



University of Barcelona

Research Group on Risk in Insurance and Finance www.ub.edu/riskcenter

Working paper 2015/01 \ \ Number of pages 18

Estimación del riesgo mediante el ajuste de cópulas

Catalina Bolancé, Montserrat Guillén y Alemar Padilla



Estimación del riesgo mediante el ajuste de cópulas
Catalina Bolancé, Montserrat Guillén y Alemar Padilla
Departamento de Econometría, Estadística y Economía Española
Universidad de Barcelona
Riskcenter
Diagonal, 690, 08034 Barcelona

Resumen

El propósito de éste trabajo, es presentar una forma de cuantificar el valor en riesgo de una cartera de activos mediante dos medidas de riesgo ampliamente conocidas como lo son el VaR y el TVaR. Para ello, utilizamos cópulas paramétricas bivalentes y el método de simulación de Monte Carlo.

Palabras claves: dependencia, copulas, Value at Risk, Tail Value at Risk, cuantificación del riesgo.

1. Introducción a las copulas

En una situación de riesgo, cuando la pérdida total, a la que denominaremos L (loss), está generada por las posibles pérdidas ocasionadas por k sub-riesgos L_1, \dots, L_k , por ejemplo $L = L_1 + \dots + L_k$, el riesgo de pérdida total depende de la relación entre los sub-riesgos. Las cópulas (palabra que deriva de Latín y significa “unión”, “link”) permiten modelizar un amplio rango de estructuras de dependencia, lineales o no lineales. Describiremos las cópulas más comunes y para simplificar la exposición nos ceñimos al caso bivalente¹, $k = 2$.

Sea (L_1, L_2) un vector aleatorio bivalente que representa dos pérdidas dependientes y sean $F_1(l_1) = P(L_1 \leq l_1)$ y $F_2(l_2) = P(L_2 \leq l_2)$ sus dos funciones de distribución continuas, siguiendo el teorema de Sklar² existe una única cópula $C_\theta: [0,1]^2 \rightarrow [0,1]$, con parámetro de dependencia θ tal que, siendo $F(l_1, l_2) = P(L_1 \leq l_1, L_2 \leq l_2)$ la función de distribución conjunta, se cumple que:

$$F(l_1, l_2) = C_\theta(u_1, u_2), \forall l_1, l_2 \in \mathcal{R},$$

donde $u_1 = F_1(l_1)$ y $u_2 = F_2(l_2)$ son, respectivamente, los valores de dos variables aleatorias U_1 y U_2 con distribución Uniforme(0,1).

¹ Para la mayoría de cópulas la generalización a más de dos dimensiones ($k > 2$) no es directa. Una alternativa consiste en unir pares de variables hasta relacionarlas todas. Ver, por ejemplo, Aas, K., Czado, C., Frigessic, A. y Bakken, H. (2009) “Pair-copula constructions of multiple dependence” *Insurance: Mathematics and Economics*, 44(2), 182-198.

² Sklar, A. (1959) “Fonctions de répartition á n dimensions et leurs marges”. *Publications de l'Institut de Statistique de l'université de Paris*, 8, 229-231.

Existen diversas formas de clasificar los tipos de cópulas, una de las más comunes es la que divide las cópulas en:

- Cópulas implícitas: su forma funcional coincide con una función de distribución conocida (Normal, t-Student)
- Cópulas explícitas: posee una forma funcional propia y sencilla.

Otra clasificación muy utilizada es la que diferencia las cópulas Arquimedianas de las que no lo son. La particularidad de las cópulas Arquimedianas está en que existe una función que las genera, lo que se denomina el generador de la cópula³.

En la Tabla 1 clasificamos las cópulas que se describen en este documento.

Tabla 1: Clasificación de las cópulas analizadas.

