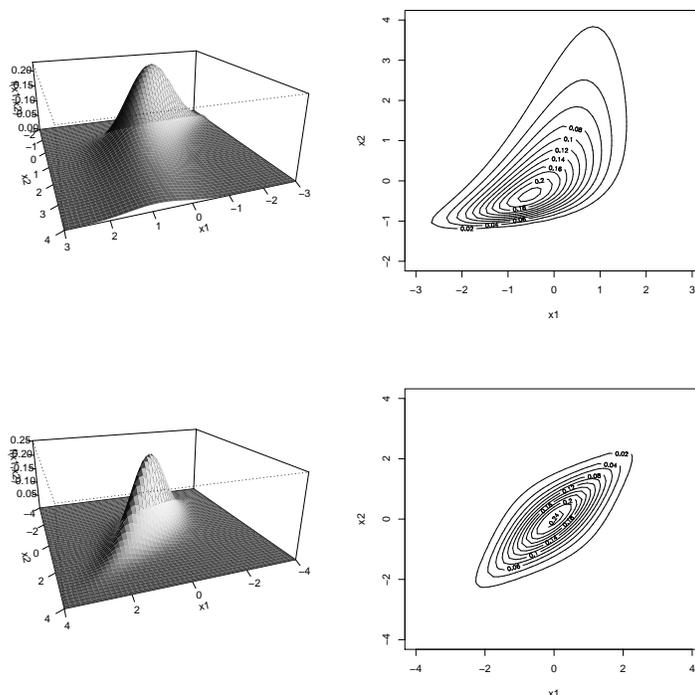


GRÁFICO 3.11. Superficies y funciones de densidad asociadas a dos distribuciones elípticas bivariadas.

presentan el coeficiente de dependencia en las colas y las correlaciones de rango, ρ_S de Spearman y ρ_τ de Kendall, para estas dos cópulas. Cabe destacar que a diferencia de la cópula Gaussiana, no se conoce una fórmula simple para el ρ_S de Spearman de la cópula t . En consecuencia, este valor debe ser obtenido por integración numérica a partir de las ecuaciones (3.17) y (3.33) [Liu, 2006].

	Cópula Gaussiana	Cópula t de student
$\hat{\rho}_\tau$ (Kendall)	$\frac{2}{\pi} \arcsen(\rho_{ij})$	$\frac{2}{\pi} \arcsen(\rho_{ij})$
$\hat{\rho}_S$ (Spearman)	$\frac{6}{\pi} \arcsen(\frac{1}{2}\rho_{ij})$	—
$\lambda_l = \lambda_u$	0,0	$2t_{v+1} \left(-\sqrt{\frac{(v+1)(1-\rho)}{1+\rho}} \right)$

TABLA 3.4. Medidas de dependencia para las cópulas normal y t_v

3.5.2. *Cóputas de Arquímedes.* Las cóputas de Arquímedes conforman una de las familias más representativas y estudiadas en el ámbito teórico y aplicado. En general, su popularidad se encuentra asociada a su sencilla definición, la cual ha permitido construir un numeroso grupo de funciones que pertenecen a esta familia [Nelsen, 2006].

Sea $H(r_1, r_2)$ la función de distribución bivariada de las variables aleatorias R_1 y R_2 , con distribuciones marginales F_1 y F_2 , respectivamente. Si bien sólo en el caso de independencia entre R_1 y R_2 se observa que $H(r_1, r_2) = F_1(r_1)F_2(r_2)$, es posible encontrar funciones para las cuales se satisface la siguiente condición:

$$\lambda(H(r_1, r_2)) = \lambda(F_1(r_1)) \lambda(F_2(r_2)) \quad (3.35)$$

Para alguna función λ que debe ser positiva en el intervalo $(0, 1)$. Si se define la función $\phi(t)$ como $\phi(t) \equiv -\log \lambda(t)$, para la cóputa $C(u_1, u_2)$, la ecuación (3.35) queda de la forma:

$$\phi(C(u_1, u_2)) = \phi(u_1) + \phi(u_2) \quad (3.36)$$

despejando $C(u_1, u_2)$ de la ecuación anterior, se obtiene $C(u_1, u_2) = \phi^{[-1]}(\phi(u_1) + \phi(u_2))$ donde $\phi^{[-1]}(s)$ es la pseudo-inversa de $\phi(t)$, la cual se define como:

$$\phi^{[-1]}(s) = \begin{cases} \phi^{-1}(s), & \text{si } 0 \leq s \leq \phi(0) \\ 0, & \text{si } \phi(0) \leq s \leq +\infty \end{cases} \quad (3.37)$$

Las consideraciones anteriores motivan la siguiente definición para las cóputas de Arquímedes. Sea $\phi: [0, 1] \rightarrow \mathbb{R}^+$ una función continua, estrictamente decreciente, cóncava y que satisface $\phi(1) = 0$. Se dice que la cóputa C es una cóputa de Arquímedes si se puede representar de la siguiente forma:

$$C(u_1, u_2) = \phi^{[-1]}(\phi(u_1) + \phi(u_2)) \quad (3.38)$$

donde ϕ es conocido como el “generador” de la cóputa C y $\phi^{[-1]}$ es la pseudo inversa de ϕ , definida en (3.37). En el caso en que $\phi(0) \rightarrow +\infty$, entonces se dice que $\phi(t)$ es un generador estricto y $\phi^{[-1]}(s) \equiv \phi^{-1}(s)$.

Algunos ejemplos de cóputas bivariadas que satisfacen la condición (3.38) se encuentran en las cóputas de contramonotonidad e independencia. En el primer caso, su generador $\phi_l(t) = 1 - t$ es un generador no estricto y, en consecuencia, al utilizar la definición (3.37) se tiene que $\phi_l^{[-1]}(s) = 1 - s$ para $s \in [0, 1]$ y 0 para $s > 1$, es decir $\phi_l^{[-1]}(s) = \max\{1 - s, 0\}$. Así, a partir de (3.38) se obtiene:

$$\begin{aligned} C(u_1, u_2) &= \max\{1 - (1 - u_1 + 1 - u_2), 0\} \\ &= \max\{u_1 + u_2 - 1, 0\} \\ &= C^l(u_1, u_2) \end{aligned}$$

A diferencia del caso de contramonotonidad, para la cópula de independencia $\phi_I(t) = -\log(t)$ es un generador estricto por lo que $\phi_I^{[-1]}(s) = \phi_I^{-1}(s) = \exp(-s)$. Utilizando (3.38), se tiene que:

$$\begin{aligned} C(u_1, u_2) &= \exp(-(-\log(u_1) - \log(u_2))) \\ &= \exp(\log(u_1) + \log(u_2)) \\ &= C^I(u_1, u_2) \end{aligned}$$

A diferencia de las cópulas elípticas, las cuales pueden ser definidas para dimensiones arbitrariamente grandes, existen cópulas de Arquímedes bivariadas que no pueden ser generalizadas para dimensiones mayores que dos. Un ejemplo importante se encuentra en la cópula de contramonotonidad, que si bien representa uno de los límites de Fréchet para cópulas de cualquier dimensión d , no puede ser considerada una cópula para $d \geq 3$ [McNeil et al., 2005]. Para construir cópulas de Arquímedes de dimensiones mayores a 2, es necesario incorporar una restricción adicional sobre las derivadas de $\phi^{[-1]}(t)$. Específicamente, $\phi^{[-1]}(t)$ debe satisfacer:

$$(-1)^k \frac{d^k}{dt^k} \phi^{[-1]}(t) \geq 0 \quad k = 0, 1, 2, \dots \quad (3.39)$$

sobre el intervalo $[0, \infty)$. Para que esto suceda, $\phi(t)$ debe ser un generador estricto y en consecuencia $\phi^{[-1]}(t) = \phi^{-1}(t)$ [Nelsen, 2006]. En caso que la condición (3.39) se mantenga, entonces la función:

$$C(u_1, \dots, u_d) = \phi^{-1}(\phi(u_1) + \dots + \phi(u_d)) \quad (3.40)$$

es una cópula de Arquímedes para dimensiones mayores o iguales a 2 [Kimberling, 1974]. Teniendo en cuenta este resultado, la cópula de independencia es una cópula de Arquímedes para $d \geq 3$, mientras que la cópula de contramonotonidad no lo es.

Por último, Marshall y Olkin [1988] muestran que una forma sencilla de construir generadores estrictos para cópulas de Arquímedes, es a través de la transformación de Laplace de la función de distribución asociada a una variable aleatoria $\gamma > 0$, la cual se define como:

$$\psi(t) = E(e^{-t\gamma}) = \int_0^\infty e^{-tx} dG(x) \quad (3.41)$$

donde $G(x)$ corresponde a la función de distribución de γ , la cual debe satisfacer $G(0) = 0$. A partir de (3.41), es posible mostrar que $\phi(t) \equiv \psi^{-1}(t)$ [Cherubini et al., 2004].

Ejemplos de cópulas de Arquímedes. Como ya se mencionó, las cópulas de Arquímedes son la familia de funciones más numerosa. La mayoría de cópulas que pertenecen a esta familia son funciones de uno o dos parámetros, lo que si bien permite representar fácilmente diferentes tipos de dependencia, también implica una de sus mayores limitaciones, ya que resulta complicado describir relaciones de dependencia complejas con un número reducido de parámetros, especialmente en dimensiones altas [Nelsen, 2006]. A continuación, se mencionan cuatro de las cópulas de Arquímedes más relevantes. Una discusión detallada de sus propiedades se encuentra en Joe [1997].

- *Cópula Clayton*. También conocida como la cópula Cook-Johnson o Kimeldorf-Sampson, el generador de esta familia es $\phi(t) = \frac{1}{\theta} (t^{-\theta} - 1)$ con $\theta > 0$. La cópula de Clayton viene de la forma:

$$C^{Cl}(u_1, \dots, u_d) = \left(u_1^{-\theta} + \dots + u_d^{-\theta} - d + 1 \right)^{\frac{1}{\theta}} \quad (3.42)$$

- *Cópula Gumbel*. También conocida como cópula de Gumbel-Hougaard, su generador está definido por $\phi(t) = (-\log(t))^\theta$ con $\theta > 1$. En este caso, la cópula Gumbel toma la forma:

$$C^{Gu}(u_1, \dots, u_d) = \exp \left(- \left(\tilde{u}_1^\theta + \dots + \tilde{u}_d^\theta \right)^{\frac{1}{\theta}} \right) \quad (3.43)$$

donde $\tilde{u}_i \equiv -\log u_i$.

- *Cópula Frank*. El generador de la cópula Frank está definido por $\phi(t) = -\log \left(\frac{\exp(-\theta t) - 1}{\exp(-\theta) - 1} \right)$ con $\theta > 0$. La cópula asociada a este generador es:

$$C^{Fr}(u_1, \dots, u_d) = -\frac{1}{\theta} \log \left[1 + \frac{\prod_{i=1}^d (e^{-\theta u_i} - 1)}{(e^{-\theta} - 1)^{d-1}} \right] \quad (3.44)$$

- *Cópula Clayton generalizada*. Esta es un ejemplo de cópula de Arquímedes de dos parámetros. El generador de esta cópula es $\phi(t) = \theta^{-\delta} (t^{-\theta} - 1)^\delta$, donde $\theta > 0$ y $\delta \geq 1$. Para el caso bivariado, la cópula Clayton generalizada es:

$$C^{GC}(u_1, u_2) = \left\{ \left[\left(u_1^{-\theta} - 1 \right)^\delta + \left(u_2^{-\theta} - 1 \right)^\delta \right]^{1/\delta} + 1 \right\}^{-\frac{1}{\theta}} \quad (3.45)$$

Algunos resultados asociados a cada una de las cópulas de Arquímedes reportadas en los párrafos anteriores, como lo son el τ de Kendall y los coeficientes de dependencia en las colas se encuentran definidos en la Tabla 3.5, la cual es tomada de McNeil et al. [2005].

Como se puede observar, si bien todas estas cópulas son de Arquímedes, describen tipos de dependencia completamente diferentes. Por ejemplo, la cópula Gumbel muestra dependencia en los extremos únicamente en la cola superior, mientras que la cópula Clayton la exhibe en la cola contraria. Por su parte, la cópula Frank no muestra dependencia en ninguna de las dos colas y la cópula Clayton generalizada muestra dependencia en los extremos para ambas colas, aunque pueden ser de intensidades diferentes. De otra parte, en todos los ejemplos, el τ de Kendall se expresa como función de los parámetros, propiedad que resultará útil para la estimación de los parámetros de este tipo de cópulas a través del método de momentos (Sección 4.1.2).

3.5.3. Cópulas de valor extremo. Otra familia de cópulas de uso común es la de valor extremo. Estas cópulas son derivadas de la estructura de dependencia de la distribución generalizada de valor extremo multivariada, *MGEV*. Cabe destacar que debido a la importancia del análisis de los valores extremos en la medición de riesgos, los principales

Cópula	τ de Kendall	λ_u	λ_l
Clayton	$\frac{\theta}{\theta+2}$	0	$2^{-1/\theta}$
Frank	$1 - \frac{4}{\theta} \left[1 - \frac{1}{\theta} \int_0^\theta \frac{t}{\exp(t)-1} dt \right]$	0	0
Gumbel	$1 - \frac{1}{\theta}$	$2 - 2^{1/\theta}$	0
Clayton Generalizada	$\frac{(2+\theta)\delta-2}{(2+\theta)\delta}$	$2 - 2^{1/\delta}$	$2^{-1/(\theta\delta)}$

TABLA 3.5. Coeficiente τ de Kendall y coeficientes de dependencia en las colas para algunas cópulas de Arquímedes

resultados asociados a las cópulas pertenecientes a esta familia se discutirán en detalle en la Sección 4.5.

En general, una cópula de valor extremo satisface la siguiente relación:

$$C(u_1^t, \dots, u_d^t) = [C(u_1, \dots, u_d)]^t \quad (3.46)$$

para $t > 0$.

Otra forma de representar esta familia de cópulas para el caso bivariado¹⁰ es mediante la expresión:

$$C(u_1, u_2) = \exp \left[\log(u_1 u_2) A \left(\frac{\log(u_1)}{\log(u_1 u_2)} \right) \right] \quad (3.47)$$

donde $A(\cdot) : [0, 1] \rightarrow [\frac{1}{2}, 1]$, conocida como la función de dependencia, es una función convexa tal que $\max(\omega, 1 - \omega) \leq A(\omega) \leq 1$ para todo $\omega \in [0, 1]$.

A partir de la condición (3.46) es fácil verificar que las cópulas Gumbel, de independencia, y de comonotonidad pertenecen a esta clase. Estas cópulas y otras que pertenecen a la familia de valor extremo, así como sus funciones de dependencia, son presentadas en la Tabla 4.2.

3.5.4. *Cópulas Arquimax.* Capéraà et al. [2000] combinaron las cópulas de valor extremo y las de Arquímedes en una clase denominada cópulas Arquimax¹¹. Estas cópulas tienen la siguiente forma:

$$C(u_1, u_2) = \phi^{-1} \left[(\phi(u_1) + \phi(u_2)) A \left(\frac{\phi(u_1)}{\phi(u_1) + \phi(u_2)} \right) \right] \quad (3.48)$$

donde $A(t)$ es una función de dependencia similar a la definida para las cópulas de valor extremo y $\phi(t)$ es el generador de las cópulas de Arquímedes. Cabe destacar, que de

¹⁰Esta representación es presentada en la Sección 4.5 para el caso multivariado.

¹¹En inglés, cópulas *Archimax*.

acuerdo con la definición (3.48), tanto las cópulas de Arquímedes como las de valor extremo pueden ser consideradas como casos particulares de la familia Arquimax. En efecto, si $A(t) = 1$, la expresión (3.48) toma la forma de una cópula de Arquímedes, mientras que las cópulas de valor extremo se obtienen cuando $\phi(t) = -\log(t)$.

No obstante, la afirmación anterior no implica que toda cópula Arquimax deba pertenecer a las familias de valor extremo o de Arquímedes. Una cópula de Arquimax que no pertenece a ninguna de estas dos familias es la cópula BB4 propuesta por Joe [1997]. Esta se encuentra descrita en la Tabla C.3, donde:

$$A(t) = 1 - \left[t^{-\theta} + (1-t)^{-\theta} \right]^{-1/\theta} \quad (3.49)$$

$$\phi(t) = t^{-\delta} - 1 \quad (3.50)$$

Capéraà et al. [2000], entre otros resultados, encuentran que el coeficiente de correlación $\tau_{\phi,A}$ de Kendall tiene la siguiente forma para las cópulas Arquimax:

$$\tau_{\phi,A} = \tau_A + (1 - \tau_A) \tau_{\phi} \quad (3.51)$$

donde τ_A es el τ de Kendall asociado a una cópula de valor extremo con función de dependencia A (C_A) y τ_{ϕ} es el τ de Kendall de una cópula de Arquímedes con función generadora ϕ (C_{ϕ}).

Adicionalmente, estos autores demuestran que las cópulas Arquimax, $C_{\phi,A}$, que satisfacen la condición $\lim_{t \rightarrow \infty} \frac{\phi(1-\frac{1}{st})}{\phi(1-\frac{1}{t})} = s^{-m} \forall s > 0$, exhiben dependencia en la cola superior. El coeficiente asociado a esta dependencia, descrito en la Sección 3.3.2, para esta familia de cópulas es:

$$\lambda_u = 2 - \left[2A \left(\frac{1}{2} \right) \right]^{1/m} \leq 1 \quad (3.52)$$

Los valores que toman este coeficiente indican que existe dependencia en la cola superior, excepto cuando $m = 1$ y $A \equiv 1$, ya que en este caso λ_u es cero.

3.6. Simulación. Entre los principales usos de la cópula se destaca la simulación de variables aleatorias con diferentes estructuras de dependencia en la modelación de riesgos. Por ejemplo, el Gráfico 3.12 presenta 1000 observaciones obtenidas a partir de la simulación de un vector de tres variables aleatorias uniformes $(0, 1)$, cuya distribución multivariada es la cópula t con 5 grados de libertad y correlaciones $\rho_{12} = 0,5$, $\rho_{13} = 0,3$ y $\rho_{23} = 0,7$. En este caso, cada uno de los vectores fue transformado de tal forma que siguiera una distribución marginal diferente (normal estándar, logística con parámetros $(0, 1)$ y t con 7 grados de libertad), aplicando la función inversa de distribución. Con este ejemplo, se muestra la versatilidad de la cópula para generar variables aleatorias con diferentes grados de dependencia y gran variedad de comportamientos marginales. A continuación se describen los principales algoritmos para la generación de números aleatorios uniformes que sigan una cópula específica.

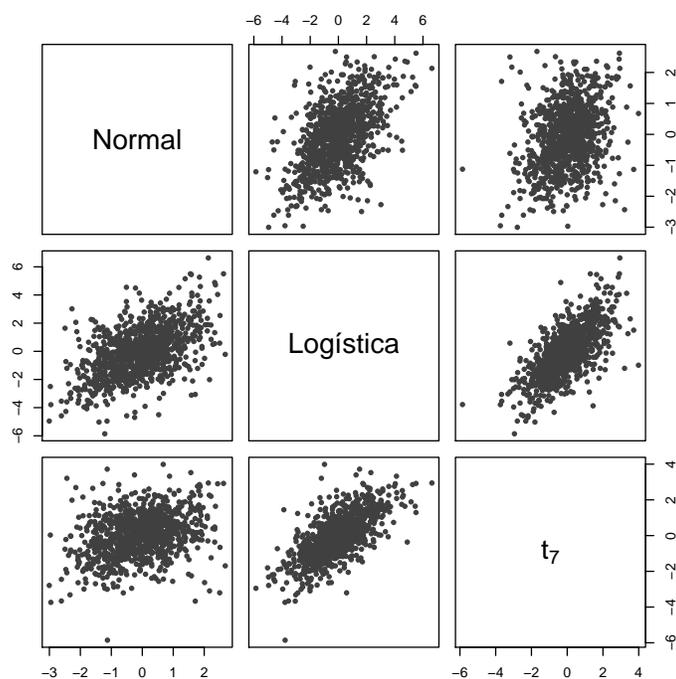


GRÁFICO 3.12. Simulación de 1000 observaciones a partir de una cópula t_5 tridimensional cuyas marginales distribuyen normal estándar, logística con parámetros $(0, 1)$ y t_7

- *Algoritmo de simulación para la cópula Gaussiana.*

La idea detrás de este algoritmo se fundamenta en dos hechos puntuales. La función de distribución normal multivariada depende exclusivamente de sus dos primeros momentos y para este caso en particular, la matriz de varianzas y covarianzas juega un papel fundamental, ya que ésta determina la relación de dependencia entre las variables. Además, a partir de la ecuación (3.32) este proceso puede ser entendido como un procedimiento en el que se obtienen d variables aleatorias uniformes a partir de la simulación de un vector aleatorio que sigue una distribución normal multivariada con media cero y matriz de varianzas y covarianzas Σ . El algoritmo de simulación para este tipo de cópula es el siguiente:

- i) Encuentre la matriz A , la cual es la descomposición de Cholesky de la matriz de varianzas y covarianzas Σ , que satisface la condición $AA' = \Sigma$.
- ii) Simule n realizaciones del vector $\{\mathbf{z}_i\}_{i=1}^n$, el cual está compuesto por d variables aleatorias normales estándar, las cuales son independientes entre ellas. En consecuencia, $\mathbf{z}_i = [z_{1i}, \dots, z_{di}] \sim N(\mathbf{0}, I_d)$.

- iii) Calcule $\mathbf{x}_i = A\mathbf{z}_i$. De aquí, $\mathbf{x}_i = [x_{1i}, \dots, x_{di}] \sim N(\mathbf{0}, \Sigma)$.
- iv) Genere el vector $\mathbf{u}_i = [u_{1i}, \dots, u_{di}]$, con $u_{ji} = \Phi(x_{ji})$, $i = 1, \dots, n$ $j = 1, \dots, d$, donde Φ corresponde a la función de distribución normal. De esta forma, \mathbf{u}_i sigue una cópula Gaussiana con matriz de varianzas y covarianzas Σ .
- *Algoritmo de simulación para la cópula t de student.* La simulación de variables aleatorias uniformes que sigan este tipo de cópula, se basan en que toda variable aleatoria t con v grados de libertad se puede ver como el cociente:

$$T = \sqrt{v} \frac{Z}{\sqrt{J}} \quad (3.53)$$

donde Z es una variable aleatoria normal estándar y J es una variable chi cuadrado con v grados de libertad. De esta manera, nuevamente la simulación dependerá de la generación de variables normales multivariadas, como se muestra a continuación.

- i) Encuentre la matriz A , la cual es la descomposición de Cholesky de la matriz de varianzas y covarianzas Σ , que satisface la condición $AA' = \Sigma$.
- ii) Simule n realizaciones del vector $\{\mathbf{z}_i\}_{i=1}^n$, el cual esta compuesto por d variables aleatorias normales estándar, las cuales son independientes entre ellas.
- iii) Calcule $\mathbf{x}_i = A\mathbf{z}_i$.
- iv) Genere una variable aleatoria s_i , la cual sigue una distribución chi cuadrado con v grados de libertad, independiente de \mathbf{z}_i .
- v) Calcule $\mathbf{y}_i = \frac{\sqrt{v}}{\sqrt{s_i}} \mathbf{x}_i$
- vi) Genere el vector $\mathbf{u}_i = [u_{1i}, \dots, u_{di}]$, con $u_{ji} = t_v(y_{ji})$, $i = 1, \dots, n$ $j = 1, \dots, d$, donde t_v corresponde a la función de distribución t con v grados de libertad. Como resultado de esta simulación \mathbf{u}_i sigue una cópula t_v con matriz de varianzas y covarianzas Σ .
- *Algoritmo para la generación de números aleatorios que sigan una cópula dada, C .* Este método se basa en la distribución condicional del vector aleatorio \mathbf{u} . En el caso bivariado se tiene que:

$$P(U_2 \leq u_2 | U_1 = u_1) = C_{2|1}(u_1, u_2) = \frac{\partial C(u_1, u_2)}{\partial u_1} \quad (3.54)$$

La expresión (3.54) corresponde a la distribución marginal de u_2 condicionada al valor de u_1 . Con este principio, el algoritmo para simular este tipo de variables es el siguiente:

- i) Se simulan n realizaciones de dos variables aleatorias uniformes $(0, 1)$ independientes, w_{i1} y w_{i2} con $i = 1, \dots, n$.
- ii) Posteriormente, se define el vector $\mathbf{u}_i = [u_{i1}, u_{i2}]$ de la siguiente forma $u_{i1} = w_{i1}$ y $u_{i2} = C_{2|1}^{-1}(u_{i1}, w_{i2})$. Por consiguiente, \mathbf{u}_i tiene como distribución conjunta a la cópula C .

Si bien este procedimiento parece simple y general, su implementación puede resultar complicada dependiendo de qué tan sencillo sea obtener las distribuciones marginales condicionales $C_{2|1}(u_1, u_2)$ y sus inversas. La extensión al caso d dimensional utiliza el mismo principio que el algoritmo bidimensional y se encuentra en Romano [2002].

- *Algoritmo para simular variables aleatorias que sigan cópulas de Arquímedes.*

Una aplicación explícita del algoritmo anterior se encuentra en las metodologías utilizadas para la simulación de cópulas de Arquímedes. Para esta familia de cópulas, es posible mostrar que la distribución marginal de u_k condicionada a los valores de u_1, \dots, u_{k-1} es:

$$C_{k|k-1, \dots, 1}(u_1, \dots, u_k) = \frac{\phi^{-1(k-1)}(\phi(u_1) + \phi(u_2) + \dots + \phi(u_k))}{\phi^{-1(k-1)}(\phi(u_1) + \phi(u_2) + \dots + \phi(u_{k-1}))} \quad (3.55)$$

donde $\phi^{-1(k-1)}$ es la $(k-1)$ -ésima derivada de la función inversa del generador ϕ de la cópula de Arquímedes C , con $k = 2, \dots, d$. Esta forma funcional permite obtener algoritmos explícitos para las cópulas de Arquímedes más usadas. Por ejemplo, para la cópula Clayton, su generador es $\phi(t) = t^{-\theta} - 1$ y por lo tanto $\phi^{-1}(t) = (t+1)^{-1/\theta}$ ¹². De aquí, el procedimiento para simular variables aleatorias uniformes que sigan este tipo de cópula es [Cherubini et al., 2004]:

i) Genere n simulaciones de d variables aleatorias independientes que sigan una distribución $U(0, 1)$, w_{i1}, \dots, w_{id} , y defina $u_{i1} = w_{i1}$, con $i = 1, \dots, n$.

ii) Se define $w_{i2} \equiv C_{2|1}(u_{i1}, u_{i2})$. A partir de (3.55) se tiene que $w_{i2} = \left(\frac{u_{i1}^\theta + w_{i2}^\theta - 1}{u_{i1}^\theta}\right)^{-1/\theta-1}$.

De aquí, $u_{i2} = \left(u_{i1}^{-\theta} \left(w_{i2}^{-\theta/(1+\theta)} - 1\right) + 1\right)^{-1/\theta}$.

iii) Para $k = 3, \dots, d$, la variable aleatoria u_{ik} esta definida como:

$$u_{ik} = \left(\left(u_{i1}^{-\theta} + \dots + u_{i(k-1)}^{-\theta} - k + 2\right) \left(w_{ik}^{\theta/(\theta(1-k)-1)} - 1\right) + 1\right)^{-1/\theta}$$

donde el vector $\mathbf{u}_i = [u_{i1}, \dots, u_{id}]$ sigue una cópula Clayton d dimensional con parámetro θ , $i = 1, \dots, n$.

Por su parte, para la cópula Frank, el generador es $\phi(u) = -\log\left(\frac{e^{-\theta u} - 1}{e^{-\theta} - 1}\right)$, de aquí, $\phi^{-1}(t) = -\frac{1}{\theta} \log(1 + e^{-t}(e^{-\theta} - 1))$. Utilizando estos resultados, el algoritmo para generar variables aleatorias uniformes que sigan esta cópula son:

i) Genere n simulaciones de d variables aleatorias independientes que sigan una distribución $U(0, 1)$, w_{i1}, \dots, w_{id} , y defina $u_{i1} = w_{i1}$, con $i = 1, \dots, n$.

¹²Vale la pena destacar que el generador de la cópula Clayton es $\phi(t) = \theta^{-1}(t^{-\theta} - 1)$. No obstante, el factor θ^{-1} es una constante que no resulta relevante para el algoritmo de simulación.

ii) Como en el caso anterior, defina $w_{i2} \equiv C_{2|1}(u_{i1}, u_{i2})$. Usando las definiciones anteriores y la ecuación (3.55), $w_{i2} = e^{-\theta u_1} \frac{e^{-\theta u_2} - 1}{e^{-\theta} - 1 + (e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}$. Despejando u_2 , se tiene que $u_{i2} = -\frac{1}{\theta} \log \left(1 + \frac{w_{i2}(1 - e^{-\theta})}{w_{i2}(e^{-\theta u_1} - 1) - e^{-\theta u_1}} \right)$.

iii) Análogamente, defina $w_{i3} \equiv C_{3|1,2}(u_{i1}, u_{i2}, u_{i3})$. Utilizando nuevamente (3.55), se obtiene que $w_{i3} = \frac{(e^{-\theta} - 1)[(e^{-\theta} - 1) + (e^{-\theta u_{i1}} - 1)(e^{-\theta u_{i2}} - 1)]^2 (e^{-\theta u_{i3}} - 1)}{[(e^{-\theta} - 1)^2 + (e^{-\theta u_{i1}} - 1)(e^{-\theta u_{i2}} - 1)(e^{-\theta u_{i3}} - 1)]^2}$. Esta función es una ecuación polinomial de grado 2 sobre $x = e^{-\theta u_{i3}} - 1$. Este es un ejemplo en el cual deben ser utilizados métodos numéricos para encontrar un valor aproximado de u_{i3} .

Este procedimiento es extendido para u_{ik} , $k = 4, \dots, d$. Nótese que como en el punto anterior, la implementación de este procedimiento implica la solución de una ecuación polinomial de grado $k - 1$, lo que imprime un mayor nivel de dificultad para la implementación de este algoritmo. El vector $\mathbf{u}_i = [u_{i1}, \dots, u_{id}]$ sigue una cópula Frank d dimensional con parámetro θ , $i = 1, \dots, n$.

Este último punto destaca la complejidad que puede alcanzar la implementación de algunos algoritmos que a simple vista parecen sencillos. Incluso si se tiene una forma simple de obtener las funciones inversas de las distribuciones marginales condicionales, $C_{k|1, \dots, k-1}$, como en el caso de la cópula Clayton, repetir este procedimiento para un gran número de simulaciones (denotadas por el subíndice i en las técnicas mencionadas), se convierte en una tarea altamente costosa desde un punto de vista computacional. McNeil et al. [2005] proponen otro tipo de algoritmos para simular variables aleatorias que sigan cópulas de Arquímedes y presentan metodologías específicas para las cópulas Clayton, Frank y Gumbel.

Para el caso de la cópula Clayton, se generan n replicas de una variable aleatoria v_i que siga una distribución gamma con parámetros $(1/\theta, 0)$, $\theta > 0$. Posteriormente, se define el vector $\mathbf{u}_i = [\phi^{-1}(-\log(w_{i1})/v_i), \dots, \phi^{-1}(-\log(w_{id})/v_i)]$, donde $\phi^{-1}(t) = (t + 1)^{-1/\theta}$ es la función inversa del generador de la cópula Clayton y w_{i1}, \dots, w_{id} son variables aleatorias uniformes $(0, 1)$ independientes. El vector $\mathbf{u}_i = [u_{i1}, \dots, u_{id}]$ sigue una cópula Clayton d dimensional con parámetro θ , $i = 1, \dots, n$.

En el caso de la cópula Gumbel, se generan n replicas de una variable aleatoria v_i que siga una distribución estable positiva con parámetros $(1/\theta, 1, \gamma, 0)$, donde $\gamma = (\cos(\pi/(2\theta)))^\theta$ y $\theta > 1$ ¹³. Posteriormente, se define el vector $\mathbf{u}_i = [\phi^{-1}(-\log(w_{i1})/v_i), \dots, \phi^{-1}(-\log(w_{id})/v_i)]$, donde $\phi^{-1}(t) = \exp(-t^{1/\theta})$ es la función inversa del generador de la cópula Gumbel y w_{i1}, \dots, w_{id} son variables aleatorias uniformes $(0, 1)$ independientes. El vector $\mathbf{u}_i = [u_{i1}, \dots, u_{id}]$ sigue una cópula Gumbel d dimensional con parámetro θ , $i = 1, \dots, n$.

¹³Para una descripción del algoritmo de simulación de variables aleatorias que sigan una distribución estable, véase Nolan [2007].

Finalmente, para la cópula Frank, se generan n replicas de una variable aleatoria discreta v_i cuya función de densidad es $P(V = v) = (v\theta)^{-1} (1 - e^{-\theta})^v$, con $v = 1, 2, \dots$ y $\theta > 0$. Posteriormente, se define el vector \mathbf{u}_i como $\mathbf{u}_i = [\phi^{-1}(-\log(w_{i1})/v_i), \dots, \phi^{-1}(-\log(w_{id})/v_i)]$, donde $\phi^{-1}(t) = -\frac{1}{\theta} \log(1 + e^{-t}(e^{-\theta} - 1))$ es la función inversa del generador de la cópula Frank y w_{i1}, \dots, w_{id} son variables aleatorias uniformes $(0, 1)$ independientes. El vector $\mathbf{u}_i = [u_{i1}, \dots, u_{id}]$ sigue una cópula Frank d dimensional con parámetro θ , $i = 1, \dots, n$.

- *Algoritmo para simular cópulas Arquimax.* La simulación de variables aleatorias uniformes u_1 y u_2 que sigan una cópula Arquimax $C_{\phi,A}$ se basa en la proposición 5.1 de Capéraà et al. [2000]. El procedimiento sugerido por estos autores consta de cuatro pasos:
 - i) Se genera una realización de una variable aleatoria Z, z_i , que siga una distribución H , donde $H(z) \equiv \text{prob}(Z \leq z)$ y $Z = \frac{\phi(U)}{\phi(U) + \phi(V)}$.
 - ii) Posteriormente, se genera una realización de una variable aleatoria uniforme estándar, v_i .
 - iii) Si $v_i \leq p(z_i)$ se genera w_i a partir de una distribución uniforme estándar, de lo contrario, w_i se obtiene de la distribución K_ϕ , donde $p(z) = \frac{z(1-z)A''(z)}{H'(z)A(z)}$ y K_ϕ corresponde a la función K definida en (4.14).
 - iv) Finalmente, u_1 y u_2 son calculadas de la siguiente forma: $u_1 = \phi^{-1}\left[\frac{z\phi(w)}{A(z)}\right]$ y $u_2 = \phi^{-1}\left[\frac{(1-z)\phi(w)}{A(z)}\right]$.
- *Algoritmo para simular variables aleatorias a partir de la cópula empírica.*

A partir del principio de bootstrapping, Romano [2002] menciona el siguiente algoritmo para la generación de variables aleatorias basado en la cópula empírica.

- i) Obtenga de forma aleatoria (con reemplazamiento), n vectores observados de la base de datos histórica $\mathbf{r}_i = [r_{i1}, \dots, r_{id}]$, $i = 1, \dots, n$.
- ii) Defina el vector $\mathbf{u}_i = [\hat{F}_1(r_{i1}), \dots, \hat{F}_d(r_{id})]$, donde \hat{F}_j corresponde a la función de distribución empírica para la variable j . Por lo tanto, el vector $\mathbf{u}_i = [u_{i1}, \dots, u_{id}]$ es un vector cuya relación de dependencia esta descrita por la cópula empírica, $i = 1, \dots, n$.

4. ESTIMACIÓN.

En la Sección 3 se mencionaron algunos conceptos generales de la cópula, su definición, propiedades y sus categorías más relevantes. En los párrafos siguientes se describen diferentes métodos de estimación de los parámetros asociados a una cópula, dado un conjunto de realizaciones de un vector aleatorio *i.i.d.* con una función de distribución multivariada F y una cópula asociada C , la cual por simplicidad se asume conocida. Posteriormente, se introducen algunos criterios de selección que facilitan la identificación de

C entre las diferentes familias de cópula. Adicionalmente, se relaja el supuesto en el cual las observaciones provienen de un vector aleatorio *i.i.d.*, permitiendo dependencia en el tiempo.

4.1. Métodos de estimación.

4.1.1. *Estimación no paramétrica.* Al igual que en el caso de funciones de distribución univariadas, los estimadores no paramétricos desempeñan un papel importante para las cópulas. Al interior de esta categoría, el estimador más importante es la cópula empírica, inicialmente definida por Deheuvels [1979, 1981] bajo el nombre "función de dependencia empírica", pertenece a este tipo de estimadores.

Para una muestra de tamaño n , $\{u_{1,k}, \dots, u_{d,k}\}_{k=1}^n$, de una distribución continua d -dimensional, la cópula empírica es definida de la siguiente forma:

$$\widehat{C}_{emp} \left(\frac{i_1}{n}, \dots, \frac{i_d}{n} \right) = \frac{1}{n} \sum_{k=1}^n I_{\{u_{1,k} \leq u_1(i_1), \dots, u_{d,k} \leq u_d(i_d)\}} \quad (4.1)$$

donde $u_m(i_k)$ son las estadísticas de orden asociadas a u_m , $i_k = 1, \dots, n$ y $k, m = 1, \dots, d$.

Al igual que cualquier cópula, la expresión (4.1) corresponde a una función de distribución. Siguiendo la notación de Nelsen [2006], la función de densidad asociada a la cópula empírica se define como:

$$\widehat{c}_{emp} \left(\frac{i_1}{n}, \dots, \frac{i_d}{n} \right) = \sum_{k_1=1}^2 \dots \sum_{k_d=1}^2 (-1)^{\sum_{j=1}^d k_j} \times \widehat{C}_{emp} \left(\frac{i_1 - k_1 + 1}{n}, \dots, \frac{i_d - k_d + 1}{n} \right) \quad (4.2)$$

Adicionalmente, Deheuvels [1978, 1981] prueban que la cópula empírica converge uniformemente a la cópula verdadera de la cual fue seleccionada la muestra. Este resultado implica que cuando la cópula verdadera es desconocida, un criterio de selección puede ser el comparar cada una de las cópulas candidatas con la empírica, tal como se muestra en la Sección 4.3.

Sin embargo, la definición (4.1) no corresponde al único estimador no paramétrico de cópulas, Cherubini et al. [2004] discuten otros estimadores en este contexto. En particular, estos autores consideran métodos basados en aproximaciones polinómicas y en funciones *kernel*.

4.1.2. *Estimación a través de las correlaciones de rango (Método de momentos).* Este tipo de metodologías se basan en las relaciones existentes entre dos medidas no paramétricas, como son el coeficiente de correlación de Spearman (ρ_S) y el τ de Kendall (ρ_τ), y las cópulas (ver Sección 3.3.1). Esta metodología tiene la ventaja que no necesita información acerca de las distribuciones marginales (individuales) de los datos, como se muestra a continuación:

Si se tienen n realizaciones de un vector aleatorio *i.i.d.* de dimensión d , $R_t = [R_{1t}, \dots, R_{dt}]'$, R_1, \dots, R_n , este método se puede implementar en dos etapas:

- i) Se calcula de forma no paramétrica la matriz de correlaciones de rango $\tilde{\mathbf{P}} \in \mathbb{R}^{d \times d}$, la cual es simétrica y semidefinida positiva¹⁴. En esta matriz, el elemento que se encuentra en la posición (i, j) corresponde al coeficiente de correlación de rango de Spearman (o el τ de Kendall) para las series R_{it} y R_{jt} , $t = 1, \dots, n$, que pertenecen al vector aleatorio $R_t = [R_{1t}, \dots, R_{dt}]'$. Un estimador del coeficiente de correlación de rango de Spearman se puede calcular como:

$$r_S(R_i, R_j) = 1 - \frac{6}{n(n-1)} \sum_{t=1}^n d_t^2 \quad (4.3)$$

donde d_t es la diferencia entre los rangos R_{it} y R_{jt} . El rango de R_{it} corresponde a la posición que tiene R_{it} dentro del conjunto de observaciones ordenadas de la muestra, R_{i1}, \dots, R_{in} . De otro lado, un estimador a partir de rangos del τ de Kendall para las variables R_i y R_j es:

$$\hat{\tau}(R_i, R_j) = \binom{n}{2}^{-1} \sum_{t=1}^{n-1} \sum_{s=t+1}^n \text{signo}((R_{it} - R_{is})(R_{jt} - R_{js})) \quad (4.4)$$

- ii) De forma similar al método de momentos, se forma un sistema de ecuaciones a partir de la igualación del resultado teórico obtenido de las integrales definidas en (3.17) (o (3.18)), con los coeficientes de correlación de rango estimados en (4.3) (o (4.4)). Dado que los resultados de estas integrales dependen de los parámetros de la cópula, en los casos que el sistema de ecuaciones se encuentra identificado, se puede obtener un estimador de los parámetros de la cópula.

Un ejemplo sencillo de esta metodología se tiene en el caso de las cópulas de Arquímedes bivariadas de un solo parámetro. En el caso de la cópula Gumbel, un estimador del parámetro θ basado en este método es $\hat{\theta} = 1/(1 - r_\tau(R_1, R_2))$, dado que $\rho_\tau(R_1, R_2) = 1 - 1/\theta$ ¹⁵.

4.1.3. Método de máxima verosimilitud. Sea $\mathbf{R} = [R_1, \dots, R_d]'$ un vector aleatorio conformado por d variables aleatorias continuas con distribuciones marginales F_1, \dots, F_d , con función de distribución conjunta F y cópula C . Si se tiene una muestra de n realizaciones *i.i.d.* del vector aleatorio \mathbf{R} , $\mathbf{r}_1, \dots, \mathbf{r}_n$, es posible obtener una estimación de los parámetros de la cópula a través de la maximización de su función de log-verosimilitud, la cual

¹⁴En la práctica, puede que la matriz de correlación de rango estimada no sea semidefinida positiva. En estos casos, es posible utilizar el algoritmo del valor propio. Para una mayor discusión acerca de éste algoritmo, véase Rousseeuw y Molenberghs [1993]

¹⁵Véase Sección 3.5.2.

esta definida como:

$$\ell(\theta | \mathbf{r}_1, \dots, \mathbf{r}_n) = \sum_{t=1}^n \log c(F_1(r_{1t}), \dots, F_d(r_{dt})) \quad (4.5)$$

$$\ell(\theta | \mathbf{u}_1, \dots, \mathbf{u}_n) = \sum_{t=1}^n \log c(u_{1t}, \dots, u_{dt})$$

donde θ es el vector de parámetros de la cópula C , $u_{it} = F_i(r_{it})$ y $\mathbf{u}_t = [u_{1t}, \dots, u_{dt}]'$. De los resultados presentados en (3.4) y (3.7), se sabe que $u_{it} \sim U(0, 1)$ y $\mathbf{u}_t \stackrel{i.i.d.}{\sim} C$. Adicionalmente, c corresponde a la función de densidad de la cópula C , la cual se define como:

$$c(u_1, \dots, u_d) = \frac{\partial^d C(u_1, \dots, u_d)}{\partial u_1 \dots \partial u_d} \quad (4.6)$$

expresando (4.6) en términos de la función de densidad del vector \mathbf{R} se obtiene:

$$c(u_1, \dots, u_d) = \frac{f(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))}{f_1(F_1^{-1}(u_1)) \dots f_d(F_d^{-1}(u_d))}$$

donde f corresponde a la función de densidad conjunta de \mathbf{R} y f_1, \dots, f_d son las funciones de densidad marginales [McNeil et al., 2005].

La implementación de este procedimiento es más complicada que la basada en las correlaciones de rango. Por un lado, se requiere la función de densidad de la cópula c , la que en algunos casos es difícil de derivar; además, este procedimiento necesita estimaciones de las funciones de distribución marginales de \mathbf{R} . Por otro lado, en la mayoría de los casos el proceso de maximización de la función de verosimilitud se lleva a cabo mediante procedimientos numéricos, lo que implica un mayor costo computacional.

Sin embargo, a diferencia del método de correlación de rango, el cual es aplicable solo en los casos en que el sistema de ecuaciones resultante se encuentra identificado¹⁶, el método de máxima verosimilitud se puede aplicar a cualquier familia de cópula. Otra ventaja de este último método es que, basado en las propiedades asintóticas que comparten todos los estimadores de máxima verosimilitud, ofrece información adicional acerca de los estimadores, tales como sus errores estándar y sus distribuciones asintóticas.

Como se mencionó anteriormente, dada una cópula C el método de máxima verosimilitud basado en (4.5) requiere información acerca de una muestra aleatoria $\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_n$, la cual se encuentra relacionada con la muestra aleatoria inicial $\mathbf{r}_1, \dots, \mathbf{r}_n$ a través de las funciones de distribución marginales F_1, \dots, F_d mediante la siguiente expresión: $\hat{\mathbf{u}}_t = [\hat{F}_1(r_{1t}), \dots, \hat{F}_d(r_{dt})]' = [\hat{u}_{1t}, \dots, \hat{u}_{dt}]'$, $t = 1, \dots, n$. Esta nueva muestra es llamada una pseudo-muestra aleatoria.

¹⁶En los casos en que este sistema no se encuentra plenamente identificado es posible utilizar una aproximación que combine los métodos de correlación de rango y de máxima verosimilitud. Un ejemplo de esta metodología aplicado a la cópula t se encuentra en Mashal y Zeevi [2002] y Demarta y McNeil [2005].

Si la cópula C y las funciones de distribución marginales son conocidas, se podrían incluir estas últimas dentro de (4.5) y estimar los parámetros de la cópula (θ) y los de las distribuciones marginales (ϕ_1, \dots, ϕ_d) de forma conjunta, a partir de la maximización de la siguiente función de verosimilitud¹⁷:

$$\ell(\theta, \phi_1, \dots, \phi_d | \mathbf{r}_1, \dots, \mathbf{r}_n) = \sum_{t=1}^n \log c(F_1(r_{1t}), \dots, F_d(r_{dt})) \quad (4.7)$$

de esta manera, la pseudo muestra se encuentra dentro de la función de verosimilitud (4.7). Este procedimiento paramétrico es conocido como el método de "Inference Functions for Margins (IFM)" [Joe, 1997].

El IFM requiere hacer supuestos sobre las distribuciones marginales, y resulta costoso en términos computacionales. Para reducir estos costos, Joe [1997] sugiere implementar este procedimiento en dos etapas: en la primera de ellas se generan las pseudo muestras $\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_n$ a partir de la estimación de los parámetros de las funciones de distribución marginales (ϕ_i) por máxima verosimilitud, la cual se hace de manera separada para $i = 1, \dots, d$. Finalmente, se maximiza (4.5) para obtener los parámetros de la cópula.

McNeil et al. [2005] describe un procedimiento de dos etapas en el que se generaliza el método IFM. En la primera etapa, se estiman las funciones de distribución marginales de forma paramétrica (IFM) o no paramétrica, para generar la pseudo muestra $\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_n$ y posteriormente, en la segunda etapa se estiman los parámetros de la cópula a partir de la maximización de (4.5).

Uno de los posibles estimadores no paramétricos que se pueden usar para generar la pseudo muestra $\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_n$, y quizá el más natural, se encuentra asociado con la función de distribución empírica F_i^* . Esta función se encuentra definida de la siguiente forma:

$$F_i^*(r) = \frac{1}{n+1} \sum_{t=1}^n I_{\{R_{i,t} \leq r\}} \quad i = 1, \dots, d \quad (4.8)$$

donde es $I_{\{D\}}$ es una función indicadora que toma el valor de uno si se cumple la condición D y cero en otro caso [McNeil et al., 2005].

Carmona [2004] destaca que la función de distribución empírica no es apropiada para la estimación de las funciones de distribución marginales en las colas, ya que existen muy pocas observaciones para obtener una correcta aproximación en los extremos de la distribución. En consecuencia, utiliza un estimador semi-paramétrico que involucra la teoría del valor extremo (EVT por sus siglas en inglés). En esta metodología las colas de la distribución son modeladas usando una distribución generalizada de Pareto, mientras que el centro de la distribución es modelado utilizando la función de distribución empírica [Zivot y Wang, 2006].

¹⁷Por ejemplo, asumiendo que las distribuciones marginales son normales, se tiene que $F_i(r_{it}) \sim N(\mu_i, \sigma_i^2)$, $t = 1, \dots, n$, $i = 1, \dots, d$. Por lo tanto, $\phi_i = [\mu_i, \sigma_i^2]'$.

4.2. Verificación de supuestos. Es importante destacar que la estimación de la cópula C por el método de máxima verosimilitud supone que:

- i)* La cópula C es conocida.
- ii)* \mathbf{r}_t y \mathbf{u}_t son *i.i.d.*
- iii)* $u_{it} = \widehat{F}_i(r_{it})$ sigue una distribución uniforme entre 0 y 1.

El supuesto *i)* resulta complicado de satisfacer, ya que en la mayoría de los casos no se conoce a priori cuál es la cópula más apropiada para los datos; en estos casos es común estimar varias cópulas y buscar un criterio de selección para escoger la más apropiada (véase Sección 4.3). Por otro lado, el supuesto *ii)* tiene grandes implicaciones en la estimación, ya que si los datos no son *i.i.d.*, la función de verosimilitud presentada en (4.5) no es adecuada y por consiguiente la estimación pierde validez; el desarrollo de modelos dinámicos que abordan la dependencia temporal en \mathbf{r}_t es tratado en la Sección 4.4.

El supuesto *iii)* implica que las funciones de distribución marginales se especificaron correctamente. En los casos en que la función de distribución esta mal especificada, no es posible garantizar que la transformación $\widehat{F}_i(r_{it})$ siga una distribución uniforme por lo cual el Teorema 3.1 no está correctamente representado.

Para ilustrar las implicaciones que tiene la mala especificación de la función de distribución en la posterior definición de la pseudo-muestra, en el Gráfico 4.1 se evaluaron 2,000 realizaciones de una variable aleatoria normal estándar en dos funciones de distribución, G_1 y G_2 . Si el supuesto *iii)* se mantiene ante una mala especificación de la distribución marginal, al realizar los gráficos de Q-Q de $u_1 = G_1(z)$ y $u_2 = G_2(z)$ contra una distribución uniforme con parámetros 0 y 1 estos deberían formar una línea recta con un ángulo de 45° .

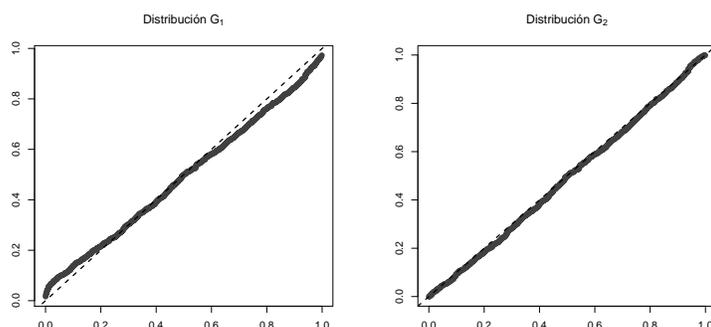


GRÁFICO 4.1. Gráficos de Q-Q que muestran los efectos de la mala especificación de las distribuciones marginales en la construcción de la pseudo muestra.

En el panel izquierdo del Gráfico 4.1 se utiliza una distribución t con $v = 3$ (G_1), mientras que en el panel derecho se usa la distribución normal estándar (G_2). Es posible notar que en el caso que se especifica mal la función de distribución, la transformación $G_1(z)$ se aleja del comportamiento de una variable aleatoria uniforme.

Aunque la cópula es invariante ante transformaciones monótonas crecientes de los datos, y la función de distribución acumulada es una función monótona creciente, las estimaciones de los parámetros de la cópula no son adecuadas ante malas especificaciones de las funciones de distribución marginales. Esto se debe a que, como se mostró anteriormente, la transformación $\hat{F}_i(r_{it})$ no sigue una distribución uniforme en los casos en que la función de distribución se encuentra mal especificada¹⁸.

Para verificar que \mathbf{u}_t es *i.i.d.*, Patton [2001] sugiere el uso de la prueba propuesta por Diebold et al. [1998] en la que se prueba independencia de los primeros cuatro momentos de u_{1t}, \dots, u_{dt} ¹⁹. Para el caso bivariado, se realizan las regresiones de $\tilde{u}_{1t}^k = (u_{1t} - \bar{u}_1)^k$ y $\tilde{u}_{2t}^k = (u_{2t} - \bar{u}_2)^k$ contra Q rezagos de \tilde{u}_{1t}^k y \tilde{u}_{2t}^k para $k = 1, 2, 3, 4$ y se calcula el estadístico $J = (T - 2Q) * R^2$ para cada regresión. Bajo la hipótesis nula todos los coeficientes de estas regresiones son cero, y el estadístico J distribuye χ^2 con $2Q$ grados de libertad.

Adicionalmente, Patton [2001] utiliza la prueba de Kolmogorov–Smirnov para probar si la serie u_{it} sigue una distribución uniforme entre 0 y 1 (supuesto *iii*). Sin embargo, Patton [2001] menciona dos desventajas acerca del uso de este procedimiento: primero, estas pruebas contrastan separadamente los supuestos *ii*) y *iii*), cuando sería deseable utilizar una prueba conjunta; segundo, la prueba de Kolmogorov–Smirnov tiene baja potencia en las colas de la distribución, las cuales tienen gran importancia en el cálculo de medidas de riesgo tales como el valor en riesgo, VaR ²⁰.

4.3. Selección de la cópula. En la sección anterior, se presentaron diferentes procedimientos para la estimación de una cópula de forma paramétrica o semi-paramétrica. Sin embargo, estos procedimientos suponen que se conoce de antemano el tipo de cópula que describe correctamente la estructura de dependencia que existe entre las series analizadas. En general, este supuesto no se cumple, por lo que es necesario obtener algún tipo de criterio que permita escoger una cópula específica entre el conjunto de todas las cópulas posibles. Algunos de los criterios más conocidos son:

4.3.1. Log verosimilitud y criterios de información. Joe [1997] propone el uso de dos estadísticos como criterios para comparar varias cópulas y entre estas seleccionar la más apropiada. El primero de ellos es el valor de la función de log verosimilitud evaluada

¹⁸No obstante, este inconveniente no se presenta cuando se realiza la estimación de los parámetros de la cópula a través de métodos que no realicen supuestos sobre F_i , tales como el método de correlaciones de rango y en el caso en que se utiliza la función de distribución empírica F_i^* .

¹⁹Esta prueba fue extendida al caso multivariado por Diebold et al. [1999].

²⁰Una prueba alternativa a la de Kolmogorov–Smirnov es el “Hit test”, el cual fue planteado por Diebold et al. [1998].

en $\tilde{\theta}$ (ecuación (4.5)), donde $\tilde{\theta}$ corresponde al estimador de máxima verosimilitud²¹. El segundo estadístico sugerido por Joe [1997] es el criterio de información de Akaike (AIC por sus siglas en inglés), definido como $AIC = \ell - n_p$, donde n_p corresponde al número de parámetros estimados en el modelo. Esta última definición difiere de la definición usual de este criterio de información donde $AIC = -2\ell + 2n_p$; sin embargo, Joe [1997] propone la primera definición de este criterio, como una especie de log verosimilitud penalizado por el número de parámetros estimados.