Nombre de la Cópula	Clasificación
Gaussiana	Implícita y no Arquimediana
t-Student	Implícita y no Arquimediana
Gumbel	Explícita y Arquimediana
Clayton	Explícita y Arquimediana
Frank	Explícita y Arquimediana

2. Medidas de dependencia

Uno de los aspectos fundamentales relacionados con las cópulas es que éstas han servido para definir nuevas medidas de dependencia, alternativas al coeficiente de correlación lineal de Pearson. A continuación, además de este coeficiente de correlación, definimos las medidas de dependencia alternativas y la forma en que se calculan a partir de una muestra de pérdidas observadas.

El coeficiente de correlación de Pearson

El coeficiente de correlación de Pearson es una medida de dependencia lineal que teóricamente equivale a:

$$\rho_P = \frac{E[(L_1 - \mu_1)(L_2 - \mu_2)]}{\sqrt{E[(L_1 - \mu_1)^2]E[(L_2 - \mu_2)^2]}}$$

donde $E[\cdot]$ es la esperanza matemática, siendo $\mu_1 = E[L_1]$ y $\mu_2 = E[L_2]$ las pérdidas medias teóricas. A partir de una muestra de N pérdidas observadas (L_{i1}, L_{i2}) , $i = 1, \dots, N$, el coeficiente de correlación de Pearson se estima como:

³ Nelsen, R. (2006) *An Introduction to Copulas*. Springer, USA.

$$\hat{\rho}_P = \frac{\sum_{i=1}^N (L_{i1} - \hat{\mu}_1)(L_{i2} - \hat{\mu}_2)}{\sqrt{\sum_{i=1}^N (L_{i1} - \hat{\mu}_1)^2 \sum_{i=1}^N (L_{i2} - \hat{\mu}_2)^2}}$$

donde $\hat{\mu}_1 = \frac{\sum_{i=1}^N L_{i1}}{N}$ y $\hat{\mu}_2 = \frac{\sum_{i=1}^N L_{i2}}{N}$ son las medias muestrales de ambas pérdidas.

El coeficiente de correlación de Pearson es una medida de dependencia lineal, por lo que si la relación entre las pérdidas no es lineal este coeficiente no mide correctamente la dependencia y el hecho de que tome valor cero no implica independencia⁴. Ante ello, a partir de las cópulas se plantean coeficientes de correlación alternativos:

- La τ de Kendall.
- El coeficiente de correlación de Spearman.
- Los coeficientes de dependencia de las colas (*tail dependence*).

La τ de Kendall

La τ de Kendall es una medida de asociación cuyo valor puede deducirse a partir de la cópula:

$$\tau = 4 \int_0^1 \int_0^1 C_\theta(u_1, u_2) dC_\theta(u_1, u_2) - 1.$$

A partir de una muestra de N pérdidas observadas $(L_{i1}, L_{i2}), i = 1, \dots, N$, la τ de Kendall se estima tal y como se describe a continuación. Para cada par de observaciones consecutivas (L_{i1}, L_{i2}) y (L_{j1}, L_{j2}) se concluye que son concordantes si $L_{i1} \geq L_{j1}$ y $L_{i2} \geq L_{j2}$ o si $L_{i1} \leq L_{j1}$ y $L_{i2} \leq L_{j2}$, en caso contrario las observaciones son discordantes. En el caso $L_{i1} = L_{j1}$ y $L_{i2} = L_{j2}$ las observaciones no son ni concordantes ni discordantes. Una vez identificados los pares concordantes y discordantes, la τ de Kendall se estima como:

$$\hat{\tau} = \frac{\text{Número de pares concordantes} - \text{Número de pares discordantes}}{\frac{1}{2}N(N-1)}.$$

El coeficiente de correlación de Spearman

El coeficiente de correlación de Spearman también puede obtenerse a partir de la cópula:

⁴ Un valor del coeficiente de correlación de Pearson igual a cero implica independencia cuando la distribución conjunta de las pérdidas es Normal multivariante.

$$\rho_s = 12 \int_0^1 \int_0^1 C_\theta(u_1, u_2) du_1 du_2 - 3.$$

La estimación de este coeficiente se obtiene a partir del orden de las observaciones. Es decir, sean $(R_{i1}, R_{i2}), i = 1, \dots, N$, dos variables que indican las posiciones que toman en la muestra las pérdidas $(L_{i1}, L_{i2}), i = 1, \dots, N$, si las ordenamos de mayor a menor, el coeficiente estimado se obtiene como:

$$\hat{\rho}_s = \frac{6}{N(N^2 - 1)} \sum_{i=1}^N (R_{i1} - R_{i2})^2.$$

Además, el coeficiente de correlación de Spearman se puede obtener estimando la correlación lineal entre $F_1(L_{i1})$ y $F_2(L_{i2}), i = 1, \dots, N$.