Adicional a los anteriormente mencionados, Zivot y Wang [2006] utilizan otros dos criterios de información para la selección de la cópula. Estos son el criterio de información Bayesiano (BIC por sus siglas en inglés) y el criterio de información de Hannan y Quinn (HQ). Definidos de la forma:

$$BIC = -2\ell + n_p \log(n) \quad (4.9)$$

$$HQ = -2\ell + 2n_p \log(\log(n)) \quad (4.10)$$

donde n_p representa el número de parámetros estimados y n es el tamaño de la muestra.

4.3.2. *Comparación entre cópulas paramétricas y la cópula empírica.* Otro posible criterio de selección es el uso de estadísticos de bondad de ajuste (GOF por sus siglas en inglés), en los que se determina cuál es la cópula que describe de mejor manera el comportamiento observado en los datos. Romano [2002] propone una estadística que pretende generalizar las estadísticas usadas en las pruebas GOF univariadas, calculando la distancia entre la cópula teórica estimada y la cópula empírica, de la forma:

$$d(C_k, C_{emp}) = \sqrt{\sum_{t_1=1}^n \cdots \sum_{t_d=1}^n \left(\hat{C}_k \left(\frac{t_1}{n}, \dots, \frac{t_d}{n} \right) - \hat{C}_{emp} \left(\frac{t_1}{n}, \dots, \frac{t_d}{n} \right) \right)^2} \quad (4.11)$$

donde \hat{C}_{emp} corresponde a la cópula empírica (ver ecuación (4.1)) y \hat{C}_k es la cópula teórica estimada. Así, la cópula mas apropiada será la que minimice (4.11). No obstante, Romano [2002] no analiza las propiedades de este tipo de estadístico.

En este mismo orden de ideas, otras medidas que se utilizan para comparar la cópula teórica estimada y la cópula empírica se basan en la estadística de Anderson y Darling [1952]. Al respecto, Junker y May [2005] y Ané y Kharoubi [2003] utilizan los siguientes

²¹No obstante, Patton [2001] llama la atención acerca de la comparación de los valores de log verosimilitud para la selección de la cópula, ya que en el caso de modelos no anidados esta aproximación no es valida. Se dice que la cópula C_1 se encuentra anidada dentro de la cópula C_2 si es posible encontrar un conjunto de restricciones sobre los parámetros de C_2 tal que esta sea igual a C_1 .

estadísticos:

$$AD = \max_{1 \leq t_1 \leq n, \dots, 1 \leq t_d \leq n} \frac{\sqrt{n} \left| \widehat{C}_k \left(\frac{t_1}{n}, \dots, \frac{t_d}{n} \right) - \widehat{C}_{emp} \left(\frac{t_1}{n}, \dots, \frac{t_d}{n} \right) \right|}{\sqrt{\widehat{C}_k \left(\frac{t_1}{n}, \dots, \frac{t_d}{n} \right) \left(1 - \widehat{C}_k \left(\frac{t_1}{n}, \dots, \frac{t_d}{n} \right) \right)}} \quad (4.12)$$

$$IAD = \frac{1}{n} \sum_{t_1=1}^n \dots \sum_{t_d=1}^n \frac{\left[\widehat{C}_k \left(\frac{t_1}{n}, \dots, \frac{t_d}{n} \right) - \widehat{C}_{emp} \left(\frac{t_1}{n}, \dots, \frac{t_d}{n} \right) \right]^2}{\widehat{C}_k \left(\frac{t_1}{n}, \dots, \frac{t_d}{n} \right) \left(1 - \widehat{C}_k \left(\frac{t_1}{n}, \dots, \frac{t_d}{n} \right) \right)} \quad (4.13)$$

Junker y May [2005] utilizan los valores críticos tabulados por Giles [2000] para evaluar la bondad de ajuste de la cópula estimada a partir de la estadística IAD. Sin embargo, Junker y May mencionan que estos estadísticos deben ser interpretados con cautela, ya que la prueba IAD tiene problemas en estas aplicaciones y por lo tanto, los valores críticos solo deben servir como un indicador adicional.

Finalmente, Malevergne y Sornette [2006] presentan varias medidas de bondad de ajuste adicionales; sin embargo, la mayoría de pruebas se desarrollan para la cópula Gaussiana. Cabe destacar que una de las estadísticas consideradas por estos autores tiene un enfoque diferente. En este caso, la prueba se basa en la estadística de Anderson y Darling para la función K (ecuación (4.14)) y no directamente sobre las cópulas.

Adicional a las técnicas anteriormente reseñadas, existen otras metodologías gráficas y paramétricas para seleccionar cópulas en el caso bivariado. Estas metodologías se basan en la función de distribución de la variable aleatoria $V = C(F_1(R_1), F_2(R_2))$. Genest y Rivest [2001] mostraron que a diferencia del caso univariado, donde la definición (3.4) asegura que la variable aleatoria $F_i(R_i) \sim U(0, 1)$, la función de distribución de la variable aleatoria V , $K(v)$, no necesariamente sigue esa misma distribución. En este caso, $K(v)$ tiene la forma:

$$K(v) = v - \lambda(v) \quad (4.14)$$

donde la función $\lambda(v) : [0, 1] \rightarrow [v - 1, 0]$ se encuentra determinada por la familia de cópulas a la cual pertenezca C^{22} . Algunos casos específicos de formas funcionales de $\lambda(v)$ y $K(v)$ se presentan en la Tabla 4.1.

En esta tabla, ϕ representa la función generadora, $\tau_A = \int_0^1 t(1-t)/A(t)A''(t) dt$ es el τ de Kendall, y A es la función de dependencia, tal como se encuentra definida en la Sección 3.5.3. Las dos últimas filas de la Tabla 4.1 corresponden a las distribuciones límite de $K(v)$ cuando las variables aleatorias R_1 y R_2 son comonotónicas y contramonotónicas, respectivamente [Genest y Rivest, 1993, 2001; Cebrian et al., 2002]²³.

A partir de los resultados de Genest y Rivest sobre $K(v)$, se han desarrollado algunos métodos gráficos y pruebas estadísticas que pueden aportar información acerca del ajuste

²²La forma en la que se deriva la función $\lambda(v)$ se encuentra en Genest y Rivest [2001].

²³En los casos de cópulas de más de dos dimensiones, véase Genest y Rivest [2001].

Tipos de cópula	$\lambda(v)$	$K(v)$
Arquímedes	$\lambda_\phi(v) = \phi(v)\phi'(v)$	$v - \phi(v)\phi'(v)$
Valor extremo	$\lambda_A(v) = (1 - \tau_A)v \log(v)$	$v - (1 - \tau_A)v \log(v)$
Arquimax	$\lambda_{\phi,A}(v) = (1 - \tau_A)\lambda_\phi(v)$	$v - (1 - \tau_A)\lambda_\phi(v)$
Comonotonicidad	$\lambda(v) = 0$	v
Contramonicidad	$\lambda(v) = v - 1$	1

TABLA 4.1. Formas funcionales de $\lambda(v)$ y $K(v)$ para algunos tipos de cópulas.

de los datos a una familia de cópulas específica. Por ejemplo, si la función de distribución de los datos se encuentra relacionada con una cópula de la Familia de Arquímedes, Frees y Valdez [1998] plantean la aplicación de gráficos de Q-Q para la identificación de la función generadora de la cópula de Arquímedes que ajuste mejor a los datos. Este procedimiento consta de los siguientes pasos:

- i) Se calcula el estimador no paramétrico del τ de Kendall $\hat{\tau}$, de la misma forma que en la ecuación (4.4).
- ii) Se construye un estimador no paramétrico de $K(v)$, denominado $K_n(v)$, de la forma siguiente:
 - a) Se genera un conjunto de pseudo observaciones, V_i , definidas como:

$$V_i = \frac{1}{n-1} \sum_{j=1}^n I_{\{R_{1j} < R_{1i}\}} I_{\{R_{2j} < R_{2i}\}} \quad i = 1, \dots, n \quad (4.15)$$

- b) A partir de las pseudo observaciones, se calcula un estimador de K como:

$$K_n(v) = \frac{1}{n} \sum_{i=1}^n I_{\{V_i \leq v\}} \quad (4.16)$$

- iii) Se construye un estimador de K a partir de la relación (4.14) aplicada a diferentes especificaciones de ϕ asociadas a cópulas de Arquímedes, $K_\phi(v)$ (primera fila de la Tabla 4.1), a partir de la relación existente entre los parámetros de la cópula y el $\hat{\tau}$ estimado en el primer paso (véase Sección 4.1.2).
 - iv) Finalmente, se realiza un gráfico de Q-Q comparando los cuantiles de $K_n(v)$ y $K_\phi(v)$ para cada una de las cópulas de Arquímedes consideradas en el numeral anterior.

Una aplicación de este método se presenta en el Gráfico 4.2, donde se utilizan los gráficos de Q-Q para 1,500 simulaciones de una pareja de variables aleatorias que siguen una cópula Gumbel con parámetro $\theta = 2,5$. La función de distribución empírica para esta pareja de variables, $K_n(v)$, fue comparada contra cuatro funciones de distribución $K_\phi(v)$ teóricas, asociadas a las cópulas de Arquímedes Gumbel, Clayton, Frank y de independencia, respectivamente.

Como era de esperarse de las cópulas consideradas, la Gumbel es la que presenta el mejor ajuste. Vale la pena destacar que en esta metodología se supone que la función de distribución multivariada de los datos esta asociada con una cópula de Arquímedes.

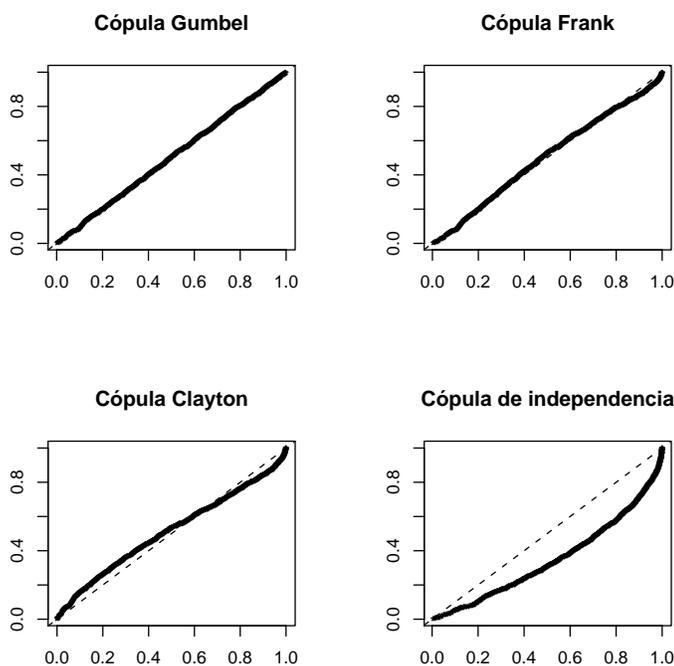


GRÁFICO 4.2. Selección de la cópula de Arquímedes a través del método de Frees y Valdez [1998] basado en Gráficos Q-Q.

Junto al método de Romano [2002] que puede ser utilizado para evaluar cualquier tipo de cópula y el método gráfico de Frees y Valdez [1998] para cópulas de Arquímedes, Ghoudi et al. [1998] proponen una prueba para identificar si la cópula bivariada asociada a los datos pertenece a la familia de cópulas de valor extremo. Para los retornos $(R_{1i}, R_{2i})_{i=1, \dots, n}$ esta prueba se basa en el estadístico S_n :

$$\begin{aligned}
 S_n = & \frac{8}{n(n-1)} \sum_{i=1}^n \sum_{j=1, j \neq i}^n I_{\{R_{1i} \geq R_{1j}\}} I_{\{R_{2i} \geq R_{2j}\}} \\
 & - \frac{9}{n(n-1)(n-2)} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \sum_{k=1, k \neq j}^n I_{\{R_{1i} \geq R_{1j}\}} I_{\{R_{2i} \geq R_{2j}\}} I_{\{R_{1i} \geq R_{1k}\}} I_{\{R_{2i} \geq R_{2k}\}} - 1
 \end{aligned}
 \tag{4.17}$$

donde $I_{\{\Lambda\}}$ representa una función que toma el valor de uno si la condición Λ se satisface y cero en caso contrario. Bajo la hipótesis nula el estadístico $\widehat{S}_n / (\widehat{V}_{Jack}(\widehat{S}_n))^{1/2}$ sigue asintóticamente una distribución normal estándar, donde $\widehat{V}_{Jack}(\widehat{S}_n)$ representa el estimador de la varianza de \widehat{S}_n a través del método de *jackknife*. Al evaluar esta estadística sobre los datos simulados usados en el ejemplo del Gráfico 4.2, el valor del estadístico de prueba fue 1,53 (*p-value* 0,13). Este resultado no es sorprendente, ya que la cópula Gumbel pertenece a la familia Arquimax. Otra aplicación de esta prueba se encuentra en Cebrian et al. [2002].

Una ventaja que presenta el uso de este estadístico es que ofrece un criterio de selección más objetivo que los métodos gráficos, no obstante, su implementación únicamente indica si los datos provienen de una cópula de valor extremo, lo que incluye también a las Arquimax. Por lo tanto, es recomendable usar este criterio en conjunto con los métodos expuestos anteriormente.

4.4. Estimación de cópulas en presencia de dependencia temporal. Tal como se mencionó en la Sección 4.2 uno de los supuestos utilizados en la estimación de los parámetros de la cópula es que los datos provienen de una muestra *i.i.d.* En particular, este supuesto implica que las observaciones no deben presentar ningún tipo de dependencia temporal, lo que resulta poco probable en datos financieros [Mikosch, 2006]. Como resultado, se han desarrollado metodologías denominadas “cópulas dinámicas” que tienen en cuenta esta característica de las series financieras.

En el caso más general, las cópulas dinámicas utilizan algún tipo de regla de comportamiento para describir la evolución de las funciones de distribución (marginales y conjunta) condicionadas al conjunto de información disponible hasta el momento t , denominado \mathfrak{S}_t . Así, para este tipo de modelos la relación establecida por el teorema de Sklar en la ecuación (3.3), condicionada a \mathfrak{S}_t , es de la forma [Patton, 2006]:

$$F_t(r_{1t}, \dots, r_{dt} | \mathfrak{S}_{t-1}) = C_t(F_{1t}(r_{1t} | \mathfrak{S}_{t-1}), \dots, F_{dt}(r_{dt} | \mathfrak{S}_{t-1}) | \mathfrak{S}_{t-1}) \quad (4.18)$$

Patton [2001] muestra que si las distribuciones marginales y multivariada están condicionadas sobre el mismo conjunto de información, la ecuación (4.18) puede ser entendida como la versión condicionada del teorema de Sklar. A partir de (4.18) es posible modelar múltiples tipos de dependencia, usando especificaciones particulares de $C_t = C_t(\cdot | \mathfrak{S}_{t-1})$ y $F_{it} = F_{it}(r_{it} | \mathfrak{S}_{t-1})$.

En el caso más sencillo, los retornos y la cópula no dependen del tiempo, es decir, $C_t = C$ y $F_{it} = F$ para todo t . En consecuencia, la relación (4.18) toma su forma no condicionada (ecuación (3.3)), por lo que es posible usar las técnicas de estimación referenciadas en secciones anteriores sin necesidad de realizar ningún tipo de modelación adicional. Aplicaciones de este tipo de modelos se encuentran en Joe [1997], Nelsen [2006] y McNeil et al. [2005], entre otros.

Otra posible especificación de (4.18) es aquella en que la forma funcional de la cópula y sus parámetros no dependen del tiempo ($C_t = C$), pero las distribuciones marginales describen algún tipo de dependencia, por ejemplo, dependencia en el primer y segundo momento de las series del tipo $AR - GARCH$. El uso de este tipo de cópulas dinámicas es común en aplicaciones financieras, por ejemplo Poon et al. [2001], Chollete y Heinen [2006] y los modelos SCOMDY de Chen y Fan [2005], los cuales serán discutidos en detalle en la Sección 4.4.1.

Una tercera aproximación es aquella que modela dependencias en las distribuciones marginales y en la cópula. Adicional a una dependencia marginal del tipo $AR - GARCH$, Patton [2001, 2006] menciona dos posibles formas para las variaciones temporales de la cópula: en la primera de ellas, la forma funcional de la cópula se mantiene, pero se permite que el conjunto de parámetros de la cópula evolucione en el tiempo, es decir $C_t = C^0(\theta_t)$ donde θ_t representa el conjunto de parámetros de la cópula C^0 .

En la segunda especificación mencionada por Patton [2006], la cual fue usada también en Rodríguez [2007], tanto la forma funcional de la cópula como su conjunto de parámetros cambian en el tiempo $C_t = C_t^i(\theta_t^i)$ de acuerdo con diferentes regímenes $i = 1, \dots, k$. Aunque esta forma funcional resulta ser la más general de las presentadas anteriormente, también es la más costosa en términos de selección del modelo e implementación. A continuación se presentan las metodologías usadas por Chen y Fan [2005] y la primera metodología usada por Patton [2006].

4.4.1. Modelos dinámicos multivariados basados en la cópula (SCOMDY). Estos modelos fueron propuestos originalmente por Chen y Fan [2005], quienes encuentran las propiedades asintóticas de los estimadores utilizados en este tipo de modelos²⁴. Los modelos SCOMDY suponen que la cópula no depende del tiempo, mientras que el vector de retornos sigue un proceso de la forma:

$$\mathbf{r}_{it} = \mu_t(\theta_{o1}) + \sqrt{H_t(\theta_o)}\varepsilon_t \quad (4.19)$$

donde $\mu_t(\theta_{o1}) = [\mu_{1t}(\theta_{o1}), \dots, \mu_{dt}(\theta_{o1})]'$ representa el valor esperado del vector de retornos \mathbf{r}_t condicionado al conjunto de información \mathfrak{S}_{t-1} , mientras $H_t(\theta_o)$ describe el segundo momento condicionado del proceso y esta dado por:

$$H_t(\theta_o) = \text{diag}(h_{1t}(\theta_o), \dots, h_{dt}(\theta_o)) \quad (4.20)$$

donde $h_{it}(\theta_o) = h_{it}(\theta_{o1}, \theta_{o2}) = E[(r_{it} - \mu_{it}(\theta_{o1}))^2 | \mathfrak{S}_{t-1}]$ para $i = 1, \dots, d$ ²⁵. En esta especificación, el vector de parámetros es $\theta_o = (\theta'_{o1}, \theta'_{o2})'$, donde θ'_{o1} y θ'_{o2} no tienen elementos en común. En la ecuación (4.19) $\varepsilon_t = [\varepsilon_{1t}, \dots, \varepsilon_{dt}]'$ representa un vector de

²⁴No obstante, este tipo de modelos habían sido utilizados en trabajos anteriores como Poon et al. [2001] y Patton [2001]. Sin embargo, ninguno de estos trabajos desarrolla las propiedades de los estimadores de estos modelos.

²⁵Aunque la especificación original de los modelos SCOMDY utiliza la ecuación (4.20) para especificar la varianza condicionada del proceso, otros autores como Patton [2006] y Fernández [2003] utilizan matrices de varianzas y covarianzas no diagonales.

innovaciones independiente de \mathfrak{S}_{t-1} , idénticamente distribuido con media cero y varianza uno para cada uno de sus componentes, el cual sigue una distribución conjunta $F(\varepsilon_t) = C(F_1(\varepsilon_{1t}), \dots, F_d(\varepsilon_{dt}))$ donde F_i representa la función de distribución marginal de ε_{it} , $i = 1, \dots, d$ y C es una cópula.

Existen gran variedad de modelos SCOMDY, dependiendo de las formas funcionales que se especifiquen para las ecuaciones (4.19) y (4.20). Un ejemplo de este tipo de modelos es el $AR(p) - GARCH(1, 1)$ con cópula normal, el cual es de la forma:

$$\begin{aligned} r_{it} &= \zeta_i + \sum_{k=1}^p \delta_{ik} r_{i,t-k} + \nu_{it} \\ \nu_{it} &= \varepsilon_{it} \sqrt{h_{it}} \\ h_{it} &= \omega_i + \alpha_i \nu_{i,t-1}^2 + \beta_i h_{i,t-1} \end{aligned} \quad (4.21)$$

donde $\omega_i > 0$, α_i, β_i son mayores o iguales que cero y $\alpha_i + \beta_i < 1$ para todo $i = 1, \dots, d$. En este caso, la cópula asociada al vector de los errores estandarizados del modelo (ε_{it}) es la cópula normal referenciada en la Sección 3.5.1.

En general, las especificaciones $AR(p) - GARCH(1, 1)$ tienen una gran aceptación y uso en aplicaciones financieras, aunque el proceso de selección de la forma funcional de la cópula depende del comportamiento multivariado de ε_t . Por ejemplo, cuando se observa una mayor dependencia en los valores extremos de ε_t , es recomendable contrastar varios modelos SCOMDY con cópulas diferentes a la normal, y seleccionar la que muestre un mejor ajuste. Para la selección del modelo SCOMDY más apropiado, Chen y Fan (2005) sugieren una prueba de pseudo razón de verosimilitud (PLR por sus siglas en inglés), aunque también podrían ser usadas las técnicas presentadas en la Sección 4.3.

De la misma forma que en la estimación de los parámetros de la cópula, existen diferentes procedimientos para la estimación de los modelos SCOMDY. Chen y Fan [2005] aplican en forma univariada una técnica de dos etapas usual en este tipo de modelos, la cual se describe a continuación:

- i) En una primera etapa, se estima el modelo que describe el comportamiento del primer y segundo momento condicional de \mathbf{r}_t , por ejemplo un modelo $AR(p) - GARCH(1, 1)$ para cada uno de los elementos de \mathbf{r}_t [Chen y Fan, 2005], o modelos multivariados, tales como el $VAR - DVEC$ [Fernández, 2003] o el $VAR - BEKK$ [Patton, 2001].

Adicionalmente, si se tienen en cuenta las propiedades asintóticas de los estimadores de los modelos $GARCH$, es posible estimar separadamente los parámetros del primer y segundo momento condicional de \mathbf{r}_t . Otro punto importante para tener en cuenta es que en esta etapa la forma funcional de la cópula es desconocida, por lo tanto la mayoría de aplicaciones utilizan estimadores de quasi máxima verosimilitud²⁶.

²⁶Para una discusión más detallada de estimadores de quasi máxima verosimilitud y sus propiedades, véase Gouriéroux y Jasiak [2001].

- ii) En la segunda etapa se estiman los parámetros de la cópula sobre los residuales de la primera etapa con alguna de las técnicas presentadas en la Sección 4.1. Adicionalmente, es posible acompañar esta segunda etapa con los criterios de selección de la cópula presentados en la Sección 4.3.

4.4.2. *Cópulas con parámetros cambiantes en el tiempo.* Aunque los modelos planteados en la sección anterior son ampliamente utilizados y han tenido un buen desempeño, estos modelos no toman en cuenta algunas consideraciones importantes, como por ejemplo cuando los parámetros de la cópula siguen algún tipo de patrón paramétrico en el tiempo (parámetros dinámicos), o cuando existe un cambio estructural. En consecuencia, se han desarrollado otros tipos de modelos que generalizan las metodologías presentadas en la sección anterior, los cuales se presentan a continuación.

Patton [2006] propone una dinámica similar a la de un modelo *ARMA* sobre los parámetros de la cópula, para la modelación de los retornos diarios de las tasas de cambio del marco alemán y el yen japonés con respecto al dólar, durante el período comprendido entre enero de 1991 y diciembre de 2001. Esta dinámica fue aplicada sobre los parámetros de dos cópulas, la cópula Gaussiana y la cópula Joe-Clayton²⁷, de la siguiente manera:

$$\lambda_{u,t} = \Lambda \left(\omega_u + \beta_u \lambda_{u,t-1} + \frac{\alpha_u}{10} \sum_{j=1}^{10} |u_{t-j} - v_{t-j}| \right) \quad (4.22)$$

$$\lambda_{l,t} = \Lambda \left(\omega_l + \beta_l \lambda_{l,t-1} + \frac{\alpha_l}{10} \sum_{j=1}^{10} |u_{t-j} - v_{t-j}| \right) \quad (4.23)$$

$$\rho_t = \tilde{\Lambda} \left(\omega_\rho + \beta_\rho \rho_{t-1} + \frac{\alpha_\rho}{10} \sum_{j=1}^{10} \Phi^{-1}(u_{t-j}) \Phi^{-1}(v_{t-j}) \right) \quad (4.24)$$

donde $u_t = F_1(r_{1t})$, $v_t = F_2(r_{2t})$ y Φ^{-1} es la función inversa de distribución normal, $\lambda_{u,t}$, $\lambda_{l,t}$ y ρ_t son los coeficientes de dependencia en las colas superior e inferior (para la cópula Joe-Clayton), y el coeficiente de correlación (para la cópula Gaussiana), respectivamente. Además, $\Lambda(x) \equiv (1 + e^{-x})^{-1}$ y $\tilde{\Lambda}(x) \equiv \tanh(x/2)$ son las transformaciones logística y logística modificada, que son usadas para garantizar que $\lambda_{u,t}$, $\lambda_{l,t}$ y ρ_t se mantengan entre (0, 1) y (-1, 1), respectivamente.

Para este caso, la estimación es similar a la presentada en la sección anterior, con la diferencia que en estos modelos se utiliza una cópula que tenga en cuenta algún tipo de dinámica en sus parámetros, por ejemplo la descrita por las ecuaciones (4.22) y (4.23) para la cópula Joe-Clayton. Patton [2006] usa en su aplicación técnicas de máxima verosimilitud.

²⁷La cópula Joe-Clayton corresponde a una generalización de la cópula *BB7* presentada en la Tabla C.3 con $\delta = \frac{1}{\log_2(2-\lambda_u)}$, $\theta = \frac{-1}{\log_2(\lambda_l)}$ y $\lambda_u \in (0, 1)$, $\lambda_l \in (0, 1)$.

Una ventaja que ofrecen este tipo de modelos es que permiten observar cómo evolucionan los parámetros de la cópula a través del tiempo, lo que podría brindar información adicional acerca de posibles cambios en las relaciones de dependencia. Por ejemplo, Patton [2006] encuentra un cambio en los patrones de dependencia entre las tasas de cambio del marco alemán y el yen japonés con respecto al dólar, a partir de la introducción del euro y la fijación de una tasa de cambio fija entre esta moneda y el marco alemán en enero de 1999²⁸.

De otro lado, Dias y Embrechts [2004] utilizan cópulas con parámetros dinámicos para tener en cuenta posibles cambios estructurales en las relaciones de dependencia entre el marco alemán y el yen japonés, desde abril de 1986 hasta octubre de 1998, con seis diferentes frecuencias. En una primera etapa, después de filtrar la dependencia temporal a partir de modelos *ARMA – GARCH*, Dias y Embrechts [2004] identifican a la cópula t como aquella que describe mejor la dependencia entre estas dos tasas de cambio. Posteriormente, estiman por máxima verosimilitud una cópula t en la que sus parámetros (\tilde{v}_t y ρ_t) evolucionan en el tiempo de la siguiente forma:

$$\tilde{v}_t = \tilde{v} \text{ para todo } t \quad (4.25)$$

$$\rho_t = h^{-1}(\omega_0 + \alpha_1 h(\rho_{t-1}) + \beta_1 u_{t-1} v_{t-1}) \quad (4.26)$$

donde u_t y v_t son las funciones de distribución evaluadas en los retornos de cada moneda y $h(\rho) = \log\left(\frac{1+\rho}{1-\rho}\right)$ es la transformación de Fisher sobre ρ . De manera similar a (4.24), (4.26) muestra una evolución del tipo *ARMA* sobre el coeficiente de correlación. A partir de la evolución de ρ_t , Dias y Embrechts [2004] identifican varios puntos en los cuales se pudo haber presentado un cambio estructural en la relación de dependencia de estas dos series.