El coeficiente de dependencia de la cola

Se pueden calcular dos coeficientes de dependencia de la cola (*tail dependence*), el de la cola derecha de la distribución λ_u (*upper tail dependence*) y el de la cola izquierda λ_l (*lower tail dependence*). A partir de la cópula se obtiene que el coeficiente de dependencia de la cola derecha es:

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

donde, $\bar{C}_\theta(u, u) = 1 - C_\theta(u, u)$ es la cópula de supervivencia y el límite indica que los argumentos tienden a 1 por la derecha, lo que equivale a que ambas pérdidas tienden a $+\infty$ por la izquierda.

Por su parte, el coeficiente de dependencia de la cola izquierda es:

$$\lambda_l = \lim_{u \rightarrow 0^+} \frac{C_\theta(u, u)}{u},$$

donde el límite indica que los argumentos tienden a $-\infty$ por la derecha.

3. Estimación

Los métodos de estimación para las cópulas son tres:

- Método máximo verosímil.
- Método pseudo-máximo verosímil.
- Método de los momentos (*Inference function for margins*, IFM).

Una de las ventajas de las cópulas es que permiten estimar de forma independiente el parámetro de la cópula y las marginales, en esto consisten los métodos pseudo-maximo verosímil e IFM.

Método máximo verosímil

El método máximo verosímil consiste en maximizar el logaritmo de la función de verosimilitud asociada a la cópula y las marginales conjuntamente. Sea $\hat{\theta}$ el parámetro de la cópula estimado y sean $\hat{\beta}_1$ y $\hat{\beta}_2$ los parámetros estimados que definen las funciones de distribución marginales, el logaritmo de la función de verosimilitud es:

$$l(\hat{\theta}, \hat{\beta}_1, \hat{\beta}_2) = \sum_{i=1}^N \ln\{c_{\hat{\theta}}(U_{i1}, U_{i2})\} + \ln\{f_1(L_{i1}; \hat{\beta}_1)\} + \ln\{f_2(L_{i2}; \hat{\beta}_2)\},$$

donde $U_{i1} = F_1(L_{i1}; \hat{\beta}_1)$, $U_{i2} = F_2(L_{i2}; \hat{\beta}_2)$, $c_{\hat{\theta}}$ es la densidad de la cópula y f_1 y f_2 son las funciones de densidad de las marginales⁵. El método máximo verosímil no suele utilizarse dado que no suele ser fácil encontrar una configuración de parámetros de cópula y marginales que maximicen $l(\hat{\theta}, \hat{\beta}_1, \hat{\beta}_2)$.

Método pseudo-máximo verosímil

Concretamente, el método pseudo-máximo verosímil consiste en generar lo que se denomina las pseudo-observaciones $(\tilde{U}_{i1}, \tilde{U}_{i2})$, $i = 1, \dots, N$, las cuales se obtienen a partir de la distribución empírica corregida:

$$\tilde{U}_{i1} = \frac{N}{N+1} \sum_{j=1}^N I(L_{j1} \leq L_{i1}) \text{ y } \tilde{U}_{i2} = \frac{N}{N+1} \sum_{j=1}^N I(L_{j2} \leq L_{i2}). \quad (1)$$

Utilizando las pseudo-observaciones se obtiene la pseudo-verosimilitud a maximizar respecto al parámetro de la cópula:

$$l(\hat{\theta}) = \sum_{i=1}^N \ln\{c_{\hat{\theta}}(\tilde{U}_{i1}, \tilde{U}_{i2})\}. \quad (2)$$

Tras estimar el parámetro de la cópula se estiman las marginales de forma independiente. A este método se le denomina pseudo-verosímil, debido a que el resultado no maximiza la verosimilitud total, sino que maximiza las verosimilitudes parciales, de la cópula y las marginales.

⁵ La función de densidad es la derivada de la función de distribución.

Método IFM

El método de IFM también parte de las pseudo-observaciones y consiste en estimar el parámetro de la cópula por el método de los momentos. Normalmente, se puede encontrar una relación entre este parámetro y una medida de dependencia como el coeficiente de correlación de Spearman o la τ de Kendall. Posteriormente, para cada caso concreto, analizaremos la relación entre los parámetros de la cópula y los coeficientes de dependencia que se utilizan en la estimación de momentos. En este caso las marginales también se obtienen por separado.

4. Simulación

Para la cuantificación del riesgo a partir de cópulas se utiliza el **método de simulación de Monte Carlo**, que consiste en generar pérdidas simuladas asociadas a los valores aleatorios provenientes del modelo ajustado a nuestros datos y seguidamente estimar el riesgo de forma empírica.

En general, para generar valores a partir de una cópula bidimensional el procedimiento puede dividirse en los siguientes 5 pasos:

1. Calcular las pseudo-observaciones tal y como se describía en la expresión (1).
2. Estimar el parámetro de la cópula maximizando la expresión (2) o por el método IFM.
3. Simular a partir de la cópula los valores $(\hat{U}_{r1}, \hat{U}_{r2})$, $r = 1, \dots, R$, donde R es el número de réplicas simuladas, el cual tiene que ser lo suficientemente grande (por ejemplo, $R = 100000$), cuyo valor dependerá de la capacidad de cálculo computacional.
4. Obtener las pérdidas simuladas como: $\hat{L}_{r1} = F_1^{-1}(\hat{U}_{r1}; \hat{\beta}_1)$ y $\hat{L}_{r2} = F_2^{-1}(\hat{U}_{r2}; \hat{\beta}_2)$.

En general, para simular los valores procedentes de una cópula bivalente (paso 3) se utiliza el método basado en la distribución condicionada:

$$P(U_2 \leq u_2 | U_1 = u_1) = \frac{\partial C_{\theta}(u_2)}{\partial u_1} = C_{\theta; u_1}(u_2).$$

El algoritmo se implementa del siguiente modo:

- (1) Generar valores de dos variables aleatorias independientes con distribución *Uniforme*(0,1), \hat{u}_1 y t .
- (2) Obtener $\hat{u}_2 = C_{\theta; u_1}^{[-1]}(t)$, donde el subíndice $[-1]$ indica pseudo-inversa, que equivale a:

$$C_{\theta;u_1}^{[-1]}(t) = \begin{cases} \inf\{x|C_{\theta;u_1}(x) \leq t\}, & \text{if } t = 0 \\ C_{\theta;u_1}^{[-1]}, & \text{if } t \in (0,1) \\ \inf\{x|C_{\theta;u_1}(x) \leq t\}, & \text{if } t = 1 \end{cases}$$

Sin embargo, la cópula condicionada $C_{\theta;u_1}(u_2)$ no siempre es invertible, como sucede con el caso de la cópula de Gumbel, en este caso se puede utilizar el algoritmo de⁶ Genest y Rivest (1993) que consiste en lo siguiente:

- (1) Generar valores de dos variables aleatorias independientes con distribución *Uniforme*(0,1), \tilde{u}_1 y t .
- (2) Obtener $z = K^{-1}(t)$, donde $K(s) = s - \frac{\varphi(s)}{\varphi'(s)}$, donde el superíndice ' indica derivada y $\varphi(\cdot)$ es el generador de la cópula arquimediana.
- (3) Calcular $\hat{u}_1 = \varphi^{-1}(\tilde{u}_1\varphi(z))$ y $\hat{u}_2 = \varphi^{-1}((1 - \tilde{u}_1)\varphi(z))$.