Adicionalmente, estos dos autores proponen una metodología basada en pruebas de razón de verosimilitud (LR) para detectar el posible cambio estructural en la relación de dependencia. En este caso, se realiza una estimación recursiva de los parámetros de la cópula, para posteriormente calcular el estadístico de prueba como el máximo de los estadísticos LR calculados recursivamente²⁹. Los valores críticos de esta prueba son obtenidos a partir de la aproximación asintótica realizada por Csörgö y Horváth [1997].

No obstante, es posible que exista más de un cambio estructural dentro de la relación de dependencia. Para tener en cuenta esta posibilidad, Dias y Embrechts [2004] utilizan la metodología de Vostrikova [1981]. En una primera instancia, se calcula el estadístico LR recursivo sobre toda la muestra y si se encuentra evidencia de cambio estructural, se divide la muestra en dos, antes y después del punto en que se presenta el cambio, y se vuelve a realizar la prueba LR recursiva sobre las dos submuestras. Este procedimiento se

²⁸La muestra utilizada por Patton [2006] incluye el período de transición del marco alemán al euro (primero de enero de 1999 hasta el 31 de diciembre de 2001). Durante este período, se estableció una tasa de cambio fija de 1,95583 marcos por cada euro.

²⁹Bajo la hipótesis nula de esta prueba los parámetros de la cópula son constantes en el tiempo mientras que en la alterna algunos parámetros cambian.

repite sucesivamente sobre las nuevas submuestras, hasta no tener evidencia de cambio estructural.

Al aplicar este procedimiento, Dias y Embrechts [2004] encontraron evidencia de cuatro cambios estructurales dentro de la muestra analizada, y estimaron los parámetros de la cópula asociados a cada uno de los períodos obtenidos. Este tipo de estimación se presenta como una alternativa a la estimación de cópulas con parámetros dinámicos, ya que en este caso se estiman cópulas sobre subperíodos “estables”. No obstante, Dias y Embrechts [2004] señalan la baja potencia que tiene la prueba LR recursiva para detectar cambios estructurales pequeños, por lo que en estos casos sería mejor utilizar cópulas con parámetros dinámicos.

4.5. Modelación de valores extremos a través de cópulas. En secciones anteriores se han descrito diversas metodologías para tratar de caracterizar el comportamiento multivariado de los datos a través de la cópula. Sin embargo, en algunas aplicaciones (e.g. finanzas), es posible que se desee analizar únicamente ciertas características de la distribución conjunta, tales como las relaciones de dependencia entre los eventos extremos de las series.

Por ejemplo, en el Gráfico 3.1 se exhiben cuatro diferentes estructuras de dependencia, que a su vez muestran varios tipos de relaciones entre los retornos extremos. Si estas estructuras describen la relación entre los retornos de dos activos financieros, aquellas que presenten una mayor dependencia en la cola izquierda resultarán más riesgosas, ya que a medida que se observen pérdidas muy grandes en un activo, es probable que también suceda lo mismo en el otro.

En una primera etapa, estas relaciones de dependencia entre observaciones extremas, o relaciones de “dependencia extrema”, pueden ser aproximadas a partir de medidas específicas que se desprenden de la cópula, tales como el coeficiente de dependencia en las colas, descrito en la Sección 3.3.2. No obstante, el coeficiente de dependencia en las colas es una medida asintótica, que no ofrece mayor información acerca del comportamiento probabilístico de las observaciones extremas.

De la misma forma que en el caso univariado, existen distribuciones límite que describen el comportamiento de la distribución conjunta de los extremos de las series de estudio. Estas distribuciones a su vez tienen una cópula asociada, la cual pertenece a la familia de valor extremo. A continuación se describen los principales resultados relacionados con el comportamiento de los extremos de las series de estudio, para el caso univariado y multivariado.

4.5.1. Teoría del valor extremo: caso univariado. Análogo al teorema central del límite, que estudia el comportamiento de la suma de variables aleatorias independientes e idénticamente distribuidas, la Teoría del Valor Extremo, *EVT* por sus siglas en inglés, se concentra en resultados asintóticos acerca de cuál es la distribución límite hacia la cual converge la distribución del valor máximo de una variable aleatoria. El teorema de Fisher y Tippett

[1928] y Gnedenko [1943], indica que el máximo de una variable aleatoria puede ser estabilizado a través de secuencias de parámetros de localización y escala, de tal manera que converja en distribución a una distribución GEV ³⁰, $H_\xi(x)$, la cual viene de la forma:

$$H_\xi(x) = \begin{cases} \exp\left(- (1 + \xi x)^{-1/\xi}\right) & \text{si } \xi \neq 0 \\ \exp(-e^{-x}) & \text{si } \xi = 0 \end{cases} \quad (4.27)$$

En este sentido, es importante anotar que la estimación de los parámetros de la GEV requiere de la obtención de una serie de valores extremos a partir de los datos observados. Una metodología fácil de implementar es la de máximos por bloques, en la cual se identifican los valores extremos como los eventos que tengan el valor más alto durante un período de tiempo específico.

Otra forma de analizar el comportamiento de los valores extremos es a partir de la cola de la distribución de la variable aleatoria. En esta aproximación el teorema de Balkema y de Haan [1974] y Pickands [1975], indica que para aquellas distribuciones en las que sus máximos muestrales convergen en distribución a la distribución GEV , la distribución de excesos sobre un umbral converge a una distribución GPD ³¹, $G_{\xi,\beta}(x)$, la cual es de la forma:

$$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi x/\beta)^{-1/\xi} & \text{si } \xi \neq 0 \\ 1 - \exp(-x/\beta) & \text{si } \xi = 0 \end{cases} \quad (4.28)$$

Dado que en este caso se desea aproximar el comportamiento de la cola de la distribución a través de una GPD , la forma de seleccionar los valores extremos es a través del método de picos sobre el umbral (POT por sus siglas en inglés). Como su nombre lo indica, en este método los valores extremos se definen como los valores de la serie que sean mayores que un umbral predeterminado v . Una discusión más detallada acerca de la EVT , su implementación y aplicaciones a la medición de riesgo de mercado se encuentra en Melo y Becerra [2006].

4.5.2. *Teoría del valor extremo: caso multivariado.* De la misma forma que en el caso univariado, la teoría del valor extremo multivariada, $MEVT$ por sus siglas en inglés, muestra el comportamiento asintótico de la distribución del máximo \mathbf{M}_n de una familia de vectores aleatorios $\mathbf{X}_1, \dots, \mathbf{X}_n$ donde $\mathbf{X}_t \in \mathbb{R}^d$ con $t = 1, \dots, n$. En este caso, \mathbf{M}_n corresponderá al vector de máximos de los d componentes de $\mathbf{X}_1, \dots, \mathbf{X}_n$, es decir, $\mathbf{M}_n = [M_{1n}, \dots, M_{dn}]'$ donde $M_{jn} = \max(X_{j1}, \dots, X_{jn})$ $j = 1, \dots, d$. Vale la pena destacar que en la mayoría de los casos \mathbf{M}_n no coincidirá con una observación de \mathbf{X}_t .

En la $MEVT$ se busca la distribución límite de los máximos normalizados a través de vectores de parámetros de localización ($\mathbf{b}_n \in \mathbb{R}^d$) y escala ($\mathbf{c}_n \in \mathbb{R}^d$, $\mathbf{c}_n > \mathbf{0}$), tal que:

$$\lim_{n \rightarrow \infty} P\left(\frac{\mathbf{M}_n - \mathbf{b}_n}{\mathbf{c}_n} \leq \mathbf{x}\right) = \lim_{n \rightarrow \infty} F^n(\mathbf{c}_n \mathbf{x} + \mathbf{b}_n) = H(\mathbf{x}) \quad (4.29)$$

³⁰Distribución generalizada de valor extremo, GEV por sus siglas en inglés.

³¹Distribución generalizada de Pareto, GPD por sus siglas en inglés.

si la relación (4.29) se mantiene, entonces se dice que la función de distribución F pertenece al dominio de atracción en los máximos de H , donde H corresponderá a una función de distribución multivariada de valor extremo, *MEV* por sus siglas en inglés³².

Vale la pena mencionar que a diferencia del caso univariado, en el que la distribución del máximo de la variable aleatoria converge a una sola función de distribución, en el análisis multivariado este resultado no se mantiene. Por el contrario, en la *MEVT* la distribución límite es una familia de distribuciones multivariadas, la *MEV*, la cual puede ser aproximada por una familia de cópulas.

Si la relación (4.29) se mantiene, usando la definición del teorema de Sklar (Sección 3.2) se tiene que:

$$H(\mathbf{x}) = C_0\left(\tilde{F}_1(x_1), \dots, \tilde{F}_d(x_d)\right) \quad (4.30)$$

donde $\tilde{F}_1(x_1), \dots, \tilde{F}_d(x_d)$ son las funciones de distribución de los máximos de cada uno de los componentes de la familia de vectores aleatorios $\mathbf{X}_1, \dots, \mathbf{X}_n$, por lo que, de acuerdo con el teorema de Fisher y Tippett [1928] y Gnedenko [1943] (Sección 4.5.1), cada una de éstas converge en distribución a una distribución *GEV*. A partir de estos resultados, es posible definir el dominio de atracción en los máximos en términos de la cópula, tal como se muestra en el Teorema 4.1 [McNeil et al., 2005]:

Teorema 4.1. *Sea $F(\mathbf{r}) = C(F_1(r_1), \dots, F_d(r_d))$, para funciones marginales continuas F_1, \dots, F_d y alguna cópula C . Adicionalmente, sea $H(\mathbf{r})$ una distribución *MEV* tal que satisface (4.30) para alguna cópula C_0 . Se dice que F pertenece al dominio de atracción en los máximos de H , si y solo si cada F_i pertenece al dominio de atracción en los máximos de la distribución *GEV*, con $i = 1 \dots, d$ y si además se satisface la siguiente condición:*

$$\lim_{t \rightarrow \infty} C^t(\mathbf{u}^{1/t}) = C_0(\mathbf{u}), \quad \mathbf{u} \in [0, 1]^d \quad (4.31)$$

Tal como lo destacan McNeil et al. [2005], el resultado obtenido en el teorema anterior implica que la cópula límite C_0 se encuentra únicamente determinada por la cópula C , en consecuencia, aunque las distribuciones marginales determinarán el comportamiento de la distribución *MEV*, su estructura de dependencia no se verá afectada por éstas.

En los casos en que la condición (4.31) se mantiene, se dice que la cópula C pertenece al dominio de atracción de los máximos de la cópula C_0 . La familia de cópulas límite que cumple esta condición se denomina de valor extremo, la cual satisface la siguiente propiedad:

$$C(\mathbf{u}^t) = C^t(\mathbf{u}), \quad \forall t > 0 \quad (4.32)$$

donde $\mathbf{u} = [u_1, \dots, u_d]'$ con $u_i = \tilde{F}_i(x_i)$. Algunas de las cópulas más importantes, tales como la de independencia y la de comonotonicidad, hacen parte de esta familia, junto

³²Para una discusión detallada sobre los dominios de atracción en los máximos, tanto para el caso univariado como multivariado, véase Demarta [2002].

con otras cópulas versátiles, por ejemplo, la Gumbel y la Galambos, las cuales resultarán de gran importancia para la aplicación de la *MEVT*, tal como se mostrará más adelante.

Adicionalmente, existe otra representación de esta familia de cópulas que se basa en el siguiente resultado obtenido por Pickands [1981].

Teorema 4.2 (Representación de Pickands). *se dice que la cópula C es una cópula de valor extremo si y sólo si ésta tiene la siguiente representación:*

$$C(\mathbf{u}) = \exp \left\{ B(\omega_1, \dots, \omega_d) \left(\sum_{i=1}^d \tilde{u}_i \right) \right\} \quad (4.33)$$

donde $\tilde{u}_i = \log u_i$ y $\omega_i = \tilde{u}_i / \sum_{j=1}^d \tilde{u}_j$ para $i = 1, \dots, d$ y B es una función convexa, la cual esta definida como:

$$B(\omega_1, \dots, \omega_d) = \int_{S_d} \text{máx}(x_1\omega_1, \dots, x_d\omega_d) dU(\mathbf{x}) \quad (4.34)$$

con U siendo una medida finita definida sobre el simplex unitario d -dimensional, es decir, sobre el conjunto $S_d = \{\mathbf{x}, x_i \geq 0, i = 1, \dots, d, \sum_i x_i = 1\}$ [McNeil et al., 2005].

La función B es denominada función de dependencia y generalmente toma una forma sencilla sólo en el caso bidimensional, cuando se puede expresar como una función A evaluada en la variable ω_1 . En tal caso, $A(\omega_1) = A(\omega) = B(\omega, 1 - \omega)$ con $0 \leq \omega \leq 1$. En este caso la representación (4.33) de la cópula de valor extremo viene de la forma:

$$C(u_1, u_2) = \exp \left\{ (\log u_1 + \log u_2) A \left(\frac{\log u_1}{\log u_1 + \log u_2} \right) \right\} \quad (4.35)$$

De la misma manera que en la función B , A es una función convexa, que satisface la siguiente condición:

$$\text{máx}(\omega, 1 - \omega) \leq A(\omega) \leq 1 \quad (4.36)$$

nuevamente, los límites de la función $A(\omega)$ mostrados en la condición (4.36) se encuentran asociados a las cópulas de independencia ($A(\omega) = 1$ para todo $\omega \in [0, 1]$) y de comonotonicidad ($A(\omega) = \text{máx}(\omega, 1 - \omega)$). Una de las utilidades de la representación (4.35) radica en el hecho que esta expresión brinda una forma relativamente sencilla de encontrar la función $A(\omega)$ a través de la siguiente definición:

$$A(\omega) = -\log C \left(e^{-\omega}, e^{-(1-\omega)} \right) \quad \omega \in [0, 1] \quad (4.37)$$

Bouyé et al. [2000] muestran algunas de las funciones de dependencia asociadas con varias de las cópulas bivariadas de valor extremo más utilizadas, las cuales se reproducen en la Tabla 4.2

En la Tabla 4.2, $v(u_1, u_2; \theta) = \Phi \left(\frac{1}{\theta} + \frac{1}{2}\theta \log \frac{\log u_1}{\log u_2} \right)$, $\xi(\omega; \theta) = \Phi \left(\frac{1}{\theta} + \frac{1}{2}\theta \log \frac{\omega}{(1-\omega)} \right)$ y $\tilde{u} = -\log u$.

Familia	θ	$C(u_1, u_2)$	$A(\omega)$
Cópula de Independencia		$u_1 u_2$	1
Gumbel	$[1, \infty)$	$\exp \left[- \left((\tilde{u}_1)^\theta + (\tilde{u}_2)^\theta \right)^{\frac{1}{\theta}} \right]$	$(\omega^\theta + (1 - \omega)^\theta)^{\frac{1}{\theta}}$
Gumbel II	$[0, 1]$	$u_1 u_2 \exp [\theta \tilde{u}_1 \tilde{u}_2 / (\tilde{u}_1 + \tilde{u}_2)]$	$\theta \omega^2 - \theta \omega + 1$
Galambos	$[0, \infty)$	$u_1 u_2 \exp \left[(\tilde{u}_1^{-\theta} + \tilde{u}_2^{-\theta})^{-\frac{1}{\theta}} \right]$	$1 - (\omega^{-\theta} + (1 - \omega)^{-\theta})^{-\frac{1}{\theta}}$
Husler-Reiss	$[0, \infty)$	$\exp [-\tilde{u}_1 v(u_1, u_2; \theta) - \tilde{u}_2 v(u_2, u_1; \theta)]$	$\omega \xi(\omega; \theta) + (1 - \omega) \xi(1 - \omega; \theta)$
Marshall-Olkin	$[0, 1]^2$	$u_1^{1-\theta_1} u_2^{1-\theta_2} \min(u_1^{\theta_1}, u_2^{\theta_2})$	$\max(1 - \theta_1 \omega, 1 - \theta_2(1 - \omega))$
Cópula de Comonotonidad		$\min(u_1, u_2)$	$\max(\omega, 1 - \omega)$

TABLA 4.2. Cópulas paramétricas de valor extremo y sus funciones de dependencia

4.5.3. *Implementación MEVT.* Al igual que en el caso univariado, la implementación de la MEVT requiere la obtención de una serie de máximos. Una alternativa es el uso del método de máximos por bloques, el cual identifica los valores extremos como los eventos que tengan el valor más alto durante un período de tiempo específico. En este procedimiento, los datos son divididos en m bloques no traslapados de longitud n y se define el vector \mathbf{M}_{nk} como el vector de d máximos (componente por componente) dentro del bloque k , para $k = 1, \dots, m$. Así, se obtiene una serie de m vectores de valores extremos $\mathbf{M}_{n1}, \dots, \mathbf{M}_{nm}$.

En general, la estimación de los parámetros de la cópula de valor extremo es realizada a través del uso de máxima verosimilitud. De la misma forma que en la Sección 4.1.3, se estiman dos conjuntos de parámetros, el vector θ de parámetros de la cópula y los vectores ξ , μ y σ , que agrupan los parámetros de forma, localización y escala de las distribuciones marginales GEV. Este proceso puede implementarse en una o en dos etapas, tal como se explicó en la Sección 4.1.3.

Otra forma de encontrar valores extremos es a partir del método de picos sobre umbrales (POT). Dados unos vectores aleatorios R_1, \dots, R_n , con una función de distribución conjunta $F(\mathbf{r}) = C(F_1(r_1), \dots, F_d(r_d))$, una cópula C y unas funciones de distribución marginales F_1, \dots, F_d , la metodología POT multivariada (MPOT) afirma que si F pertenece al dominio de atracción de una distribución MEV, es posible aproximarse a la cola superior de $F(\mathbf{r})$ para valores de las variables que sean mayores que un vector de umbrales $v = [v_1, \dots, v_d]'$, en cuyo caso las distribuciones marginales serán distribuciones

GPD, ecuación (4.28), las cuales toman la forma [McNeil, 1999]:

$$\tilde{F}_j(r_j) = 1 - \lambda_j \left(1 + \xi_j \frac{r_j - v_j}{\beta_j} \right)^{-1/\xi_j} \quad j = 1, \dots, d \quad (4.38)$$

donde β_j y ξ_j son los parámetros de la distribución *GPD* asociada a los excesos sobre el umbral de r_j y λ_j corresponde a la probabilidad de que las observaciones de r_j superen el umbral. Esto sugiere que para $\mathbf{r} \geq \mathbf{v}$ se puede usar la aproximación $\tilde{F}(\mathbf{r}) \approx C_0 \left(\tilde{F}_1(r_1), \dots, \tilde{F}_d(r_d) \right)$, donde C_0 representa la cópula límite descrita en el Teorema 4.1.

Para la implementación práctica de esta metodología, es necesario obtener en una etapa inicial un vector de “excesos” de retorno, a partir de la estimación del conjunto de umbrales $\mathbf{v} = [v_1, \dots, v_d]'$. En general, en la identificación de estos umbrales es posible utilizar metodologías univariadas sobre cada una de las series, por ejemplo, las metodologías (paramétricas o no paramétricas) basadas en el estimador de Hill o los gráficos del “promedio de los excesos”, los cuales brindan información acerca de la posible ubicación del umbral³³. No obstante, estas metodologías pueden resultar poco manejables en dimensiones grandes, por lo que también es posible utilizar un umbral fijo para todas las series, por ejemplo, el percentil 95 de la distribución. Aunque esta última alternativa genera ganancias en términos de tiempo, puede llevar a sesgos en la estimación.

Después de haber obtenido el vector de umbrales, el siguiente paso es la estimación de los parámetros de la cópula (θ) y de las distribuciones marginales (μ y β) a través de máxima verosimilitud, lo cual nuevamente, se puede hacer en una o dos etapas. En este caso, la función de verosimilitud tiene en cuenta datos censurados y, para dos dimensiones, se construye de acuerdo con el siguiente esquema [Longin y Solnik, 2001]:

- i) Se separan los retornos en cuatro conjuntos, dependiendo si exceden o no los umbrales determinados, v_1 y v_2 . Es decir, se definen los conjuntos A_{jk} , donde $j = I_{\{R_1 > v_1\}}$, $k = I_{\{R_2 > v_2\}}$ e $I_{\{\cdot\}}$ representa la función indicadora. Por ejemplo, en el Gráfico 4.3 estos conjuntos corresponden a las zonas A_{00} , A_{01} , A_{10} y A_{11} , donde A_{00} contiene a las parejas de retornos (r_{1t}, r_{2t}) en las que ninguno de los dos componentes superan los umbrales.
- ii) Se define la función de verosimilitud para cada una de las cuatro regiones. Así, el aporte a la verosimilitud total de la pareja (r_{1t}, r_{2t}) dependerá de en cual conjunto se ubique, por lo tanto, existirán cuatro tipos de contribuciones a la verosimilitud total,

³³Para una discusión detallada de estos métodos, véanse por ejemplo Embrechts et al. [1997] y McNeil et al. [2005] entre otros.

L_{00} , L_{01} , L_{10} y L_{11} las cuales se definen como sigue:

$$\begin{aligned}
 L_{00}(r_{1t}, r_{2t}) &= \tilde{F}(v_1, v_2) \\
 L_{01}(r_{1t}, r_{2t}) &= \left. \frac{\partial \tilde{F}(r_{1t}, r_{2t})}{\partial r_{2t}} \right|_{R_{1t}=v_1} \\
 L_{10}(r_{1t}, r_{2t}) &= \left. \frac{\partial \tilde{F}(r_{1t}, r_{2t})}{\partial r_{1t}} \right|_{R_{2t}=v_2} \\
 L_{11}(r_{1t}, r_{2t}) &= \frac{\partial^2 \tilde{F}(r_{1t}, r_{2t})}{\partial r_{1t} \partial r_{2t}}
 \end{aligned} \tag{4.39}$$

donde $\tilde{F}(r_{1t}, r_{2t}) = C_0(\tilde{F}_1(r_{1t}), \tilde{F}_2(r_{2t}))$ y \tilde{F}_j es una función de distribución GPD tal como se definió en la ecuación (4.38).

iii) Finalmente, se define la función de verosimilitud para una secuencia de vectores aleatorios $\mathbf{r} = (r_{1t}, r_{2t})_{t=1, \dots, n}$ como la multiplicatoria de las contribuciones a la verosimilitud de cada una de las observaciones, de la siguiente forma:

$$L(\mathbf{r}; \theta, \xi, \beta) = \prod_{t=1}^n \sum_{j,k=0}^1 L_{jk}(r_{1t}, r_{2t}) I_{jk}(r_{1t}, r_{2t}) \tag{4.40}$$

donde $I_{jk}(r_{1t}, r_{2t}) = I_{\{(r_{1t}, r_{2t}) \in A_{jk}\}}$.

Vale la pena recordar que esta metodología supone que los retornos son independientes e idénticamente distribuidos. En caso que no se presente esta situación, se hace necesario el modelamiento de las dependencias existentes en las series, tal cómo se describe en la Sección 4.4.

Longin y Solnik [2001] aplican la metodología de MPOT a los retornos mensuales de índices accionarios de Inglaterra, Francia, Alemania y Japón entre 1959 y 1996 (456 observaciones), con el fin de comprobar si existe dependencia entre los retornos extremos de estas series y los retornos extremos de un índice del mercado accionario de Estados Unidos. Adicionalmente, este ejercicio tiene en cuenta si el mercado se encuentra a la baja (*bear market*) o al alza (*bull market*). En este documento, los autores usan como cópula límite a la Gumbel, $C(u_1, u_2) = \exp \left[- \left((-\log u_1)^\theta + (-\log u_2)^\theta \right)^{1/\theta} \right]$, ya que su parámetro $\theta \in [1, \infty)$ describe diferentes tipos de dependencia, desde independencia (en el caso en que $\theta \rightarrow 1$) hasta comonotonicidad (cuando $\theta \rightarrow \infty$).

Los resultados de la estimación para la cola izquierda de la distribución conjunta de los retornos de los índices alemán y estadounidense se encuentran reportados en la Tabla 4.3. Con base en estos datos se pueden realizar algunos cálculos sobre probabilidades de ocurrencia de eventos extremos. Por ejemplo, la probabilidad estimada de que los dos

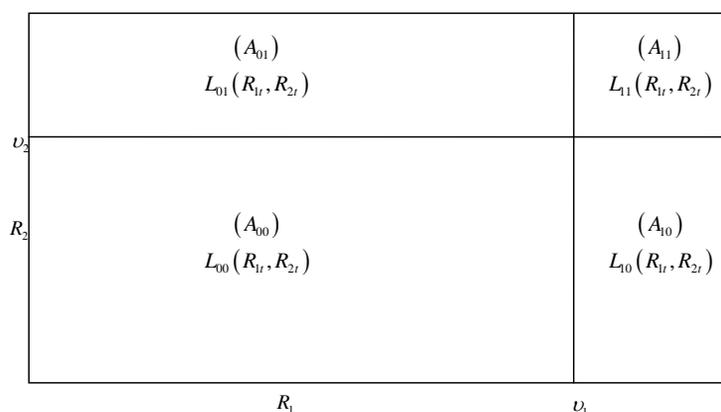


GRÁFICO 4.3. Regiones usadas para la definición de la función de verosimilitud en la metodología MPOT bivariada.

Parámetro	Estados Unidos	Alemania
v	-6,12 %	-7,84 %
$obs > v$	25	16
ξ	0,153 (0,187)	0,260 (0,674)
β	2,367 (0,652)	3,102 (1,910)
λ	0,060 (0,015)	0,043 (0,009)
θ	1,389 (0,166)	

TABLA 4.3. Resultados de la estimación de la relación de dependencia extrema para excesos negativos basados en Longin y Solnik [2001] (errores estándar entre paréntesis)

índices accionarios caigan más del 10 % es:

$$\begin{aligned}
 p_{12} &= P(R_1 \geq 0,1, R_2 \geq 0,1) \\
 &= 1 - \tilde{F}_1(0,1) - \tilde{F}_2(0,1) + C_0(\tilde{F}_1(0,1), \tilde{F}_2(0,1)) = 0,0063
 \end{aligned}$$

Adicionalmente, la probabilidad de que el índice accionario alemán caiga por debajo de 10 % dado que el índice de E.U. cae más de 10 % es:

$$p_{12}/p_1 = \frac{p_{12}}{p_1} = \frac{p_{12}}{1 - \tilde{F}_1(0,1)} = \frac{0,0063}{0,0139} = 0,4517$$

con lo cual se concluye que para el período de muestra analizado, aunque la probabilidad de observar caídas mayores a 10 % en ambos índices es reducida, cerca de la mitad de los meses en que se presente una pérdida superior a 10 % en el índice estadounidense,

ésta se encontrará acompañada por una reducción de proporciones similares en el índice alemán. No obstante, vale la pena destacar que Longin y Solnik [2001] suponen que los datos son *i.i.d.*, lo cual puede afectar la estimación y la interpretación de los resultados en caso de que este supuesto no sea apropiado.

5. APLICACIÓN DE LA CÓPULA EN MODELOS DE RIESGO.

5.1. Modelos de riesgo de mercado. Uno de los usos de los modelos expuestos anteriormente se encuentra en la valoración del riesgo de mercado de un portafolio. Específicamente, en este documento se considerarán dos medidas de riesgo ampliamente aplicadas en la literatura: el Valor en Riesgo (*VaR*, por sus siglas en inglés) y la pérdida esperada en la cola (*Expected Shortfall*, *ES*), cuyas definiciones se presentan a continuación:

- El valor en riesgo es una medida estándar para la medición de riesgo de mercado. El *VaR* corresponde a la máxima pérdida posible en el $\alpha \times 100\%$ de los mejores escenarios, o de forma equivalente, la mínima pérdida posible en el $(1 - \alpha) \times 100\%$ de las pérdidas más grandes en un período de tenencia dado (k).

En el Gráfico 5.1 se representa la distribución hipotética de las pérdidas y ganancias de un activo (o portafolio) después de mantenerse durante k días. El área sombreada corresponde al 95% de los mejores resultados (es decir, aquellos casos en que se obtiene una pérdida pequeña o una ganancia), por lo tanto, el *VaR* es el límite derecho de dicha área. En términos estadísticos esta medida se puede definir como el α -ésimo percentil de la distribución futura de las pérdidas y ganancias. Si se denota a esta distribución como F_R , entonces el *VaR* con un nivel de significancia α , VaR_α , estará definido por:

$$VaR_\alpha = \inf \{r | F_R(r) \geq \alpha\} \quad (5.1)$$

- Como su nombre lo indica, la pérdida esperada en la cola (ES_α) se define como el promedio de las pérdidas superiores al VaR_α . Al ser comparado con el valor en riesgo el ES_α ofrece ventajas adicionales, ya que si bien el VaR_α indica cuál es la menor de las peores pérdidas, no brinda información sobre la magnitud de los valores que lo superen, mientras que la definición del ES_α si le permite cumplir este requerimiento. Formalmente, el ES_α se define como:

$$ES_\alpha = E(R | R \geq VaR_\alpha) \quad (5.2)$$

Dependiendo de la distribución de las pérdidas y ganancias del activo, el VaR_α y el ES_α pueden ser fáciles de calcular. Sin embargo, la situación es un poco más complicada en presencia de un portafolio, ya que el VaR_α no es una medida coherente de riesgo en el sentido de Artzner et al. [1999], debido a que no siempre cumple la propiedad de subaditividad. Para que una medida de riesgo ρ sea subaditiva, ésta debe satisfacer la siguiente condición:

$$\rho^{(p)} \leq \sum_{j=1}^d \omega^{(j)} \rho^{(j)} \quad (5.3)$$

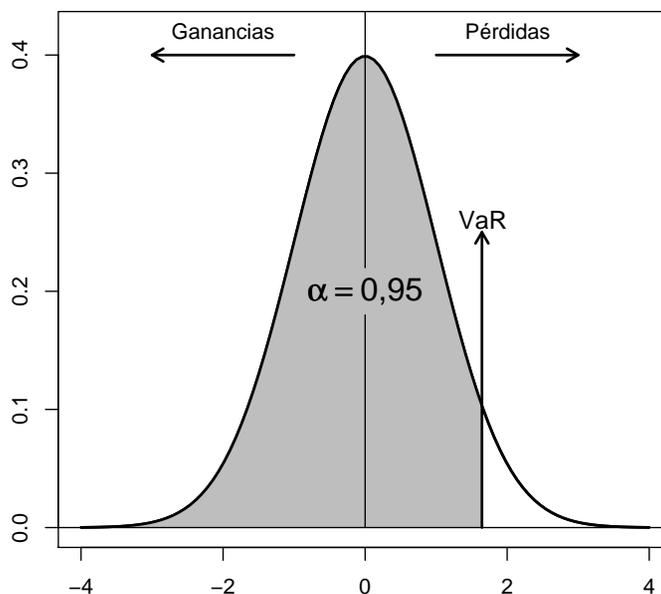


GRÁFICO 5.1. Valor en Riesgo.

donde $\rho^{(p)}, \rho^{(j)}$ $j = 1, \dots, d$, son las medidas de riesgo calculadas para el portafolio y cada uno de los activos que lo conforman y $\omega^{(j)}$ es la participación del activo j dentro de la cartera, tal que $\sum_{j=1}^d \omega^{(j)} = 1$.

La condición (5.3) tiene un significado especial dentro de la administración de riesgo, ya que se encuentra asociada con el principio de diversificación. Ante portafolios grandes, en los que resulta complicado calcular una medida de riesgo global, es útil conocer que el máximo riesgo que podría tener portafolio corresponde a la suma de los riesgos individuales. Además, esta propiedad está relacionada con la convexidad de la medida de riesgo, la cual garantiza que se puede encontrar una combinación óptima dentro del portafolio tal que el riesgo sea mínimo Acerbi y Tasche [2002]. Sin embargo, como lo anotan Embrechts et al. [1999a], el VaR_α si cumple la propiedad de subaditividad en el contexto de distribuciones elípticas³⁴.

5.1.1. Modelos de VaR_α y ES_α para un portafolio usando cópulas. En general, para calcular el VaR_α y el ES_α para portafolios se hace necesario conocer tanto las distribuciones marginales como la distribución multivariada de los factores de riesgo incluidos en él, lo que exhorta nuevamente al uso de la cópula para simplificar este trabajo, más aún si se tiene en cuenta que encontrar una fórmula explícita para determinar el VaR_α y el ES_α

³⁴Adicionalmente, cuando la dependencia entre los factores de riesgo es descrita por la cópula de comonotonidad (ecuación (3.23)), el VaR_α del portafolio es exactamente igual a la suma del VaR_α de cada uno de los activos [McNeil et al., 2005].

del portafolio puede convertirse en una tarea complicada (incluso hasta imposible) dependiendo de la función de distribución multivariada y del número de activos que hagan parte de la cartera³⁵.

Una manera relativamente sencilla y eficaz de determinar cuáles son el VaR_α y el ES_α de un portafolio es a través del uso del método de simulación de Monte Carlo, el cual se resume en los siguientes cuatro pasos:

- i) A partir de la información acerca de los retornos de los activos, $\mathbf{r}_t = [r_{1t}, \dots, r_{dt}]'$ $t = 1, \dots, n$, se estiman los parámetros de la cópula y de las distribuciones marginales, utilizando cualquiera de las técnicas descritas en la Sección 4.
- ii) Con los parámetros estimados, se simulan Q vectores (de dimensión d) de números aleatorios, \mathbf{r}_{t+1}^q $q = 1, \dots, Q$, que tengan las mismas distribuciones (marginales y conjunta) que los retornos de los activos que conforman el portafolio, cuyo procedimiento se puede resumir en dos etapas:
 - Se simulan d variables uniformes $(0, 1)$, u_1, \dots, u_d , que tengan como distribución conjunta la cópula estimada en el numeral anterior, tal como se describe en la Sección 3.6.
 - Se calculan los retornos simulados, r_1, \dots, r_d , con base en u_1, \dots, u_d y en la identidad $r_i = F^{-1}(u_i)$, donde F^{-1} es la función inversa de distribución (ecuación (3.5)).
- iii) Para cada una de las simulaciones, se calcula una trayectoria del retorno del portafolio futuro, \tilde{r}_{t+k}^q :

$$\tilde{r}_{t+k}^q = \boldsymbol{\omega}' \mathbf{r}_{t+k}^q \quad q = 1, \dots, Q \quad (5.4)$$

donde $\boldsymbol{\omega} \in \mathbb{R}^d$ es el vector de ponderaciones del portafolio.

- iv) Una vez obtenidos los Q retornos del portafolio, se calcula la distribución de pérdidas y ganancias del portafolio. El $\widehat{VaR}_{\alpha, t+k}^{(p)}$ corresponderá al α -ésimo percentil de esta distribución, mientras el $\widehat{ES}_{\alpha, t+k}^{(p)}$ será el promedio de los retornos del portafolio que superen al $\widehat{VaR}_{\alpha, t+k}^{(p)}$.

Aunque a simple vista estos pasos parecen sencillos, su implementación se puede dificultar principalmente por tres factores. En primer lugar, en caso de que exista dependencia temporal en los datos y se utilicen las técnicas de estimación presentadas en la Sección 4.4, se hace necesario realizar pronósticos de los parámetros que cambian en el tiempo, e.g. la media y la varianza condicional en el caso de modelos AR-GARCH, lo cual incrementa los costos computacionales de las simulaciones; en segundo lugar, la viabilidad de este procedimiento está directamente ligado con el número de activos que constituyen

³⁵No obstante, Malevergne y Sornette [2006] proponen una metodología para calcular el VaR_α de un portafolio a partir de aproximaciones asintóticas. Estos autores ilustran esta aproximación para el caso de una cópula Gaussiana.

el portafolio, ya que si la dimensión del proceso es muy grande, nuevamente los costos computacionales son demasiado altos. Finalmente, con un horizonte temporal alto, el proceso de simulación tiende a hacerse cada vez más dispendioso, más aún si se tiene en cuenta la dependencia temporal.

5.1.2. Pruebas de desempeño (Backtesting) para el VaR_α y el ES_α de un portafolio.

- Modelos de *Backtesting* para el VaR_α

Una etapa fundamental en la implementación de modelos de riesgo de mercado se encuentra en la aplicación de pruebas de desempeño o *backtesting*, con las que se verifica que tan efectivos son el VaR_α y el ES_α para identificar los riesgos subyacentes dentro del portafolio. Estas pruebas requieren estimaciones recursivas del VaR_α para un período de tenencia dado, tal como se describe a continuación (en este caso se supone que se tiene información hasta el período n):

- Con la información hasta el período t , se calcula el valor en riesgo del portafolio con un nivel de confianza α , para el período $t + k$, $\widehat{VaR}_{\alpha,t+k}^{(p)}$, donde k representa el tiempo de tenencia del portafolio, con $t = n_1, n_1 + 1, \dots, n - k$.
- A partir de los valores en riesgo estimados, se construye la sucesión I_{t+k} , la cual se define como:

$$I_{t+k} = \begin{cases} 1 & \text{si } r_{t+k}^{(p)} > \widehat{VaR}_{\alpha,t+k}^{(p)} \\ 0 & \text{en otro caso} \end{cases} \quad (5.5)$$

La información relacionada con I_{t+k} constituye el principal insumo para la construcción de pruebas de desempeño, ya que ésta forma una sucesión de variables aleatorias Bernoulli con parámetro p , el cual corresponde a la probabilidad de que el retorno del portafolio exceda el VaR_α . Si las observaciones de I_{t+k} son *i.i.d.*, entonces la función de verosimilitud de esta secuencia sería:

$$L(p) = \prod_{i=1}^m p^{I_{t+k}} (1-p)^{1-I_{t+k}} = p^x (1-p)^{m-x} \quad (5.6)$$

donde x es el número de excepciones y m el número de observaciones incluidas en el *Backtesting* ($m = n - (n_1 + k) + 1$). Con base en (5.6) es posible encontrar el estimador *ML* del parámetro p , el cual es $\widehat{p} = x/m$. Si el VaR_α está midiendo de manera apropiada el riesgo del portafolio, éste valor debe ser igual a $(1 - \alpha)$. A partir de estas observaciones, Kupiec [1995] propuso una prueba de razón de verosimilitud para contrastar la hipótesis nula $H_0 : p = 1 - \alpha$. El estadístico para la prueba de "proporción de fallas" de Kupiec viene de la forma:

$$LR_{pof} = -2 \log \left(\frac{L(p)}{L(\widehat{p})} \right) = -2 \log \left(\frac{p^x (1-p)^{m-x}}{\widehat{p}^x (1-\widehat{p})^{m-x}} \right) \quad (5.7)$$

donde el numerador de (5.7) corresponde al valor de (5.6) evaluado bajo H_0 . Este estadístico se distribuye asintóticamente Chi-cuadrado con un grado de libertad.

Un supuesto importante detrás de esta prueba es que las observaciones de I_{t+k} son *i.i.d.*, no obstante, en especial en modelos que no tienen en cuenta la dependencia temporal en los datos, es posible que esta secuencia no satisfaga esta condición, lo que podría llevar a resultados indeseados en la cobertura de riesgo. Por ejemplo, si las excepciones se encuentran agrupadas en una parte de la muestra, el administrador de portafolio podría predecir que si en un día se presenta una excepción del VaR_α , la probabilidad de que se presente una excepción el siguiente día es mayor a $(1 - \alpha)$, y por lo tanto, debería tener en cuenta esta información para evitar subestimar el riesgo existente en el portafolio [Christoffersen, 2003].

Es por esto, que Christoffersen [2003] propone una prueba para contrastar la independencia entre las excepciones observadas dentro del *backtesting*. ésta se basa en el supuesto que las realizaciones de la sucesión de variables aleatorias $\{I_t\}_{n_1+k}^n$ siguen una cadena de Markov de orden uno, cuya matriz de transición viene de la forma:

$$\Pi_1 = \begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix} \quad (5.8)$$

donde $\pi_{ij} = P(I_t = j | I_{t-1} = i)$ con $i, j = \{0, 1\}$ son las probabilidades de transición de la cadena de Markov, tal que $\pi_{i0} + \pi_{i1} = 1$ para $i = \{0, 1\}$. A partir de esta condición, la matriz de transición (5.8) se puede describir como:

$$\Pi_1 = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix} \quad (5.9)$$

para estimar las probabilidades de transición π_{ij} , se hace necesario construir la función de verosimilitud para esta sucesión, la cual tiene la forma:

$$L(\Pi_1) = (1 - \pi_{01})^{m_{00}} \pi_{01}^{m_{01}} (1 - \pi_{11})^{m_{10}} \pi_{11}^{m_{11}} \quad (5.10)$$

donde m_{ij} representa el número de observaciones en las cuales $I_t = j$ e $I_{t-1} = i$, con $m_{00} + m_{01} + m_{10} + m_{11} = m$, siendo m el número total de observaciones incluidas dentro del *backtesting*. A partir de la optimización de (5.10), el estimador máximo verosímil de la matriz (5.8) es:

$$\hat{\Pi}_1 = \begin{bmatrix} \hat{\pi}_{00} & \hat{\pi}_{01} \\ \hat{\pi}_{10} & \hat{\pi}_{11} \end{bmatrix} = \begin{bmatrix} \frac{m_{00}}{m_{00}+m_{01}} & \frac{m_{01}}{m_{00}+m_{01}} \\ \frac{m_{10}}{m_{10}+m_{11}} & \frac{m_{11}}{m_{10}+m_{11}} \end{bmatrix} \quad (5.11)$$

A partir de la matriz (5.8) y algunas consideraciones acerca del comportamiento de sus probabilidades de transición, es posible establecer si las excepciones son independientes. Específicamente, para este ejercicio resulta interesante comparar las diferencias entre π_{01} , la probabilidad de que mañana se presente una excepción dado que hoy no ocurrió, y π_{11} , la probabilidad de que mañana el retorno del portafolio exceda el VaR_α

condicionado a que hoy lo superó. Bajo el supuesto de independencia, se debe cumplir que $\pi_{01} = \pi_{11} = \pi$ y por lo tanto la matriz de transición tomará la forma:

$$\Pi_0 = \begin{bmatrix} 1 - \pi & \pi \\ 1 - \pi & \pi \end{bmatrix} \quad (5.12)$$

en este caso, es posible demostrar que el estimador máximo verosímil de π , restringido a que $\pi_{01} = \pi_{11} = \pi$, es igual a $\hat{\pi} = \frac{m_{01} + m_{11}}{m}$. Con base en estos resultados, Christoffersen [2003] contrasta la hipótesis nula $H_0 : \pi_{01} = \pi_{11}$ a través de la siguiente prueba de razón de verosimilitud:

$$LR_{ind} = 2 \left(\ell \left(\hat{\Pi}_1 \right) - \ell \left(\hat{\Pi}_0 \right) \right) \quad (5.13)$$

donde $\ell(\cdot) = \log L(\cdot)$. Bajo H_0 , el estadístico LR_{ind} distribuye asintóticamente χ^2 con un grado de libertad. En caso que se rechace H_0 , se diría que las excepciones del VaR_α son dependientes y por lo tanto el modelo utilizado debería modificarse para corregir estas deficiencias.

Finalmente, aunque las dos pruebas de hipótesis anteriores brindan información acerca del desempeño del VaR_α , sería deseable probar de manera conjunta si las observaciones de I_{t+k} son independientes y si el promedio de excesos en realidad se acerca al nivel de significancia supuesto para el modelo $(1 - \alpha)$. Para poder tener en cuenta esta posibilidad, Christoffersen [2003] propone la siguiente prueba de razón de verosimilitud, en la que se contrasta la hipótesis nula $H_0 : \pi_{01} = \pi_{11} = p$:

$$LR_{cc} = 2 \left(\ell \left(\hat{\Pi}_1 \right) - \ell \left(\Pi_p \right) \right) \quad (5.14)$$

nuevamente $\ell(\cdot) = \log L(\cdot)$, mientras que la matriz Π_p corresponde a la matriz de transición (5.12) evaluada en $\pi = p = (1 - \alpha)$, siendo α el nivel de confianza considerado en el modelo VaR_α . Bajo la hipótesis nula, el estadístico LR_{cc} asintóticamente se distribuye χ^2 con dos grados de libertad. Tal como lo destaca Christoffersen [2003], una forma sencilla de calcular el estadístico LR_{cc} es a partir de las ecuaciones (5.7) y (5.13), de la siguiente forma:

$$\begin{aligned} LR_{cc} &= 2 \left(\ell \left(\hat{\Pi}_1 \right) - \ell \left(\Pi_p \right) \right) \\ LR_{cc} &= 2 \left(\ell \left(\hat{\Pi}_1 \right) - \ell \left(\hat{\Pi}_0 \right) + \ell \left(\hat{\Pi}_0 \right) - \ell \left(\Pi_p \right) \right) \\ LR_{cc} &= 2 \left(\ell \left(\hat{\Pi}_1 \right) - \ell \left(\hat{\Pi}_0 \right) \right) + 2 \left(\ell \left(\hat{\Pi}_0 \right) - \ell \left(\Pi_p \right) \right) \\ LR_{cc} &= LR_{ind} + LR_{pof} \end{aligned} \quad (5.15)$$

para llegar a este resultado, nótese que cuando $\pi_{01} = \pi_{11} = \pi$, la función de verosimilitud (5.10) toma la forma (5.6) y por lo tanto el LR_{pof} se puede reescribir como $LR_{pof} = 2 \left(\ell \left(\hat{\Pi}_0 \right) - \ell \left(\Pi_p \right) \right)$.

Junto con el *Backtesting*, otra forma de evaluar el VaR_α es a través de las denominadas funciones de pérdida, las cuales tienen en cuenta no sólo el número de excepciones que

haya tenido el modelo de riesgo, sino que también consideran la magnitud de dichas excepciones. En general, para un portafolio con un período de tenencia k , un nivel de confianza α e información disponible hasta el período t , la función de pérdida L_{t+k} viene de la forma [Micocci y Masala, 2004]:

$$L_{t+k} = \begin{cases} f\left(r_{t+k}^{(p)}, \widehat{VaR}_{t+k}^{(p)}\right) & \text{si } r_{t+k}^{(p)} > \widehat{VaR}_{t+k}^{(p)} \\ g\left(r_{t+k}^{(p)}, \widehat{VaR}_{t+k}^{(p)}\right) & \text{si } r_{t+k}^{(p)} \leq \widehat{VaR}_{t+k}^{(p)} \end{cases} \quad (5.16)$$

donde $f(x, y)$ y $g(x, y)$ son dos funciones tales que $f(x, y) \geq g(x, y)$ para todo $y \in \mathbb{R}$. En otras palabras, dado un valor del VaR_α , la función de pérdida siempre castigará de manera más fuerte cuando se excede el valor en riesgo. Nótese que la función L_{t+k} toma la forma de la sucesión I_{t+k} (ecuación (5.5)) en el caso en que $f(x, y) = 1$ y $g(x, y) = 0$.

Uno de los principales usos de las funciones de pérdida es en la construcción de índices de desempeño para cada uno de los modelos de riesgo, los cuales se definen como:

$$L = \sum_{t=1}^m L_{t+k} \quad (5.17)$$

de esta manera, a medida que L es más alto, dado que L_{t+k} es no decreciente en $r_{t+k}^{(p)}$, el modelo de riesgo presenta un peor desempeño. Así, el mejor modelo será aquel que minimice (5.17) y por lo tanto, si la cópula es utilizada con el fin de calcular medidas de riesgo para el portafolio, esta observación podría ser considerada como un criterio adicional para la selección de la cópula (ver Sección 4.3). Para la construcción de estos índices, se han propuesto diferentes funciones de pérdida, siendo una de las más tradicionales la propuesta por López [1998]:

$$L_{t+k} = \begin{cases} 1 + \left(r_{t+k}^{(p)} - \widehat{VaR}_{t+k}^{(p)}\right)^2 & \text{si } r_{t+k}^{(p)} > \widehat{VaR}_{t+k}^{(p)} \\ 0 & \text{si } r_{t+k}^{(p)} \leq \widehat{VaR}_{t+k}^{(p)} \end{cases} \quad (5.18)$$

Adicionalmente, este autor comenta que las pruebas de Christoffersen, descritas anteriormente, tienen baja potencia y recomienda pruebas adicionales basadas en la función de pérdida cuadrática definida en (5.18).

Por otra parte, Caporin [2003] destaca que el uso de la función $f_L = 1 + \left(r_{t+k}^{(p)} - \widehat{VaR}_{t+k}^{(p)}\right)^2$ para la construcción del índice de desempeño puede llevar a desechar modelos apropiados para la medición de riesgo, si estos presentan excepciones de gran tamaño y, en consecuencia, propone el uso de las siguientes funciones $f(x, y)$ en la definición de L_{t+k} , las cuales son menos sensibles ante valores atípicos y establecen una medida de

distancia entre el retorno del portafolio y el valor en riesgo calculado:

$$\begin{aligned}
 f_1 \left(r_{t+k}^{(p)}, \widehat{VaR}_{t+k}^{(p)} \right) &= \left| 1 - \frac{r_{t+k}^{(p)}}{\widehat{VaR}_{t+k}^{(p)}} \right| \\
 f_2 \left(r_{t+k}^{(p)}, \widehat{VaR}_{t+k}^{(p)} \right) &= \frac{\left(\left| r_{t+k}^{(p)} \right| - \left| \widehat{VaR}_{t+k}^{(p)} \right| \right)^2}{\left| \widehat{VaR}_{t+k}^{(p)} \right|} \\
 f_3 \left(r_{t+k}^{(p)}, \widehat{VaR}_{t+k}^{(p)} \right) &= \left| r_{t+k}^{(p)} - \widehat{VaR}_{t+k}^{(p)} \right|
 \end{aligned} \tag{5.19}$$

Adicionalmente, con el fin de evitar los problemas presentados por la función de pérdida (5.18), este autor propone utilizar cada una de las funciones definidas en (5.19) sobre toda la muestra, no solamente sobre los valores que exceden el VaR_α , en otras palabras, $f_i = g_i, i = 1, 2, 3$.

- Modelos de *Backtesting* para el ES_α

De la misma forma que en el caso anterior, existen pruebas para evaluar el desempeño del ES_α en la medición de riesgos de mercado. Dichas pruebas se basan en la siguiente observación: ya que el ES_α es el valor esperado de las pérdidas del portafolio, condicionadas a que estas excedieron el VaR_α , al realizar una prueba de desempeño sobre la información histórica de las series, la diferencia promedio entre los retornos que excedan el VaR_α y el ES_α estimados para esos mismos días debería ser cero. Formalmente, para un período de tenencia k y un nivel de confianza α , se define la secuencia de “excesos de retorno” del portafolio, $H_{t+k}^{(p)}$, como:

$$H_{t+k}^{(p)} = \left(r_{t+k}^{(p)} - \widehat{ES}_{t+k}^{(p)} \right) I_{t+k} \tag{5.20}$$

donde I_{t+k} está definido como en (5.5). De aquí, $H_{t+k}^{(p)}$ es una secuencia de variables aleatorias *i.i.d* y, si el ES_α mide de manera adecuada el riesgo, $E \left(H_{t+k}^{(p)} \right) = 0$. Para contrastar esta hipótesis, McNeil y Frey [2000] proponen el uso de un estadístico t calculado a partir de técnicas *bootstrap*³⁶, las cuales no realizan supuestos sobre la distribución de los datos. Vale la pena mencionar que en aquellos casos en los que la varianza de los datos no es constante en el tiempo, es necesario reescalar (5.20) por el factor $1/\widehat{\sigma}_{t+k}$, donde $\widehat{\sigma}_{t+k}$ es el pronóstico de la desviación estándar para el período $t+k$, con la información disponible hasta el período t .

5.2. Modelos de riesgo de crédito. Al igual que en el riesgo de mercado, las cópulas pueden ser usadas para ayudar a explicar de mejor manera el riesgo crediticio. En esta

³⁶Véase Efron y Tibshirani [1993]

sección, se presenta la aproximación propuesta por Li [2000], la cual ha sido adoptada dentro de la metodología utilizada por J.P. Morgan para la medición de riesgo³⁷. Cabe destacar que aparte de la propuesta de Li, existen otras aplicaciones de cópulas para medición de riesgo de crédito, las cuales se discuten en Frey et al. [2001]³⁸.

De acuerdo con la metodología de Li [2000], la distribución de pérdidas de un portafolio invertido en instrumentos de crédito pueden ser obtenida a través de técnicas de Monte Carlo donde se simulan posibles “eventos crediticios”³⁹.

Suponga que se tiene un portafolio que contiene n activos riesgosos (contrapartes) en un momento inicial $t_0 = 0$ y un horizonte fijo T . La exposición del portafolio a cada una de las contrapartes está representada por $N_i, i = 1, \dots, n$. En el caso más simple, se tienen s escenarios en los que es posible que las contrapartes incumplan o no incumplan⁴⁰. Para cada uno de los escenarios, la pérdida que presentaría el portafolio está definida por:

$$L_{ij} = \begin{cases} 0 & \text{Si } i \text{ no incumple en el escenario } j \\ N_i(1 - R_i) & \text{Si } i \text{ incumple en el escenario } j \end{cases} \quad (5.21)$$

donde $i = 1, \dots, n, j = 1, \dots, s$ y R_i es la tasa de recobro, la que por simplicidad se supone no estocástica. En el momento T , la pérdida total del portafolio para el escenario j será:

$$L_j(\omega) = \sum_{i=1}^n L_{ij}\omega_i \quad (5.22)$$

donde $\omega = [\omega_1, \dots, \omega_n]'$ es el vector de posiciones del portafolio⁴¹. De esta manera, la pérdida esperada del portafolio es:

$$EL(\omega) = \frac{1}{s} \sum_{j=1}^s L_j(\omega) \quad (5.23)$$

A diferencia del modelo de medición de riesgo de mercado, donde VaR_α corresponde al α -ésimo cuantil de la distribución de pérdidas del portafolio, en este modelo, el VaR asociado al riesgo de crédito, $CreditVaR_\alpha$, está definido por:

$$CreditVaR_\alpha(\omega) = ML_\alpha(\omega) - EL(\omega) \quad (5.24)$$

³⁷Esta sección sigue de cerca la descripción del modelo de Li [2000] realizada por Di Clemente y Romano [2004].

³⁸Para una discusión detallada de modelos de riesgo de crédito, véase McNeil et al. [2005], Rachev et al. [2005], Nyfeler [2000], KMV Corporation [2003], JP Morgan & Incorporated [1997] y Credit Suisse Financial Products [1997].