Además, para generar valores de las cópulas Gaussiana y t-Student es más recomendable utilizar un método clásico de simulación, para asegurar que los valores simulados de la t-Student poseen una cola más pesada que los de la Gaussiana. El método, en este caso, consiste en:

- (1) Generar valores de dos variables aleatorias independientes con distribución *Uniforme*(0,1), \tilde{u}_1 y t .
 - (2) Calcular $x_1 = F_1^{-1}(\tilde{u}_1)$ y $z_2 = F_2^{-1}(t)$.
 - (3) Calcular $x_2 = \rho x_1 + \sqrt{1 - \rho^2} z_2$, donde ρ es el coeficiente de correlación lineal.
 - (4) Calcular $\hat{u}_1 = F_1(x_1)$ y $\hat{u}_2 = F_2(x_2)$.
5. Cálculo del VaR y TVaR de la pérdida total \hat{L}_r .

5. Bondad del ajuste

Para comparar el ajuste de distintas cópulas y marginales puede utilizarse una medida de información estadística clásica como es el criterio de información de Akaike (*Akaike information criterion*, AIC) que equivale a:

$$AIC = -2l(\hat{\theta}) + 2k,$$

donde k es el número de parámetros, que en el caso de la cópula bivalente es igual a 1. Otro criterio que tiene en cuenta el tamaño muestral n es el criterio de información bayesiano (*Bayesian information criterion*, BIC):

⁶ Genest, C. y Rivest, L. (1993) "Statistical inference procedure for bivariate archimedean copulas" *Journal of the American Statistical Association*, 88, 1034-1043.

$$BIC = -2l(\hat{\theta}) + k \ln(n).$$

En ambos casos, cuanto menor es el valor de la medida, sea ésta el AIC o el BIC, mejor es el ajuste. Por ejemplo, supongamos que disponemos de datos diarios, 757 días en total, sobre los rendimientos de dos inversiones dependientes y queremos comparar el ajuste de la cópula Gaussiana y la cópula t-Student. Como método de estimación se utiliza el basado en la maximización de la pseudo-verosimilitud, es decir, el **método pseudo-máximo verosímil** descrito anteriormente.

6. Cópulas utilizadas

o Cópula Gaussiana

La cópula Gaussiana debe su nombre a que coincide con la función de distribución estándar de una Normal bivalente con correlación $-1 \leq \rho \leq 1$:

$$C_{\rho}(u_1, u_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{2\rho t_1 t_2 - t_1^2 - t_2^2}{2(1-\rho^2)}\right) dt_1 dt_2,$$

donde $x_1 = F_1^{-1}(u_1)$ y $x_2 = F_2^{-1}(u_2)$, por definición de la cópula Gaussiana, las funciones de distribución marginales coinciden con la de la Normal estándar. Si observamos la expresión de la cópula Gaussiana vemos que x_1 y x_2 son valores de dos variables estandarizadas, con media cero y varianza 1. La densidad de la cópula Gaussiana es la densidad de la Normal:

$$c_{\rho}(u_1, u_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{2\rho t_1 t_2 - t_1^2 - t_2^2}{2(1-\rho^2)}\right).$$

La cópula Gaussiana no tiene expresiones cerradas para los coeficientes de Spearman y τ de Kendall, éstos han de aproximarse a partir de sus expresiones generales en función de la cópula. Además, la cópula Gaussiana supone que la dependencia en las colas es cero, es decir, $\lambda_u = \lambda_l = 0$.

o Cópula t-Student

De forma similar a la cópula Gaussiana, la cópula t-Student debe su nombre a que coincide con la función de distribución estándar de una t-Student bivalente con v grados de libertad y correlación $-1 \leq \rho \leq 1$:

$$C_{v,\rho}(u_1, u_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \frac{\Gamma\left(\frac{v+2}{v}\right)}{\Gamma\left(\frac{v+2}{v}\right) 2\pi\sqrt{1-\rho^2}} \exp\left(1 + \frac{2\rho t_1 t_2 - t_1^2 - t_2^2}{v(1-\rho^2)}\right) dt_1 dt_2,$$

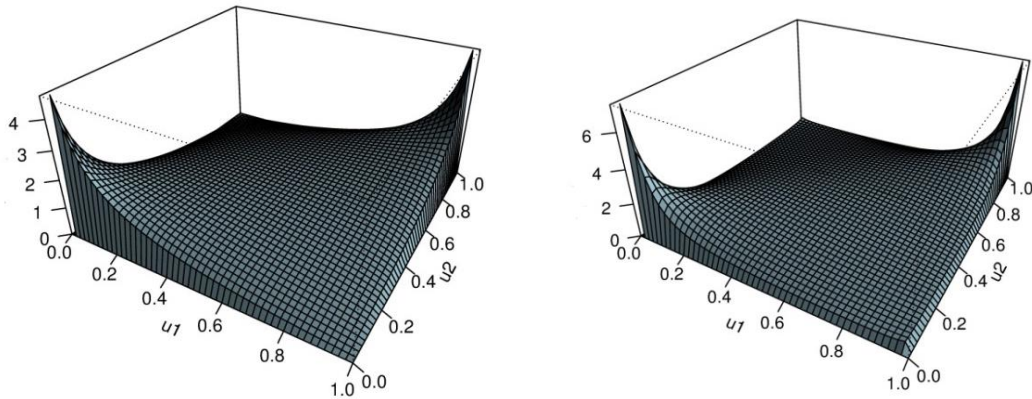
donde $x_1 = F_1^{-1}(u_1)$ y $x_2 = F_2^{-1}(u_2)$. Por definición de la cópula t-Student, las funciones de distribución marginales coinciden con la de la t-Student estándar. La densidad de la cópula t-Student es:

$$c_{v,\rho}(u_1, u_2) = \frac{\Gamma\left(\frac{v+2}{v}\right)}{\Gamma\left(\frac{v+1}{v}\right)^2 2\pi\sqrt{1-\rho^2}} \exp\left(1 + \frac{2\rho t_1 t_2 - t_1^2 - t_2^2}{v(1-\rho^2)}\right).$$

En la Figura 1 se representan las densidades de las cópula Gaussiana y t-Student para $\rho = 0,5$ y, en el caso de la t-Student, $v = 4$. Ambas cópulas son simétricas y pertenecen a la familia de las cópulas elípticas. Si observamos los valores de las densidades en el eje vertical, vemos como la cópula t-Student tiene las colas más pesadas (la densidad es más elevada en los valores próximos a 1 ó a 0).

Al igual que la cópula Gaussiana, la t-Student tampoco tiene expresiones cerradas para los coeficientes de Spearman y τ de Kendall. Por otro lado, la cópula t-Student sí que tiene dependencia en las colas, es decir, $\lambda_u = \lambda_l \geq 0$.

Figura 1: Densidades de las cópulas Gaussiana (izquierda) y t-Student (derecha).



○ Cópula de Gumbel

La cópula de Gumbel equivale a:

$$C_\theta(u_1, u_2) = \exp\left(-\left[(-\ln(u_1))^\theta + (-\ln(u_2))^\theta\right]^{\frac{1}{\theta}}\right),$$

donde $\theta \in [1, +\infty)$. En la cópula de Gumbel la dependencia es perfecta cuando $\theta \rightarrow +\infty$ y coincide con la cópula de independencia cuando $\theta = 1$. La cópula de Gumbel permite calcular el valor de la τ de Kendall de forma sencilla, ya que se cumple:

$$\tau(\theta) = 1 - \frac{1}{\theta},$$

relación que puede utilizarse para obtener el estimador de momentos a partir de la estimación empírica de la τ de Kendall. La cópula de Gumbel posee dependencia en la cola derecha, y se obtiene considerando $\lambda_u = 2 - 2^{\frac{1}{\theta}}$ y $\lambda_l = 0$.