³⁹En general, un evento crediticio se define como la ocurrencia de al menos una de las siguientes situaciones: *i*) Bancarrota; *ii*) no pago (ya sea del principal o de los intereses); *iii*) incumplimiento de la obligación; *iv*) aceleración de la obligación; *v*) repudio/moratoria y *vi*) reestructuración.

⁴⁰En este modelo, se desconoce la posibilidad de que las contrapartes “migren” entre diferentes calificaciones de riesgo.

⁴¹Di Clemente y Romano [2004] discuten la metodología para obtener las participaciones que optimizan la relación riesgo/retorno del portafolio.

donde α es un nivel de confianza dado, por ejemplo 0,95 o 0,99, y $ML_\alpha(\omega)$ es el α -ésimo cuantil de la distribución de L_j . Por su parte, el ES_α se define como:

$$ES_\alpha(\omega) = ML_\alpha(\omega) + \frac{1}{(1-\alpha)s} \sum_{j=1}^s [L_j(\omega) - ML_\alpha(\omega)] I_{\{L_j(\omega) - ML_\alpha(\omega) > 0\}} \quad (5.25)$$

Así, la correcta medición de riesgo depende de la generación de escenarios confiables para la creación de la distribución de pérdidas del portafolio. En general, esta metodología se implementa en tres pasos, los cuales se presentan a continuación.

- i) En primer lugar, se describe el comportamiento univariado de cada una de las contrapartes a través de la variable aleatoria τ_i , la cual indica el tiempo que transcurre entre el momento inicial t_0 y el incumplimiento para la i -ésima contraparte⁴². La probabilidad de que no se presente un incumplimiento hasta en el momento t para la contraparte i se encuentra determinada por su función de sobrevivencia, la cual se define como:

$$S_i(t) = P(\tau_i > t) = \exp \left\{ - \int_0^t h_i(u) du \right\} \quad (5.26)$$

donde $h_i(u)$ es la función *hazard* de la contraparte i ⁴³. Por simplicidad, es común suponer que esta función es constante en el tiempo, es decir, depende únicamente de las características propias de la contraparte i . Bajo este supuesto, τ_i sigue una distribución exponencial con parámetro h_i :

$$\begin{aligned} F_i(t) &= P(\tau_i \leq t) \\ &= 1 - \exp \left\{ - \int_0^t h_i du \right\} \\ &= 1 - e^{-h_i t} \end{aligned} \quad (5.27)$$

de esta manera, el modelamiento univariado de los tiempos hasta el incumplimiento τ_i se encuentra determinado completamente por la función h_i . En la práctica, h_i puede ser obtenido de diferentes formas, entre las que se destacan:

- El uso de probabilidades de incumplimiento publicadas por agencias calificadoras, por ejemplo, Standard and Poor's y Moody's [Di Clemente y Romano, 2004]. Si $q_i(0, t)$ es la tasa de incumplimiento promedio acumulada en el horizonte $(0, t)$, ésta se puede tomar como un estimador de $F_i(t)$ y al reemplazarla en (5.27) se obtiene:

$$h_i = \frac{-\log[1 - q_i(0, t)]}{t} \quad (5.28)$$

- A partir de la metodología de Merton, descrita en Delianedis y Geske [1998].

⁴²De acuerdo con esta definición, se dice que la contraparte i incumple cuando $\tau_i < T$.

⁴³Una discusión detallada de las funciones de supervivencia y *hazard* se encuentra en Klein y Moeschberger [2003].

- Basados en la construcción de una “curva de crédito”, donde se genera una estructura a plazos de las tasas de incumplimiento [Li, 1998, 2000].
- ii) En segundo lugar, la relación existente entre los tiempos hasta el incumplimiento de las contrapartes se modela a partir de su función de probabilidad conjunta $F_{\tau_1, \dots, \tau_n}$, definida como:

$$F_{\tau_1, \dots, \tau_n}(t_1, \dots, t_n) = P(\tau_1 \leq t_1, \dots, \tau_n \leq t_n) \quad (5.29)$$

utilizando el teorema de Sklar (Sección 3.2) sobre la ecuación anterior, se tiene que:

$$F_{\tau_1, \dots, \tau_n}(t_1, \dots, t_n) = C(F_1(t_1), \dots, F_n(t_n)) \quad (5.30)$$

para una cópula C . En esta ecuación, las funciones de distribución marginales corresponden a las distribuciones exponenciales obtenidas en el paso anterior.

- iii) Por último, la distribución de pérdidas del portafolio, el $CreditVaR_\alpha$ y el ES_α son calculados a partir de las ecuaciones (5.21) a (5.25) y los algoritmos de simulación propuestos en la Sección 3.6.

En aplicaciones particulares de esta metodología, Li [2000] propone el uso de la cópula Gaussiana para describir el comportamiento multivariado de los tiempos hasta el default (ecuación (5.29)), donde se utiliza como matriz de correlaciones la matriz estimada para los activos asociadas a las contrapartes. Esta aproximación fue extendida a otras familias de cópulas por Meneguzzo y Vecchiato [2002] y Mashal y Zeevi [2002], entre otros.

Finalmente, cabe destacar que la anterior metodología resulta apropiada incluso para la modelación de instrumentos de riesgo más estructurados, como por ejemplo los *basket default swaps* y los *collateralized debt obligation (CDO)*⁴⁴. En estos casos, la construcción de las pérdidas del portafolio (ecuación (5.21)) requiere ser ajustada para tener en cuenta las características propias del instrumento evaluado. Una aplicación de esta metodología a *basket default swaps* y *CDOs* se encuentra en Cherubini et al. [2004].

6. UNA APLICACIÓN DE LAS CÓPULAS A LA VALORACIÓN DEL RIESGO DE MERCADO

En secciones anteriores, se presentaron las principales características de la cópula y la forma de aplicarlas en la valoración de riesgo de mercado y de crédito. A continuación, se presentan los resultados de un ejercicio para calcular el Valor en Riesgo (VaR_α) de un portafolio conformado por tres activos representativos del mercado colombiano: el Índice General de la Bolsa de Colombia (IGBC), la tasa representativa del mercado (TRM) y un producto representativo del mercado de deuda, un título de deuda del Gobierno Nacional (TES) con vencimiento de 10 años. Para este ejercicio, se seleccionó el período diario comprendido entre el 8 de octubre de 2002 y el 31 de mayo de 2007, para un total de 1135 observaciones.

Los gráficos del negativo de los retornos logarítmicos (pérdidas) de cada una de las series de estudio se encuentran en el Gráfico 6.1. En una primera etapa, esta figura permite

⁴⁴Para una descripción de este tipo de instrumentos véase JP Morgan & Incorporated [2000].

identificar algunas características de las series de estudio. En primer lugar, la volatilidad de los retornos muestra patrones irregulares en el tiempo, sugiriendo la presencia de conglomerados de volatilidad; en segundo lugar, en diferentes puntos de la muestra, es notoria la presencia de valores atípicos. Como es común en el estudio de activos financieros, este comportamiento indica que la distribución de los retornos de cada uno de los activos esta descrita por una distribución de colas pesadas [Engle, 1982]. En efecto, cuando se comparan las distribuciones empíricas de cada uno de los activos con una distribución normal, las principales diferencias se encuentran en las colas de la distribución (Gráfico 6.2).

Como se mencionó en la Sección 4.4, un factor importante en la modelación de la dependencia multivariada es conocer si existe algún tipo de patrón o regularidad temporal para las series de estudio. Después de analizar funciones de autocorrelación sobre los retornos en niveles y al cuadrado, al igual que otras herramientas tradicionales en series de tiempo, se identificó un modelo VAR(10)- GARCH(1,1) sobre las tres series⁴⁵. Las pruebas de diagnóstico sobre los residuales estandarizados del modelo se encuentran en el Anexo B, en general estas pruebas muestran que no existen indicios de autocorrelación, ni de efectos ARCH remanentes.

A partir de este punto y siguiendo de cerca el enfoque de Carmona [2004], se estimaron las funciones de distribución marginales para los residuales estandarizados de cada ecuación a través de un estimador semiparamétrico, en el que en el centro de la distribución se utiliza la función de distribución empírica mientras que las colas de la distribución son ajustadas a partir de la teoría del valor extremo. Esta aproximación se justifica por dos razones. Por un lado, esta aproximación evita realizar supuestos fuertes sobre la forma funcional de las distribuciones marginales. Por otro lado, la información disponible a partir del uso de las distribuciones empíricas resulta insuficiente en el análisis de la probabilidad de ocurrencia de valores extremos en las series, en consecuencia, es mejor utilizar estimadores paramétricos o semiparamétricos en la estimación de las colas de la distribución.

La función de distribución ajustada para los residuales de la ecuación de las pérdidas del IGBC se presenta en el Gráfico 6.3. Adicional a la distribución, se presentan el histograma de la variable aleatoria uniforme derivada de la función de distribución marginal $(u_{1t} = \hat{F}(r_{1t}))$, junto con un gráfico de Q-Q en el que se comparan los cuantiles de la función de distribución estimada con respecto a los de una variable aleatoria uniforme $(0, 1)$. Como se mencionó en la Sección 4.2, un supuesto importante para que la estimación de la cópula sea apropiada es que los valores pertenecientes a la pseudo muestra $\mathbf{u}_t = [u_{1t}, \dots, u_{dt}]'$ sigan una distribución uniforme estándar. Las gráficas sugieren que la aproximación semiparamétrica de Carmona es apropiada. Esta conclusión se confirmó realizando una prueba de Kolmogorov – Smirnov, en la cual no se encuentra evidencia

⁴⁵Teniendo en cuenta que los residuales del modelo estimado no siguen una distribución normal, es necesario utilizar el método de estimación QML (*Quasi Maximum Likelihood*), ver Gouriéroux y Jasiak [2001].

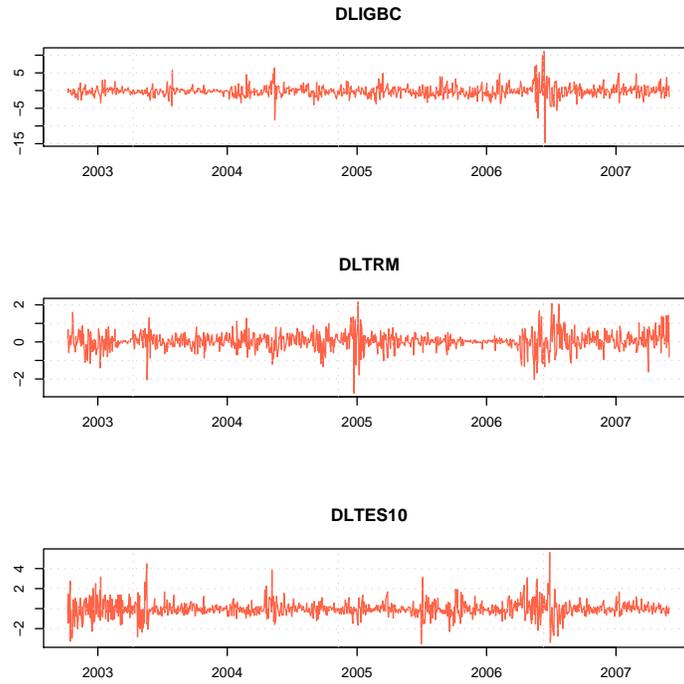


GRÁFICO 6.1. Negativo de los retornos logarítmicos de las variables seleccionadas

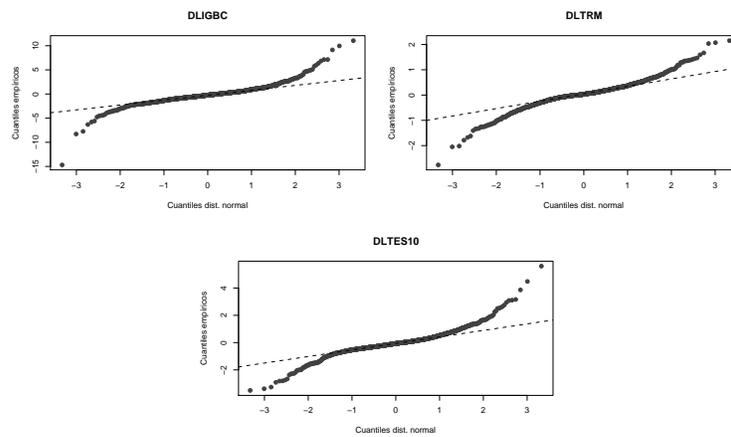


GRÁFICO 6.2. Gráficos Q-Q

estadística para rechazar la hipótesis nula de que la distribución marginal ajustada sigue una distribución uniforme estándar⁴⁶.

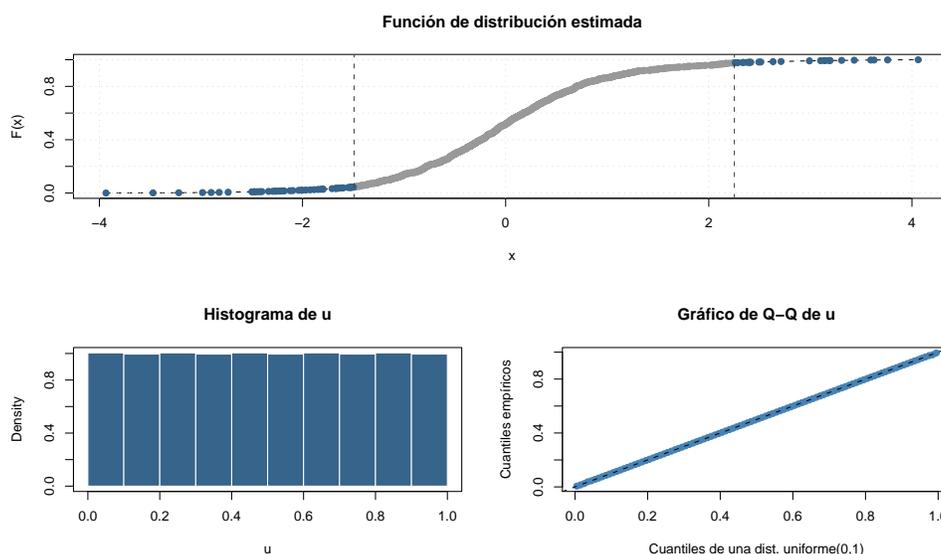


GRÁFICO 6.3. Estimación semiparamétrica de las colas de la distribución (-Dlog IGBC)

Con la pseudo muestra obtenida en el paso anterior, se procede a estimar la copula C . No obstante, no existe un criterio único ni definitivo para seleccionar una copula *a-priori* sobre estos datos. En cambio, se estiman varias copulas para describir el comportamiento conjunto de estas tres series. Las copulas seleccionadas son cinco, las cuales describen diferentes patrones de comportamiento multivariado: la de independencia, dos copulas elípticas, la normal y la t multivariada, y tres copulas de Arquímedes, la Gumbel, Clayton y Frank, las cuales se describen en la Sección 3.5.

A partir de los resultados de la estimación, se calculan los criterios de selección de la copula descritos en la Sección 4.3, los cuales se presentan en la Tabla 6.1. De acuerdo con éstos, las copulas que muestran un mejor ajuste son las copulas elípticas, ya sea la normal o la t multivariada. De los siete criterios de selección evaluados, la copula t obtiene los mejores resultados en cinco de ellos (log verosimilitud, el criterio AIC y las tres estadísticas de bondad de ajuste con respecto a la copula empírica), mientras que la normal obtiene los resultados más satisfactorios en los criterios BIC y HQ. En el Gráfico 6.4, se presentan las curvas de nivel asociadas a la copula empírica y la copula t estimada. En éste, resulta evidente el grado de correspondencia entre los datos y esta copula. Teniendo

⁴⁶Aunque en el Gráfico 6.3 se presentan únicamente los resultados para los residuales de la ecuación asociada a las pérdidas del IGBC, se obtienen conclusiones similares para los residuales de las otras dos ecuaciones.

en cuenta estos resultados, las secciones siguientes usarán las cópulas t y Gaussiana para el cálculo de las medidas de riesgo de mercado.

Cópula	Log(verosimilitud)	Grados de libertad	AIC (Joe)	BIC	HQ	Distancia (Romano)	AD	IAD
Normal	42,71	3	39,71	-64,35	-73,73	13,85	3,12	2,39
t	44,83	4	40,83	-57,16	-72,89	13,25	2,64	2,27
Gumbel	9,8	1	8,80	-12,58	-15,70	31,64	4,15	9,67
Clayton	3,23	1	2,23	0,56	-2,56	37,09	4,39	11,14
Frank	10,44	1	9,44	-13,85	-16,97	28,90	4,25	7,55
Independencia	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>	46,90	5,34	21,12

TABLA 6.1. Criterios de selección para las cópulas analizadas

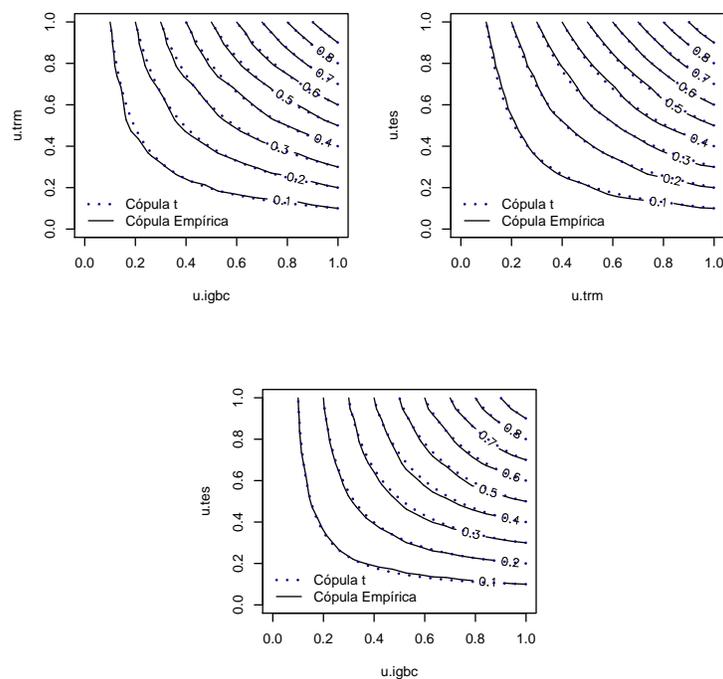


GRÁFICO 6.4. Curvas de nivel de la cópula t estimada y la cópula empírica

Por último, después de seleccionar las cópulas más apropiadas, es posible utilizar éstas para la medición de riesgo de mercado, en particular, en el cálculo del valor en riesgo, VaR_{α} , y la pérdida esperada en las colas, ES_{α} , para un portafolio conformado por estos tres activos. Los resultados obtenidos siguen la aproximación realizada por simulación de

Monte Carlo, descrita en la Sección 5.1.1, suponiendo que la participación de cada uno de los tres activos representa un tercio del portafolio. Utilizando la información disponible hasta el 31 de mayo de 2007 y un período de tenencia de un día, la distribución de pérdidas simulada a partir de la cópula t se presenta en el Gráfico 6.5, donde, para un nivel de confianza de 95 %, el VaR_{95} y el ES_{95} obtenidos son 0,815 y 1,259, respectivamente. En otras palabras, para un portafolio con un valor total de 100 millones de pesos repartido por igual entre estos tres activos, para el primero de junio de 2007, con un nivel de confianza de 95 % se esperaba que las pérdidas del portafolio no superasen 815 mil pesos y, en caso que este nivel hubiera sido sobrepasado, la pérdida promedio alcanzaría 1,26 millones de pesos.

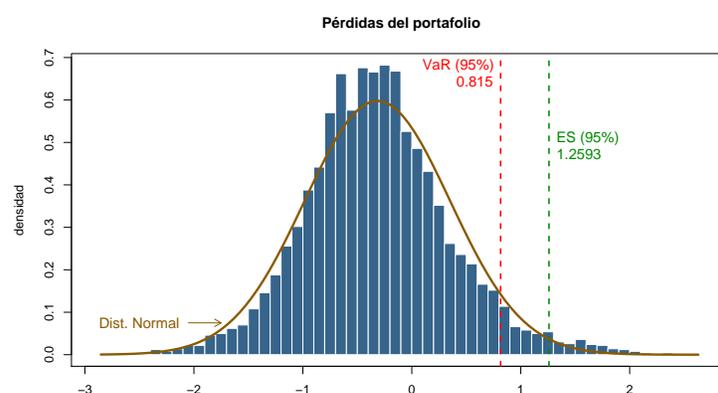


GRÁFICO 6.5. Distribución simulada del portafolio y medidas de riesgo para el 1/06/2007 a partir de la cópula t (período de tenencia un día)

6.1. Pruebas de desempeño (*Backtesting*) del VaR_{α} y el ES_{α} . En la sección anterior, se estimaron el VaR_{α} y el ES_{α} para un portafolio compuesto por tres activos: el índice accionario colombiano (IGBC), la tasa de cambio peso dólar (TRM) y un título representativo de los títulos de deuda a diez años (TES) para el primero de junio de 2007 y un nivel de confianza de 95 %. A continuación, se evalúa el uso de las cópulas en la medición de riesgo de mercado utilizando las pruebas de desempeño (*Backtesting*) presentadas en la Sección 5.1.2. Estas pruebas, por un lado, permiten establecer si las cópulas seleccionadas exhiben un desempeño satisfactorio comparado con otros métodos estándar de medición de riesgo de mercado y, por otro lado, se convierten en un criterio adicional para la selección de la cópula.

Como se mencionó en la Sección 5.1.2, las pruebas de *backtesting* juegan un papel fundamental en la medición de riesgo, ya que estas permiten establecer, sobre un rango de días observado, si las metodologías aplicadas para cuantificar el riesgo de un portafolio muestran un desempeño satisfactorio. Por ejemplo, para el valor en riesgo, se utilizan como indicadores tanto del número de ocasiones en que los retornos superan el VaR_{α}

como la magnitud de estas excepciones. Por su parte, para la pérdida esperada en las colas, se evalúan las diferencias entre la magnitud promedio de las excepciones esperada y observada.

Con el fin de establecer si las medidas de riesgo de mercado basadas en los modelos de cópulas, incluyendo la modelación del primer y segundo momento condicional a través de modelos VAR – GARCH, exhiben ventajas sobre otras metodologías, se realizan pruebas de desempeño para otros modelos simples y ampliamente utilizados en la medición de riesgo. En particular, este ejercicio evalúa el desempeño del cálculo del VaR_α por la metodología de *RiskMetrics*® y aquella que supone normalidad multivariada de los retornos. Estas dos últimas aproximaciones se describen en detalle en Melo y Becerra [2006]. Cabe destacar que mientras los modelos de cópula modelan los dos primeros momentos de las series, la metodología *RiskMetrics*® únicamente modela dependencia en el segundo momento y la metodología basada en la normalidad multivariada de los retornos, supone que las series son *i.i.d.*

Los gráficos de *Backtesting*, con un nivel de confianza de $\alpha = 0,95$, se presentan en el Gráfico 6.6. En estos, es posible observar que para la metodología que supone normalidad multivariada de los retornos de los activos, el VaR_{95} presentó 13 excepciones, que es exactamente el número de excepciones esperadas. No obstante, la mayoría de estas (11) se encuentran en los primeros 30 días de la muestra, lo que contradice el principio de independencia y además, evidencia el sesgo que posee esta metodología en la estimación del riesgo.

Por su parte, al evaluar el número de excepciones del VaR_{95} obtenido por la metodología de *RiskMetrics*®, ésta última muestra un comportamiento altamente satisfactorio, 14 excepciones. Este resultado contrasta con el número de excepciones obtenido en la aplicación de la cópula *t* y la cópula normal, en las que se presentan 18 y 20 excepciones, respectivamente. Si bien este resultado sugeriría que la metodología de *RiskMetrics*® resulta superior a la de cópula, es necesario realizar análisis adicionales, como por ejemplo, realizar pruebas de hipótesis sobre el proceso de excepciones (ver Sección 5.1.2), observar el comportamiento del valor en riesgo para niveles de confianza mayores (e.g. 99% o 99,5%) los cuales son comúnmente usados en la medición de riesgos de mercado y observar las funciones de pérdida, las cuales permitirán establecer con un mayor detalle cuál es la mejor técnica para cuantificar el riesgo del portafolio.

Los resultados de las pruebas de desempeño del VaR_α calculado por normalidad y *RiskMetrics*®, para niveles de confianza $\alpha = 0,95$, $\alpha = 0,99$ y $\alpha = 0,995$ se presentan en la Tabla 6.2. Como se comentó, debido al episodio de alta volatilidad registrado en los primeros días de la muestra, el VaR_{95} suponiendo normalidad reportó el número de excepciones esperadas, no obstante, los resultados de la prueba de independencia propuesta por Christoffersen sugieren que el número de excepciones no es un proceso independiente.

Para el VaR_α calculado bajo el supuesto de normalidad, a medida en que se aumenta los niveles de confianza, el número de excepciones observado supera de manera sistemática

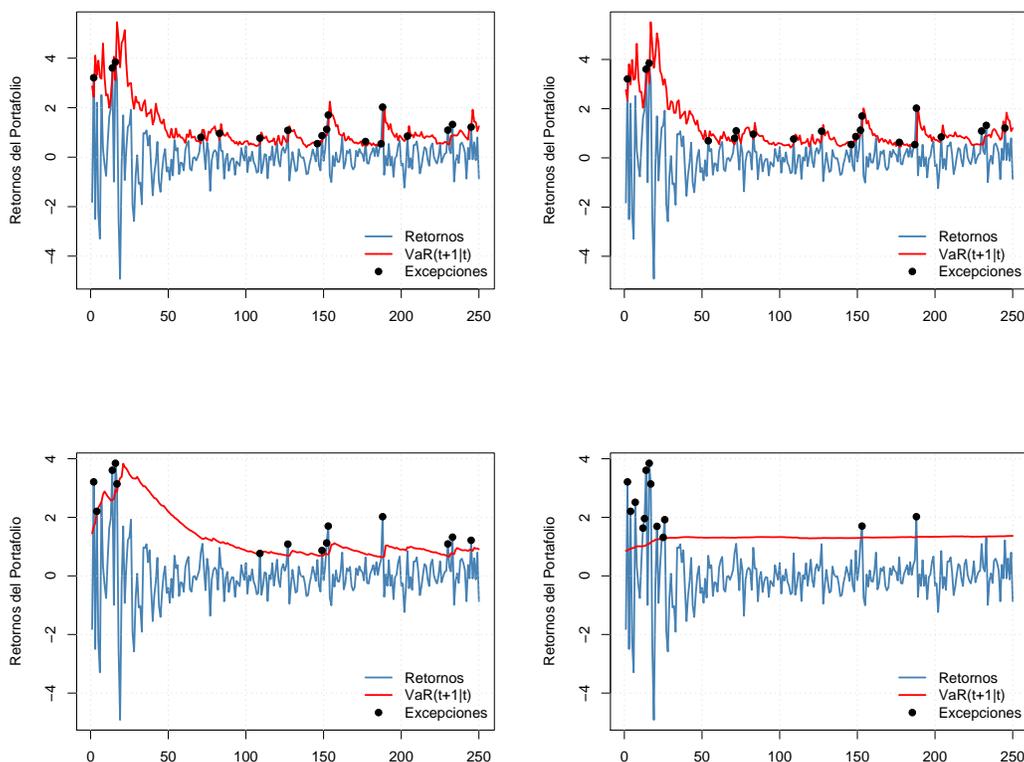


GRÁFICO 6.6. Gráficos de Backtesting para los modelos VAR-GARCH cópula t (panel superior izquierdo), VAR-GARCH cópula normal (panel superior derecho), el método de RiskMetrics[®] (panel inferior izquierdo) y el de normalidad (panel inferior derecho)

el número de excepciones esperado, lo que es ratificado por el resultado de la prueba de proporción de fallas. Al igual que en el método de normalidad, el número de excepciones obtenido por el método *RiskMetrics*[®] sobrestima el nivel de riesgo a medida que el nivel de confianza aumenta. No obstante, ya que este último método tiene en cuenta la varianza condicional de la serie, no se encuentra evidencia para suponer que el número de excepciones no es un proceso independiente (Tabla 6.2).

De otra parte, los resultados obtenidos con la cópula t y la cópula normal son satisfactorios ya que la proporción de fallas observada se encuentra cerca de su valor esperado y además, las pruebas de independencia sugieren que las excepciones se distribuyen aleatoriamente en el período de evaluación (Tabla 6.3). Relativo a las metodologías anteriores, el VaR_{α} obtenido con la combinación de modelos VAR-GARCH y cópulas exhibe

Nivel de confianza	Normal			RiskMetrics®		
	95 %	99 %	99,5 %	95 %	99 %	99,5 %
Excepciones	13	10	7	14	7	6
Excepciones esperadas	13	3	1	13	3	1
Prop. de fallas (POF)	5,2 %	4,0 %	2,8 %	5,6 %	2,8 %	2,4 %
POF esperadas	5,0 %	1,0 %	0,5 %	5,0 %	1,0 %	0,5 %

Pruebas de desempeño	Normal			RiskMetrics®		
Prop. de fallas (POF) (valor-p)	0,03 (0,87)	13,01 (<0,01)	12,80 (<0,01)	0,20 (0,66)	5,53 (0,02)	9,45 (<0,01)
Independencia (valor-p)	9,57 (<0,01)	8,45 (<0,01)	6,74 (<0,01)	1,54 (0,21)	1,85 (0,17)	0,30 (0,59)
POF e independencia (valor-p)	9,60 (<0,01)	21,47 (<0,01)	19,54 (<0,01)	1,74 (0,41)	7,34 (0,02)	9,75 (0,01)

TABLA 6.2. Resultados obtenidos suponiendo Normalidad y por el método *RiskMetrics*® (Muestra: 2006/05/19-2007/05/31)

mejores resultados, debido a que, por un lado, considera los cambios en el primer y segundo momento condicional de las series que conforman el portafolio y, por otro lado, no realiza supuestos fuertes sobre las distribuciones marginales, permitiendo que éstas muestren colas más pesadas que las de la distribución normal.

Hasta el momento, el desempeño del VaR_α obtenido por diferentes metodologías se ha evaluado a partir de indicadores derivados de la frecuencia y el orden en el cual se presentan las excepciones del VaR_α . No obstante, como se mencionó en la Sección 5.1.2, otro factor importante para la evaluación de los modelos de medición de riesgo se encuentra en las magnitudes de las pérdidas que superen al VaR_α . En consecuencia, en la Tabla 6.4 se presentan los valores de las funciones de pérdida descritas anteriormente (ecuaciones (5.18) y (5.19)) para cada una de las metodologías consideradas. Por simplicidad, se supone que para todos los casos $g(x, y) = 0$ (ecuación (5.16)).

De acuerdo con la Tabla 6.4, las funciones de pérdida para las metodologías derivadas de las cópulas son sustancialmente menores que para el VaR_α por normalidad y por *RiskMetrics*®, lo que respalda el uso de las cópulas en la medición de riesgo. Aunque para $\alpha = 0,95$ la cópula normal exhibe un desempeño ligeramente superior, la cópula t es aquella que muestra un mejor comportamiento, factor que debería ser tenido en la selección definitiva de la cópula.

Por último, otro factor importante es observar el comportamiento del ES_α para los modelos apropiados para la medición de riesgo. Al evaluar el desempeño del ES_α para la cópula t y la cópula normal utilizando la metodología propuesta en la segunda parte de

Nivel de confianza	Cópula t			Cópula normal		
	95 %	99 %	99,5 %	95 %	99 %	99,5 %
Excepciones	18	3	2	20	3	3
Excepciones esperadas	13	3	1	13	3	1
Proporción de fallas (POF)	7,2 %	1,2 %	0,8 %	8,0 %	1,2 %	1,2 %
POF esperadas	5,0 %	1,0 %	0,5 %	5,0 %	1,0 %	0,5 %

Pruebas de desempeño	Cópula t			Cópula normal		
Proporción de fallas (POF) (valor-p)	2,30 (0,13)	0,10 (0,75)	0,39 (0,53)	4,10 (0,04)	0,10 (0,75)	1,78 (0,18)
Independencia (valor-p)	0,38 (0,54)	0,07 (0,79)	0,03 (0,86)	1,19 (0,28)	0,07 (0,79)	0,07 (0,79)
POF e independencia (valor-p)	2,69 (0,26)	0,17 (0,92)	0,42 (0,81)	5,29 (0,07)	0,17 (0,92)	1,85 (0,40)

TABLA 6.3. Resultados obtenidos con la cópula t y la cópula normal (Muestra: 2006/05/19-2007/05/31)

Función f	Cópula t	Cópula Normal	RiskMetrics ®	Normalidad
	$L_{t+1 t} (\alpha = 0,95)$			
f_L	23,0	25,0	43,2	43,2
f_1	9,3	8,7	15,7	15,7
f_2	7,0	6,7	29,5	29,5
f_3	7,2	7,1	16,3	16,3
	$L_{t+1 t} (\alpha = 0,99)$			
f_L	4,4	4,7	27,8	27,8
f_1	1,8	2,1	7,3	7,3
f_2	1,7	2,2	12,2	12,2
f_3	1,6	1,6	10,6	10,6
	$L_{t+1 t} (\alpha = 0,995)$			
f_L	3,1	4,3	21,6	21,6
f_1	1,2	1,4	5,6	5,6
f_2	1,1	1,5	8,9	8,9
f_3	1,2	1,3	9,1	9,1

TABLA 6.4. Funciones de Pérdida para un período de tenencia de un día (Muestra: 2006/05/19-2007/05/31)

la Sección 5.1.2 (ecuación (5.20)), se encuentra que para todos los niveles de confianza, el ES_α obtenido por estas aproximaciones es adecuado (Tabla 6.5).

	Nivel de confianza		
	95 %	99 %	99,5 %
Cópula t	0,445	0,450	0,432
Cópula normal	0,465	0,424	0,469

TABLA 6.5. Valores-p obtenidos en las pruebas de *backtesting* para el ES_α con un período de tenencia de un día (Muestra: 2006/05/19-2007/05/31)

6.2. Otras aplicaciones del VaR_α . Además de brindar información sobre el riesgo subyacente de un portafolio, es importante analizar la sensibilidad de las medidas de riesgo ante cambios en las posiciones de los activos, ya que esto resulta de gran valor para una apropiada administración de riesgo. Con este fin, esta sección presenta dos medidas de sensibilidad del VaR_α a cambios en las posiciones de los activos, el VaR_α marginal y el VaR_α diferencial.

El VaR marginal se define como la cantidad de riesgo que la posición del activo i adiciona al portafolio P . Para un nivel de confianza α se tiene la siguiente expresión:

$$VaRM = VaR_{t+k|t}^{P-p_i}(\alpha) - VaR_{t+k|t}^P(\alpha)$$

En este caso se supone que el dinero obtenido por la liquidación del activo i se distribuye proporcionalmente dentro de los demás activos que conforman el portafolio. Si $VaRM > 0$, entonces liquidar el activo i aumenta el riesgo del portafolio en $VaRM$ unidades monetarias.

No obstante, el VaR_α marginal ofrece información sobre el cambio en el riesgo del portafolio cuando la posición del activo i es liquidada, lo que si bien es informativo, puede resultar extremo en la mayoría de los casos. Una alternativa a esta medida se encuentra en el VaR_α diferencial, el cual indica el cambio en el VaR del portafolio P ante un aumento en la posición en el activo i de $\delta \times 100\%$. Formalmente:

$$VaRD = VaR_{t+k|t}^{P+\delta p_i}(\alpha) - VaR_{t+k|t}^P(\alpha)$$

Nótese que $VaRM = VaRD$ si $\delta = -1$. Cuando $VaRD > 0$, aumentar en $\delta \times 100\%$ la posición del activo i aumenta el riesgo del portafolio en $VaRD$ unidades monetarias. Estas medidas alternativas se pueden obtener a partir del VaR_α calculado utilizando el procedimiento descrito en la Sección 5.1.1 y el VaR_α evaluado en las ponderaciones propuestas en cada caso.

Los resultados obtenidos para el VaR_α marginal y el VaR_α diferencial reportados para el período de desempeño del VaR_α son presentados en la Tabla 6.6, utilizando un nivel de confianza de $\alpha = 0,95$ y $\delta = 0,05$. Los resultados indican que dentro del portafolio, el

activo más riesgoso es el índice accionario, mientras que la tasa de cambio y el título de deuda soberana muestran comportamientos similares.

Por ejemplo, si se tiene un portafolio de 100 millones de pesos, para el período comprendido entre el 4 de agosto y el 17 de octubre de 2006, el resultado del VaR_{95} marginal implica que liquidar la posición mantenida en el índice accionario llevaría una reducción del valor en riesgo del portafolio de 241 mil pesos, mientras que en los otros dos casos, se tendría un incremento en el riesgo del portafolio de 336 mil pesos y 214 mil pesos, respectivamente. Consistentemente, el VaR_{95} diferencial sugiere que incrementar la posición en tasa de cambio y títulos de deuda contribuiría a reducir el valor en riesgo del portafolio. Al igual que en las aplicaciones anteriores, el VaR_{α} marginal y el VaR_{α} diferencial cambian en el tiempo (son más altas en períodos más volátiles), destacando la importancia de actualizar las medidas de riesgo del portafolio.

Período		VaR marginal (%)			VaR diferencial (%)		
		IGBC	TRM	TES	IGBC	TRM	TES
19-May-06	03-Ago-06	-0,929	1,237	0,658	0,0734	-0,0767	0,0474
04-Ago-06	17-Oct-06	-0,241	0,336	0,214	0,0093	-0,0051	-0,0131
18-Oct-06	03-Ene-07	-0,269	0,276	0,205	0,0162	-0,0302	-0,0381
04-Ene-07	15-Mar-07	-0,226	0,363	0,259	0,0246	-0,0151	-0,0088
16-Mar-07	31-May-07	-0,265	0,315	0,332	0,0189	-0,0111	-0,0194

TABLA 6.6. Mediana del VaR_{α} marginal y VaR_{α} diferencial con $\alpha = 0,95$, $\delta = 0,05$ y $k = 1$ día (Muestra: 2006/05/19-2007/05/31)

7. COMENTARIOS FINALES

En este documento, se presentaron la definición, implicaciones y metodologías para el desarrollo de modelos de medición de riesgo asociados con cópulas, las cuales, en líneas generales, son funciones que aproximan el comportamiento conjunto de variables aleatorias, a partir de sus comportamientos individuales (marginales). Las cópulas han ganado gran importancia en la última década, especialmente en el campo de las finanzas, ya que el conocimiento del comportamiento conjunto de diferentes factores de riesgo resulta de vital importancia para la correcta asignación de riesgos, por ejemplo, riesgos de mercado y de crédito. Para ilustrar estas metodologías, la última sección de este documento se dedica a analizar un portafolio compuesto por tres activos del mercado colombiano: el índice general de la bolsa de valores (IGBC), la tasa de cambio (TRM) y un título genérico con vencimiento a 10 años (TES).

En particular, al evaluar el valor en riesgo (VaR_{α}) derivado de la cópula y compararlas con el desempeño de otras metodologías comúnmente aplicadas para el cálculo de esta medida, los resultados son contundentes. En primer lugar, las técnicas que tienen en

cuenta que los momentos condicionales de los activos (media y varianza) no son constantes en el tiempo, muestran un desempeño superior con respecto a las metodologías que no consideran este tipo de dinámicas.

En segundo lugar, los modelos de medición de riesgo que realizan supuestos más realistas sobre las funciones de distribución marginales de las series (por ejemplo, aquellos que se alejan del supuesto de normalidad de cada uno de los retornos de los activos), si bien pueden ser más costosos en términos computacionales, cuantifican de mejor manera el riesgo del portafolio. En este sentido, los modelos derivados de la cópula ofrecen una estructura analítica flexible que resulta apropiada para la medición de riesgo.

Aunque este documento presenta buena parte de la teoría básica y aplicaciones de las cópulas, un análisis más detallado de cada sección se encuentra en diferentes textos especializados, por ejemplo, Nelsen [2006], McNeil et al. [2005] y Cherubini et al. [2004], entre otros. Finalmente, los códigos utilizados a lo largo del texto se basan en el programa de libre distribución R y sus librerías dedicadas a análisis financiero (disponibles en www.cran.r-project.org).

REFERENCIAS

- ACERBI, C., Y D. TASCHE (2002): "On the Coherence of Expected Shortfall," *Journal of Banking and Finance*, 26, 1487–1503.
- ANDERSON, T., Y D. DARLING (1952): "Asymptotic theory of certain goodness of fit criteria based on stochastic processes," *Annals of Mathematical Statistics*, 23, 193–212.
- ANÉ, T., Y C. KHAROUBI (2003): "Dependence structure and risk measure," *The Journal of Business*, 76, 411–38.
- ARTZNER, P., F. DELBAEN, J.-M. EBER, Y D. HEATH (1999): "Coherent Measures of Risk," *Mathematical Finance*, 9, 203–228.
- BALKEMA, A., Y L. DE HAAN (1974): "Residual Life Time at Great Age," *The Annals of Probability*, 2, 792–804.
- BLANCO, L. (2004): *Probabilidad*. Unibiblos.
- BOUYÉ, E., V. DURRLEMAN, A. ÑIKEGHBALI, G. RIBOULET, Y T. RONCALLI (2000): "Copulas for Finance - A Reading Guide and Some Applications," Discussion paper, available at SSRN: <http://ssrn.com/paper=1032533>.
- BRADLEY, B., Y M. TAQQU (2002): "Financial Risk and Heavy Tails," in *Handbook of Heavy Tailed Distributions in Finance*, ed. by S. Rachev. Elsevier.
- CAPÉRAÀ, P., A. FOURGÈRES, Y C. GENEST (2000): "Bivariate Distributions with Given Extreme Value Attractor," *Journal of Multivariate Analysis*, 72, 30–49.
- CAPORIN, M. (2003): "Evaluating Value-at-Risk Measures in Presence of Long Memory Conditional Volatility," Discussion Paper 0305, GRETA.
- CARMONA, R. (2004): *Statistical Analysis of Financial Data in S-Plus*. Springer-Verlag, Nueva York.
- CEBRIAN, A., M. DENUIT, Y O. SCAILLET (2002): "Testing for Concordance Ordering," FAME Research Paper Series rp41, International Center for Financial Asset Management and Engineering, available at <http://ideas.repec.org/p/fam/rpseri/rp41.html>.
- CHEN, X., Y Y. FAN (2005): "Estimation and Model Selection of Semiparametric Copula-Based Multivariate Dynamic Models under Copula Misspecification," *Journal of Econometrics*, 135, 125–154.
- CHERUBINI, U., E. LUCIANO, Y W. VECCHIATO (2004): *Copula Methods in Finance*. Wiley.
- CHOLLETE, L., Y A. HEINEN (2006): "Frequent Turbulence? A Dynamic Copula Approach," Discussion Paper 10, Department of Finance and Management Science, Norwegian School of Economics and Business Administration.
- CHRISTOFFERSEN, P. (2003): *Elements of Financial Risk Management*. Academic Press, San Diego, CA.
- CREDIT SUISSE FINANCIAL PRODUCTS (1997): "Creditrisk+: A Credit Risk Management Framework," Discussion paper, available from www.csfb.com/creditrisk.
- CSÖRGÖ, M., Y L. HORVÁTH (1997): *Limit Theorems in Change-Point Analysis*. Wiley.
- DEHEUVELS, P. (1978): "Caractérisation complète des Lois Extrêmes Multivariées et de la Convergence des Types Extrêmes," *Publications de l'Institut de Statistique de l'Université de Paris*, 23, 1–36.
- (1979): "La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance," *Académie Royale de Belgique, Bulletin de la Classe des*

- Sciences*, 65, 274–292.
- (1981): “A Nonparametric Test for Independence,” *Publications de l’Institut de Statistique de l’Université de Paris*, 26, 29–50.
- DELIANEDIS, G., Y R. GESKE (1998): “Credit Risk and Risk Neutral Default Probabilities: Information About Migrations and Defaults,” Discussion Paper 1114, Anderson Graduate School of Management, UCLA, available at <http://ideas.repec.org/p/cdl/anderf/1114.html>.
- DEMARTA, S. (2002): “Multivariate Extreme Value Theory and Copulas,” Master’s thesis, ETH Zurich.
- DEMARTA, S., Y A. MCNEIL (2005): “The t Copula and Related Copulas,” *International Statistical Review*, 73, 111–129.
- DI CLEMENTE, A., Y C. ROMANO (2004): “Measuring and Optimizing Portfolio Credit Risk: A Copula-based Approach,” *Economic Notes*, 33, 325–357.
- DIAS, A., Y P. EMBRECHTS (2004): “Dynamic Copula Models for Multivariate High-Frequency Data in Finance,” Discussion paper, available at <http://www.math.ethz.ch/baltes/ftp/papers.html>.
- DIEBOLD, F., T. GUNTHER, Y A. TAY (1998): “Evaluating Density Forecasts with Applications to Financial Risk Management,” *International Economic Review*, 39, 863–883.
- DIEBOLD, F., J. HAHN, Y A. TAY (1999): “Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High-Frequency Returns on Foreign Exchange,” *Review of Economics and Statistics*, 81, 661–673.
- EFRON, B., Y R. TIBSHIRANI (1993): *An Introduction to the Bootstrap*. Chapman & Hall.
- EMBRECHTS, P., C. KLUPPELBERG, Y T. MIKOSCH (1997): *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.
- EMBRECHTS, P., A. MCNEIL, Y D. STRAUMANN (1999): “Correlation: Pitfalls and Alternatives,” Discussion paper, ETH Zurich.
- (1999a): “Correlation and Dependency in Risk Management: Properties and Pitfalls,” *RISK Magazine*, pp. 69–71.
- ENGLE, R. (1982): “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of Uk Inflation,” *Econometrica*, 50, 987–1008.
- FERNÁNDEZ, V. (2003): “Extremal Dependence in exchange Markets,” Discussion Paper 165, Centro de Economía Aplicada, Universidad de Chile.
- FISHER, R. A., Y L. H. C. TIPPETT (1928): “Limiting Forms of the Frequency Distribution of the Largest of Smallest Member of a Sample,” *Proceedings of the Cambridge Philosophical Society*, 24, 180–190.
- FREES, E., Y E. VALDEZ (1998): “Understanding Relationships Using Copulas,” *North American Actuarial Journal*, 2, 1–25.
- FREY, R., A. MCNEIL, Y M. ÑYFELER (2001): “Copulas and Credit Models,” *Risk*, 14, 111–114.
- GENEST, C., Y L. RIVEST (1993): “Statistical Inference Procedures for Bivariate Archimedean Copulas,” *Journal of the American Statistical Association*, 88, 1034–1043.
- (2001): “On the Multivariate Probability Integral Transformation,” *Statistics & Probability Letters*, 53, 391–399.

- GHOUDI, K., A. KHOUDRAJI, Y L.-P. RIVEST (1998): "Propriétés Statistiques Des Copules De Valeurs Extrêmes Bidimensionnelles," *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 26, 187–197.
- GILES, D. (2000): "A saddlepoint approximation to the distribution function of the Anderson-Darling test statistics," Discussion Paper EWP 0005, University of Victoria.
- GNEDENKO, B. (1943): "Sur La Distribution Limite Du Terme Maximum D'une Serie Aleatoire," *Annals of Mathematics*, 44, 423–453.
- GOURIEROUX, C., Y J. JASIAK (2001): *Financial Econometrics: Problems, Models, and Methods*. Princeton University Press.
- HUBER, P. (2004): *Robust Statistics*. Wiley.
- JOE, H. (1997): *Multivariate Models and Dependence Concepts*. Chapman & Hall, Londres, UK.
- JP MORGAN & INCORPORATED (1997): "CreditMetrics Technical Document," Discussion paper, available from <http://www.jpmorgan.com>.
- (2000): "Guide to credit derivatives," *Risk*.
- JUNKER, M., Y A. MAY (2005): "Measurement of aggregate risk with copulas," *Econometrics Journal*, 8, 428–454.
- KIMBERLING, C. (1974): "A Probabilistic Interpretation of Complete Monotonicity," *Aequationes Math*, 10, 152–164.
- KLEIN, J., Y M. MOESCHBERGER (2003): *Survival Analysis: Techniques for Censored and Truncated Data*. Springer.
- KMV CORPORATION (2003): "Modelling Default Risk," Discussion paper, Technical Document, available from <http://www.kmv.com>.
- KUPIEC, P. (1995): "Techniques for Verifying the Accuracy of Risk Measurement Models," *Journal of Derivatives*, 3, 73–84.
- LI, D. (1998): "Constructing a Credit Curve," *Risk*, Special Report on Credit Risk, 40–44.
- (2000): "On Default Correlation: A Copula Function Approach," *Journal of Fixed Income*, 9, 43–54.
- LI, R., K. FANG, Y L. ZHU (1997): "Some Q-Q Probability Plots to Test Spherical and Elliptical Symmetry," *Journal of Computational and Graphical Statistics*, 6, 435–450.
- LINDSKOG, F. (2000): "Linear Correlation Estimation," Discussion paper, RiskLab Switzerland.
- LIU, W. (2006): "Currencies Portfolio Return: A Copula Methodology," Discussion paper, Universidad de Toronto, Department of Economics.
- LONGIN, F., Y B. SOLNIK (2001): "Extreme Correlation of International Equity Markets," *Journal of Finance*, 56, 649–676.
- LÓPEZ, J. A. (1998): "Methods for Evaluating Value-at-Risk Estimates," *Federal Reserve Bank of New York Economic Policy Review*, 3, 119–124.
- MALEVERGNE, Y., Y D. SORNETTE (2006): *Extreme Financial Risks: From Dependence to Risk Management*. Springer.
- MARKOWITZ, H. (1952): "Portfolio Selection," *The Journal of Finance*, 7, 77–91.
- MARSHALL, A., Y I. OLKIN (1988): "Families of Multivariate Distributions," *Journal of the American Statistical Association*, 83, 834–841.

- MASHAL, R., Y A. ZEEVI (2002): "Beyond Correlation: Extreme Co-Movements between Financial Assets," Discussion paper, Columbia Business School.
- MCNEIL, A. (1999): "Extreme Value Theory for Risk Managers," in *Internal Modelling and Cad II*, pp. 93–113. RISK Books.
- MCNEIL, A., Y R. FREY (2000): "Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach," *Journal of Empirical Finance*, 7, 271–300.
- MCNEIL, A., R. FREY, Y P. EMBRECHTS (2005): *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press.
- MELO, L. F., Y O. BECERRA (2006): *Medidas de Riesgo, Características y Técnicas de Medición: Una Aplicación del VaR y el ES a la Tasa Interbancaria de Colombia*. Universidad del Rosario.
- MENEGUZZO, D., Y W. VECCHIATO (2002): "Improvement on Value at Risk Measures by Combining Conditional Autoregressive and Extreme Value Approaches," *Review in International Business and Finance*, Special Issue on Financial Risk and Financial Risk Management, 16, 275–324.
- MICOCCI, M., Y G. MASALA (2004): "Backtesting Value-at-Risk Estimation with Non Gaussian Marginals," Discussion paper, University of Cagliari, Faculty of Economics.
- MIKOSCH, T. (2006): "Copulas: Tales and Facts," *Extremes*, 9, 3–20.
- NELSEN, R. (2006): *An Introduction to Copulas*. Springer, Nueva York.
- NOLAN, J. (2007): *Stable Distributions: Models for Heavy Tailed Data*. Birkhauser, Boston, In progress, Chapter 1 online at academic2.american.edu/~jpnolan.
- NYFELER, M. (2000): "Modelling Dependencies in Credit Risk Management," Diploma Thesis, ETH Zurich.
- PATTON, A. (2001): "Modelling Time-Varying Exchange Rate Dependence Using the Conditional Copula," Discussion Paper 17, University of California, Economics.
- (2006): "Modelling Asymmetric Exchange Rate Dependence," *International Economic Review*, 47, 527–556.
- PICKANDS, J. (1975): "Statistical Inference Using Extreme Order Statistics," *The Annals of Statistics*, 3, 119–131.
- (1981): "Multivariate Extreme Value Distributions," in *43rd Session of the International Statistical Institute*, Book 2, pp. 859–878.
- POON, S., M. ROCKINGER, Y J. TAWN (2001): "New Extreme-Value Dependence Measures and Finance Applications," Discussion Paper 2762, CEPR.
- RACHEV, S., S. ORTOBELLI, S. STOYANOV, F. FABOZZI, Y A. BIGLOVA (2005): "The Proper Use of Risk Measures in Portfolio Theory," *International Journal of Theoretical and Applied Finance*, pp. 1107–1133.
- RODRÍGUEZ, J. C. (2007): "Measuring Financial Contagion: A Copula Approach," *Journal of Empirical Finance*, 14, 401–423.
- ROMANO, C. (2002): "Calibrating and Simulating Copula Functions: An Application to the Italian Stock Market," Discussion Paper 12/2002, Centro Interdipartimentale sul Diritto e l'Economia dei Mercati.

-
- ROUSSEEUW, P., Y G. MOLENBERGHS (1993): "Transformation of Non Positive Semidefinite Correlation Matrices," *Communications in Statistics, Theory and Methods*, 22, 965–984.
- SKLAR, A. (1959): "Fonctions de Répartition à N Dimensions et Leurs Marges," *Publications de l'Institut de Statistique de l'Université de Paris*, 8, 229–231.
- VOSTRIKOVA, L. (1981): "Detecting Disorder in Multidimensional Random Process," *Soviet Mathematics Doklady*, 24, 55–59.
- WHITE, H. (1999): *Asymptotic Theory for Econometricians*. Academic Press Inc.
- ZIVOT, E., Y J. WANG (2006): *Modeling Financial Time Series with S-Plus®*. Springer.

ANEXO A. DISTRIBUCIONES ESFÉRICAS Y ELÍPTICAS

La distribución normal multivariada pertenece a una clase de distribuciones más amplia, la cual es conocida como la familia de distribuciones elípticas. Estas distribuciones son muy importantes en el área de finanzas ya que cuando un vector de variables aleatorias sigue una distribución elíptica, las metodologías estándar de administración de riesgo son válidas. Bajo este tipo de distribuciones el VaR_α es una medida de riesgo coherente y la selección óptima de portafolios lineales basada en la metodología media-varianza de Markowitz es apropiada.

Adicionalmente, las distribuciones elípticas pueden ser definidas en términos de otra familia de distribuciones, las distribuciones esféricas. Un ejemplo clásico de distribuciones esféricas es $N_d(0, I)$, mientras que uno para distribuciones elípticas es $N_d(\mu, \Sigma)$. Otros ejemplos interesantes de estos dos tipos de familias de distribuciones incluyen distribuciones de colas pesadas.

A.1. Distribuciones esféricas. Un vector de variables aleatorias $R = (R_1, R_2, \dots, R_d)'$ sigue una distribución esférica, $S(\psi)$, si para toda matriz ortogonal $U \in \mathbb{R}^{d \times d}$ (i.e. $UU' = U'U = I_d$) se tiene que:

$$UR \stackrel{d}{=} dR \quad (\text{A.1})$$

La definición (A.1) indica que la distribución de R es invariante ante rotaciones en sus coordenadas. Adicionalmente, la función característica ϕ_R ⁴⁷ asociada al vector aleatorio R satisface la siguiente relación:

$$\phi_R(t) = E\left(e^{it'R}\right) = E\left(e^{it'UR}\right) = \phi_R(U't)$$

La anterior definición implica que existe una función $\psi : \mathbb{R}^{d \times 1} \rightarrow \mathbb{R}$ tal que:

$$\phi_R(t) = \psi(t't) \quad (\text{A.2})$$

puesto que $\phi_R(t) = \psi(t't) = \psi(t'UU't) = \phi_R(U't)$. La función escalar $\psi(\cdot)$ es denominada el generador característico de la distribución esférica $S(\psi)$, $R \sim S(\psi)$ ⁴⁸.

Como se ha mencionado anteriormente, un ejemplo de una distribución esférica es la distribución normal estándar multivariada. De acuerdo con el resultado (A.2) una forma de comprobar que ésta distribución pertenece a esta familia es a través de su función característica. Para esta distribución se tiene que $\phi_R(t) = e^{-\frac{1}{2}t't}$; por lo tanto, el generador característico es $\psi(s) = e^{-\frac{1}{2}s}$. En general, las distribuciones esféricas corresponden a

⁴⁷La función característica de un vector aleatorio R , ϕ_R , se define como $\phi_R(t) = E\left(e^{it'R}\right)$ donde $i^2 = -1$ y $t \in \mathbb{R}^{d \times 1}$. Para una discusión de la definición de ϕ_R y sus propiedades véase White [1999].

⁴⁸Además de la definición (A.1), existe otra forma de caracterizar a una distribución esférica. Un vector aleatorio R tiene una distribución esférica si y solo si éste se puede representar estocásticamente como $R \stackrel{d}{=} TS$; donde T es una variable aleatoria no negativa y S es un vector aleatorio independiente de T y distribuido uniformemente sobre la hipersfera unitaria $S_d = \{s \in \mathbb{R}^{d \times 1} \mid \|s\| = 1\}$.

distribuciones de variables aleatorias no correlacionadas. Sin embargo, dentro de las distribuciones esféricas, $N_d(0, I)$ es la única distribución compuesta por variables aleatorias independientes [Lindskog, 2000].

A.2. Distribuciones elípticas. En general, una distribución esférica se puede ver como una extensión de la distribución normal multivariada estándar $N_d(0, I)$, mientras que una distribución elíptica se puede entender como una extensión de $N_d(\mu, \Sigma)$. Para entender la definición de este último tipo de distribuciones es importante recordar que, de la misma forma que en el caso univariado, existe una relación entre un vector aleatorio $R \sim N_d(\mu, \Sigma)$ y uno normal estándar $Y \sim N_k(0, I)$, como se muestra a continuación:

$$R \stackrel{d}{=} \mu + AY$$

donde $\Sigma = AA'$.

Si R sigue una distribución elíptica, entonces R satisface la siguiente condición:

$$R \stackrel{d}{=} \mu + AY \tag{A.3}$$

donde Y es un vector aleatorio de dimensión k que sigue una distribución esférica $S(\psi)$, $A \in \mathbb{R}^{d \times k}$ y $\mu \in \mathbb{R}^d$. Es decir, una distribución elíptica está completamente descrita por su vector de medias (μ), su matriz de covarianzas ($\Sigma = AA'$) y el generador característico (ψ). Si R sigue una distribución elíptica se nota $R \sim E_d(\mu, \Sigma, \psi)$.

Dado que $N_d(0, I)$ sigue una distribución esférica, de acuerdo con la definición (A.3) una función de distribución normal multivariada $N_d(\mu, \Sigma)$ es elíptica con $\Sigma = AA'$. Otras distribuciones esféricas o elípticas son la distribución t multivariada, la logística multivariada, las distribuciones multivariadas de escala mixta (scale mixture) y las distribuciones de leyes estables (stable laws) [Bradley y Taqqu, 2002].

En el Gráfico A.1 se muestran las funciones de densidad y las curvas de nivel de una distribución esférica y una elíptica. Ambos casos corresponden a una distribución t bivariada, $t_2(v_i, \mu_i, \Sigma_i)$, $i = 1, 2$, con los siguientes parámetros $v_1 = v_2 = 4$, $\mu'_1 = \mu'_2 = (0, 0)$, $\Sigma_1 = I_2$ y $\Sigma_2 = \begin{bmatrix} 1 & 0,8 \\ 0,8 & 1 \end{bmatrix}$. En estos gráficos se puede apreciar que las curvas de nivel de las funciones de densidad tienen formas circulares o esféricas y elípticas, lo que constituye una forma intuitiva de reconocer a que familia pertenecen algunos tipos de distribuciones.

Una de las ventajas de las distribuciones elípticas es que éstas comparten varias de las propiedades de la distribución normal multivariada: *i*) Cualquier combinación lineal de vectores aleatorios con distribuciones elípticas también es elíptica; *ii*) Las distribuciones marginales de distribuciones elípticas son elípticas; *iii*) Si $R \sim E_d(\mu, \Sigma, \psi)$, $R' = [R'_1, R'_2]$ y Σ es una matriz positiva definida, entonces la distribución de R_1 condicionada a R_2 también es elíptica, [Embrechts et al., 1999].

Desde el punto de vista de riesgo financiero algunas de las propiedades favorables de las distribuciones elípticas son las siguientes: *i*) Para cualquier portafolio lineal con factores

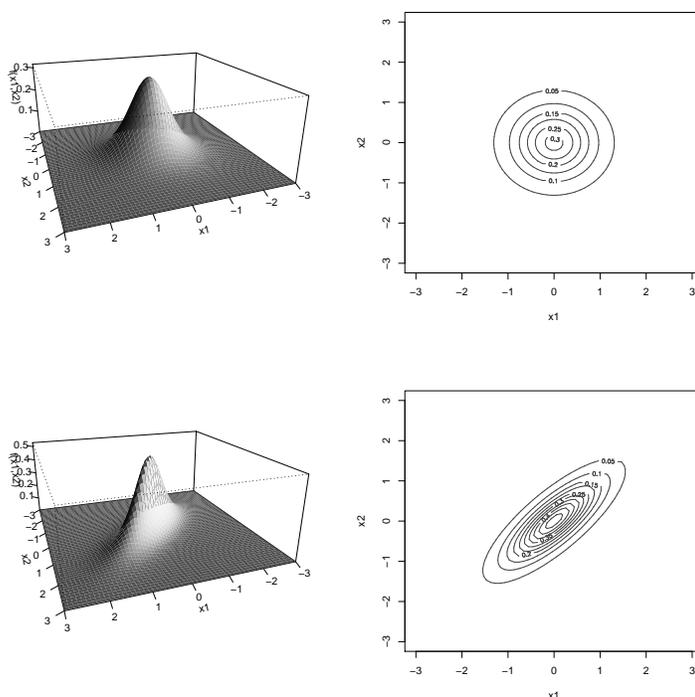


GRÁFICO A.1. Funciones de densidad para dos variables aleatorias t bivariadas, $t_2(\nu_1, \mu_1, \Sigma_1)$ y $t_2(\nu_2, \mu_2, \Sigma_2)$ con $\nu_1 = \nu_2 = 4$, $\mu'_1 = \mu'_2 = (0, 0)$, $\Sigma_1 = I_2$ y $\Sigma_2 = [1 \ 0, 8; 0, 8 \ 1]$

de riesgo elípticos se tiene que el VaR_α es una medida de riesgo coherente; *ii*) Si los factores de riesgo son elípticos entonces la correlación lineal es una medida de dependencia adecuada; *iii*) Para factores de riesgo elípticos, la selección del portafolio óptimo basada en la metodología de media-varianza de Markowitz para un nivel dado de retorno esperado es el mismo sin importar la medida de riesgo utilizada: ya sea la varianza, el VaR (valor en riesgo), el ES (Expected Shortfall) o cualquier otra medida de riesgo invariante ante traslaciones y homogénea positiva⁴⁹.

Los anteriores resultados implican que en términos de manejo de riesgo es muy relevante verificar si la distribución de los factores de riesgo utilizados tiene una distribución elíptica. A continuación se introducen algunas metodologías para identificar si la distribución de un vector aleatorio R es una distribución elíptica.

A.3. Pruebas para distribuciones elípticas. Para contrastar si una muestra de n vectores aleatorios *i.i.d.* de dimensión d , R_1, R_2, \dots, R_n siguen una distribución elíptica, Li

⁴⁹Para una demostración completa de estas propiedades véase Bradley y Taqqu [2002].

et al. [1997] proponen un método basado en los gráficos Q-Q⁵⁰, el cual se fundamenta en el siguiente lema:

Lema A.1. Sea $T(R)$ una estadística tal que, casi seguramente (a.s.):

$$T(aR) = T(R) \text{ para todo } a > 0 \quad (\text{A.4})$$

entonces $T(R)$ tiene la misma distribución para todo vector esférico $R \sim S_d^+(\psi)$ ⁵¹.

A partir del resultado anterior, Li et al. [1997] proponen utilizar dos estadísticas para $R_i \stackrel{i.i.d.}{\sim} N_d(0, I_d)$, $i = 1, 2, \dots, n$ que satisfacen la condición (A.4). Dichas estadísticas son:

$$T_1(R_i) = \frac{d^{1/2} \bar{R}_i}{\sqrt{(1/(d-1)) \sum_{j=1}^d (R_{i,j} - \bar{R}_i)^2}}, \quad \bar{R}_i = \frac{1}{d} \sum_{j=1}^d R_{i,j}$$

$$T_2(R_i) = \frac{\sum_{j=1}^k R_{i,j}^2}{\sum_{j=1}^d R_{i,j}^2}$$

donde $R'_i = (R_{i,1}, \dots, R_{i,d})$, $T_1(R_i) \sim t(d-1)$ y $T_2(R_i) \sim \text{Beta}(k/2, (d-k)/2)$. En esta prueba, el gráfico Q-Q compara los cuantiles empíricos de $T_1(R)$ y $T_2(R)$ con los cuantiles teóricos de una distribución $t(d-1)$ y $\text{Beta}(k/2, (d-k)/2)$ ⁵², respectivamente. Si estos gráficos toman la forma de una línea recta, indica evidencia a favor de la hipótesis elíptica.

Es importante anotar que aunque se desea verificar que R sigue una distribución elíptica, el Lema A.1 utilizado por Li et al. [1997] está basado en distribuciones esféricas. Por lo tanto, las estadísticas T_1 y T_2 deben ser aplicadas a una transformación R^* de R . Si R sigue una distribución elíptica, $E_d(\mu, \Sigma, \psi)$, la transformación R^* es consistente con una distribución esférica y toma la siguiente forma:

$$R_i^* = \widehat{\Sigma}^{-1/2} (R_i - \widehat{\mu}), \quad i = 1, 2, \dots, n$$

donde $\widehat{\Sigma}$ y $\widehat{\mu}$ son estimaciones de Σ y μ , respectivamente⁵³ y $\widehat{\Sigma}^{-1/2}$ corresponde al factor de la descomposición de Cholesky de $\widehat{\Sigma}$.

⁵⁰Los gráficos Q-Q, o cuantil - cuantil, comparan los cuantiles de dos distribuciones. Si las dos distribuciones son iguales se espera una línea recta. Adicional a los gráficos Q-Q, existen pruebas de carácter numérico que permiten evaluar si un vector aleatorio sigue una distribución elíptica. McNeil et al. [2005] referencia algunas de ellas.

⁵¹ $S_d^+(\psi)$ representa una subclase de distribuciones esféricas en las cuales $P(R=0) = 0$. Esta definición es utilizada ya que dentro de las distribuciones esféricas frecuentemente se excluyen las que tienen concentrada su masa en el origen.

⁵²Li et al. [1997] proponen utilizar $k = \text{int}(d/2)$.

⁵³En este contexto es conveniente utilizar estimadores robustos a datos atípicos, debido a que la mayoría de las series financieras tienen distribuciones de colas pesadas. Para tal efecto McNeil et al. [2005] recomiendan el uso de estimadores M. Para una discusión detallada de este tipo de estimadores, véase Huber [2004].

En el Gráfico A.2 se muestran los gráficos Q-Q asociados a las estadísticas T_1 y T_2 , para 5000 datos simulados de una distribución t multivariada de dimensión 10 con $v = 4$ grados de libertad. La ausencia de curvatura en estas gráficas se puede ver como una evidencia a favor de la hipótesis de que los datos siguen una distribución elíptica. Mc-Neil et al. [2005] sugiere que en la mayoría de casos empíricos la estadística T_2 es más apropiada para el gráfico Q-Q.

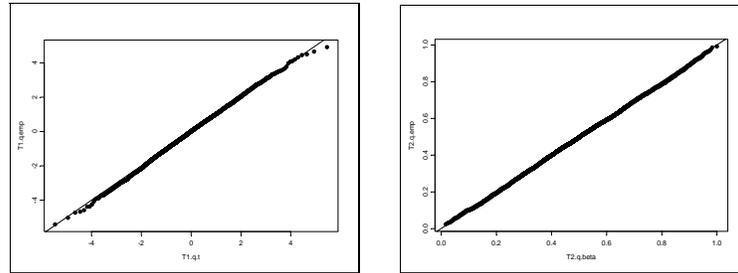


GRÁFICO A.2. Gráficos Q-Q para las estadísticas T_1 (panel izquierdo) y T_2 (panel derecho) calculadas para 5,000 realizaciones de un vector aleatorio t de dimensión $d = 10$ y con $v = 4$ grados de libertad

ANEXO B. PRUEBAS DE ESPECIFICACIÓN SOBRE EL MODELO VAR-GARCH

Prueba	Hipótesis Nula	Rezagos	Estadística	P-Value
Pruebas univariadas sobre el residuo de la ecuación de $\Delta LIGBC$				
Breusch y Godfrey (LM)	No existe autocorrelación	4	3,07	0,55
Ljung-Box	No existe autocorrelación	281	278,1	0,37
Engle	No hay efecto ARCH	12	2,42	0,99
Pruebas univariadas sobre el residuo de la ecuación de $\Delta LTRM$				
Breusch y Godfrey (LM)	No existe autocorrelación	4	6,32	0,18
Ljung-Box	No existe autocorrelación	281	254,3	0,83
Engle	No hay efecto ARCH	12	5,45	0,94
Pruebas univariadas sobre el residuo de la ecuación de $\Delta LTES10$				
Breusch y Godfrey (LM)	No existe autocorrelación	4	8,82	0,07
Ljung-Box	No existe autocorrelación	281	265,60	0,72
Engle	No hay efecto ARCH	12	4,30	0,98

TABLA B.1. Pruebas de especificación univariadas sobre los residuales estandarizados del modelo $VAR(10) - GARCH(1,1)$

Pruebas multivariadas				
Prueba	Hipótesis Nula	Rezagos	Estadística	P-Value
Breusch y Godfrey (LM)	No existe autocorrelación	4	50,82	0,05
Portmanteau	No existe autocorrelación	281	2249,1	0,99
Portmanteau (ajustado)	No existe autocorrelación	281	2564,8	0,04
Portmanteau sobre el cuadrado de los residuales	No hay efecto MGARCH	12	472,32	0,09

TABLA B.2. Pruebas de especificación multivariadas sobre los residuales estandarizados del modelo $VAR(10) - GARCH(1,1)$

ANEXO C. CÓPULAS COMÚNMENTE MENCIONADAS EN LA LITERATURA

Cópula	Fórmula	Parámetros
Normal	$\int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\theta^2}} \exp\left(-\frac{x^2-2\theta xy+y^2}{2(1-\theta^2)}\right) dx dy$ $\Phi^{-1}(\cdot)$: Inversa de la función de distribución normal estándar	$\theta \in [-1, 1]$
t	$\int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\theta^2}} \left(1 + \frac{x^2-2\theta xy+y^2}{\nu(1-\theta^2)}\right)^{-(\nu+2)/2} dx dy$ $t_\nu^{-1}(\cdot)$: Inversa de la función de distribución t con ν g.l.	$\theta \in [-1, 1]$ ν : grados de lib.
Clayton	$\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-\frac{1}{\theta}}$	$\theta \in (0, \infty)$
Clayton generalizada	$\left\{ \left[\left(u_1^{-\theta} - 1\right)^\delta + \left(u_2^{-\theta} - 1\right)^\delta \right]^{1/\delta} + 1 \right\}^{-\frac{1}{\theta}}$	$\theta \in (0, \infty)$ $\delta \in [1, \infty)$
Frank	$-\frac{1}{\theta} \log \{ [\eta - (1 - \exp(-\theta u_1)) (1 - \exp(-\theta u_2))] / \eta \}$	$\theta \in (0, \infty)$ $\eta = 1 - e^{-\theta}$
Galambos	$u_1 u_2 \exp \left\{ [(-\log(u_1))^{-\theta} + (-\log(u_2))^{-\theta}]^{-1/\theta} \right\}$	$\theta \in [0, \infty)$
Gumbel	$\exp \left\{ - [(-\log(u_1))^\theta + (-\log(u_2))^\theta]^{1/\theta} \right\}$	$\theta \in [1, \infty)$
Gumbel II	$u_1 u_2 \exp \left[\frac{\theta \log(u_1) \log(u_2)}{-(\log(u_1) + \log(u_2))} \right]$	$\theta \in [0, 1]$
Husler y Reiss	$\exp \left\{ \log(u_1) \Phi \left[\frac{1}{\theta} + \frac{\theta}{2} \log \left(\frac{\log(u_1)}{\log(u_2)} \right) \right] + \log(u_2) \Phi \left[\frac{1}{\theta} + \frac{\theta}{2} \log \left(\frac{\log(u_2)}{\log(u_1)} \right) \right] \right\}$	$\theta \in [0, \infty)$
Joe Marshall-Olkin	$1 - [(1 - u_1)^\theta + (1 - u_2)^\theta - (1 - u_1)^\theta (1 - u_2)^\theta]^{1/\theta}$ $u_1^{1-\theta_1} u_2^{1-\theta_2} \min(u_1^{\theta_1}, u_2^{\theta_2})$	$\theta \in [1, \infty)$ $\theta_1 \in [0, 1]$ $\theta_2 \in [0, 1]$
Morgens-tern	$u_1 u_2 [1 + \theta (1 - u_1) (1 - u_2)]$	$\theta \in [-1, 1]$
Plackett	$\frac{1}{2}(\theta - 1)^{-1} \left\{ 1 + (\theta - 1)(u_1 + u_2) - \left[(1 + (\theta - 1)(u_1 + u_2))^2 - 4\theta(\theta - 1)u_1 u_2 \right]^{\frac{1}{2}} \right\}$	$\theta \in [0, \infty)$
B11	$\theta \min(u_1, u_2) + (1 - \theta) u_1 u_2$	$\theta \in [0, 1]$
B12	$[\min(u_1, u_2)]^\theta (u_1 u_2)^{1-\theta}$	$\theta \in [0, 1]$

Continúa en la página siguiente

Cópula	Fórmula	Parámetros
BB1	$\left\{ 1 + \left[\left(u_1^{-\delta} - 1 \right)^\theta + \left(u_2^{-\delta} - 1 \right)^\theta \right]^{1/\theta} \right\}^{-1/\delta}$	$\theta \in [1, \infty)$ $\delta \in (0, \infty)$
BB2	$\left[1 + \theta^{-1} \log \left(\exp(\theta(u_1^{-\delta} - 1)) + \exp(\theta(u_2^{-\delta} - 1)) - 1 \right) \right]^{1/\delta}$	$\theta \in (0, \infty)$ $\delta \in (0, \infty)$
BB3	$\exp \left\{ - \left[\theta^{-1} \log \left(\exp \left[(-\log(u_1))^\delta \theta \right] + \exp \left[(-\log(u_2))^\delta \theta \right] - 1 \right) \right]^{1/\delta} \right\}$	$\theta \in (0, \infty)$ $\delta \in [1, \infty)$
BB4	$\left\{ u_1^{-\delta} + u_2^{-\delta} - 1 - \left[\left(u_1^{-\delta} - 1 \right)^{-\theta} + \left(u_2^{-\delta} - 1 \right)^{-\theta} \right]^{-1/\theta} \right\}^{-1/\delta}$	$\theta \in (0, \infty)$ $\delta \in [0, \infty)$
BB5	$\exp \left\{ - \left[(-\log(u_1))^\delta + (-\log(u_2))^\delta - \left((-\log(u_1))^{-\theta\delta} + (-\log(u_2))^{-\theta\delta} \right)^{-1/\theta} \right]^{1/\delta} \right\}$	$\theta \in (0, \infty)$ $\delta \in [1, \infty)$
BB6	$1 - \left\{ 1 - \exp \left\{ - \left[(-\log(1 - (1 - u_1)^\delta))^\theta + (-\log(1 - (1 - u_2)^\delta))^\theta \right]^{1/\theta} \right\} \right\}^{1/\delta}$	$\theta \in [1, \infty)$ $\delta \in [1, \infty)$
BB7	$1 - \left\{ 1 - \left[\left(1 - (1 - u_1)^\delta \right)^{-\theta} + \left(1 - (1 - u_2)^\delta \right)^{-\theta} - 1 \right]^{-1/\theta} \right\}^{1/\delta}$	$\theta \in (0, \infty)$ $\delta \in [1, \infty)$
BB8	$\theta^{-1} \left\{ 1 - \left\{ 1 - \left[1 - (1 - \theta)^\delta \right]^{-1} \left[1 - (1 - \theta u_1)^\delta \right] \left[1 - (1 - \theta u_2)^\delta \right] \right\}^{1/\delta} \right\}$	$\theta \in [0, 1]$ $\delta \in [1, \infty)$
BB9	$\exp \left\{ - \left[\left(\alpha - \log(u_1) \right)^\delta + \left(\alpha - \log(u_2) \right)^\delta - \alpha^\delta \right]^{1/\delta} + \alpha \right\}$	$\delta \in [1, \infty)$ $\alpha \in (0, \infty)$
BB10	$u_1 u_2 \left[1 - \delta \left(1 - u_1^{1/\alpha} \right) \left(1 - u_2^{1/\alpha} \right) \right]^{-\alpha}$	$\delta \in [0, 1]$ $\alpha \in (0, \infty)$

Tabla C.3: Algunas cópulas bivariadas

Cópula	Fórmula	Parámetros
Normal	$\frac{1}{\sqrt{(2\pi)^d \mathbf{P} }} \int_{-\infty}^{\Phi^{-1}(u_1)} \dots \int_{-\infty}^{\Phi^{-1}(u_d)} \exp\left(-\frac{1}{2} \mathbf{x}' \mathbf{P}^{-1} \mathbf{x}\right) d\mathbf{x}$ $\Phi^{-1}(\cdot)$: Inversa de la función de distribución normal estándar	\mathbf{P} es la matriz de correlaciones
t	$\frac{1}{\sqrt{ \mathbf{P} }} \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})(\pi\nu)^{d/2}} \int_{-\infty}^{t_\nu^{-1}(u_1)} \dots \int_{-\infty}^{t_\nu^{-1}(u_d)} \left(1 + \frac{\mathbf{x}' \mathbf{P}^{-1} \mathbf{x}}{\nu}\right)^{-\frac{\nu+d}{2}} d\mathbf{x}$ $t_\nu^{-1}(\cdot)$: Inversa de la función de distribución t con ν g.l. $\Gamma(\cdot)$: Función gama	\mathbf{P} es la matriz de correlaciones ν : grados de lib.
Clayton	$\left(\sum_{i=1}^d u_i^{-\theta} - d + 1\right)^{1/\theta}$	$\theta \in (0, \infty)$
Frank	$-\frac{1}{\theta} \log \left[1 + \frac{\prod_{i=1}^d (e^{-\theta u_i} - 1)}{(e^{-\theta} - 1)^d}\right]$	$\theta \in (0, \infty)$
Gumbel	$\exp \left\{ - \left[\sum_{i=1}^d (-\log(u_i))^\theta \right]^{1/\theta} \right\}$	$\theta \in (1, \infty)$
MM1	$\exp \left\{ - \left[\sum_{1 \leq i < j \leq d} \left((p_i \tilde{u}_i^\delta)^{\theta_{ij}} + (p_j \tilde{u}_j^\delta)^{\theta_{ij}} \right)^{\frac{1}{\theta_{ij}}} + \sum_{j=1}^d \nu_j p_j \tilde{u}_j^\delta \right]^{\frac{1}{\delta}} \right\}$ $p_j = (\nu_j + d - 1)^{-1}$, $\tilde{u}_j = -\log(u_j)$, $j = 1, \dots, d$	$\theta_{ij} \in [1, \infty)$ $\delta \in [1, \infty)$ $\nu_j \in [0, \infty)$
MM2	$\left[\sum_{j=1}^d u_j^{-\delta} - d + 1 - \sum_{1 \leq i < j \leq d} \left(\hat{u}_i^{-\theta_{ij}} + \hat{u}_j^{-\theta_{ij}} \right)^{-1/\theta_{ij}} \right]^{-1/\delta}$ $p_j = (\nu_j + d - 1)^{-1}$, $\hat{u}_j = p_j (u_j^{-\delta} - 1)$, $j = 1, \dots, d$	$\theta_{ij} \in (0, \infty)$ $\delta \in (0, \infty)$ $\nu_j \in [0, \infty)$
MM3	$\exp \left\{ - \left[\sum_{j=1}^d \tilde{u}_j^\delta - \sum_{1 \leq i < j \leq d} \left(p_i^{-\theta_{ij}} \tilde{u}_i^{-\delta \theta_{ij}} + p_j^{-\theta_{ij}} \tilde{u}_j^{-\delta \theta_{ij}} \right)^{-\frac{1}{\theta_{ij}}} \right]^{\frac{1}{\delta}} \right\}$ $p_j = (\nu_j + d - 1)^{-1}$, $\tilde{u}_j = -\log(u_j)$, $j = 1, \dots, d$	$\theta_{ij} \in (0, \infty)$ $\delta \in (0, \infty)$ $\nu_j \in [0, \infty)$

TABLA C.4. Algunas cópulas d -dimensionales

Cópula	$\phi(t)$	C^u	C^I	C^l
Clayton	$\frac{1}{\theta} (t^{-\theta} - 1)$	$\theta \rightarrow \infty$	$\theta \rightarrow 0$	—
Frank	$-\log \left(\frac{\exp(-\theta t) - 1}{\exp(-\theta) - 1} \right)$	$\theta \rightarrow \infty$	$\theta \rightarrow 0$	$\theta \rightarrow -\infty$
Gumbel	$(-\log(t))^\theta$	$\theta \rightarrow \infty$	$\theta = 1$	—
Joe	$-\log(1 - (1-t)^\theta)$	$\theta \rightarrow \infty$	$\theta = 1$	—
BB1	$(t^{-\delta} - 1)^\theta$	$\theta \rightarrow \infty$ o $\delta \rightarrow \infty$	$\theta \rightarrow 1$ y $\delta \rightarrow 0$	—
BB2	$\exp(\theta(t^{-\delta} - 1)) - 1$	$\theta \rightarrow \infty$ o $\delta \rightarrow \infty$	$\delta \rightarrow 0$	—
BB3	$\exp[\theta(-\log(t))^\delta] - 1$	$\theta \rightarrow \infty$ o $\delta \rightarrow \infty$	—	—
BB4	$t^{-\delta} - 1$	$\theta \rightarrow \infty$ o $\delta \rightarrow \infty$	—	—
BB6	$[-\log(1 - (1-t)^\delta)]^\theta$	$\theta \rightarrow \infty$ o $\delta \rightarrow \infty$	—	—
BB7	$[1 - (1-t)^\delta]^{-\theta} - 1$	$\theta \rightarrow \infty$ o $\delta \rightarrow \infty$	—	—

TABLA C.5. Función generadora ($\phi(t)$) y cópulas límite para algunas cópulas de Arquímedes