La densidad de la cópula de Gumbel es:

$$c_{\theta}(u_1, u_2) = C_{\theta}(u_1, u_2) u_1^{-1} u_2^{-1} \left(\left[(-\ln(u_1))^{\theta} + (-\ln(u_2))^{\theta} \right]^{-2 + \frac{2}{\theta}} \right) [\ln(u_1) \ln(u_2)]^{\theta-1} \left[1 + (\theta + 1) \left[(-\ln(u_1))^{\theta} + (-\ln(u_2))^{\theta} \right]^{\frac{1}{\theta}} \right].$$

○ Cópula de Clayton

La cópula de Clayton equivale a:

$$C_{\theta}(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}},$$

donde $\theta > 0$. En la cópula de Clayton la dependencia es perfecta cuando $\theta \rightarrow +\infty$ y coincide con la cópula de independencia cuando $\theta \rightarrow 0$. Al igual que la cópula de Gumbel, la cópula de Clayton permite calcular el valor de la τ de Kendall de forma sencilla, ya que se cumple:

$$\tau(\theta) = \frac{1}{2\theta + 1},$$

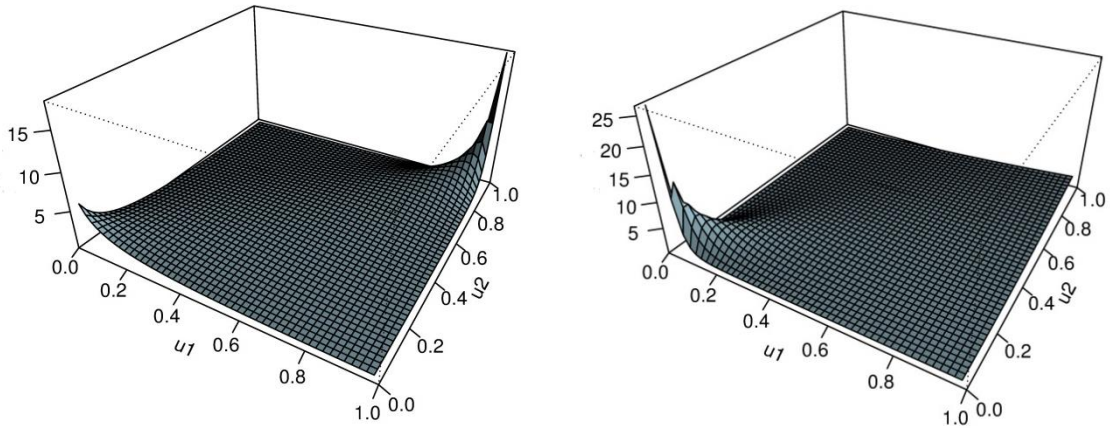
relación que puede utilizarse para obtener el estimador de momentos a partir de la estimación empírica de la τ de Kendall. Al contrario que Gumbel, la cópula de Clayton posee dependencia en la cola izquierda, y se obtiene considerando $\lambda_u = 0$ y $\lambda_l = 2^{-\frac{1}{\theta}}$.

La densidad de la cópula de Clayton es:

$$c_{\theta}(u_1, u_2) = \left(1 + \frac{1}{\theta} \right) \left[(1 - u_1)^{-\frac{1}{\theta}} + (1 - u_2)^{-\frac{1}{\theta}} - 1 \right]^{-\theta-2} [(1 - u_1)(1 - u_2)]^{-1-\frac{1}{\theta}}.$$

En la Figura 2, se representan las densidades de las cópulas de Gumbel y de Clayton, en este caso se observa como ambas cópulas son asimétricas. Concretamente, la cópula de Gumbel alcanza la densidad máxima cuando ambos argumentos tienden a 1 (dependencia en la cola derecha). Por el contrario, la cópula de Clayton alcanza la densidad máxima cuando ambos argumentos tienden a 0 (dependencia en la cola izquierda).

Figura 2: Densidades de las cópulas Gumbel (izquierda) y Clayton (derecha), ambas con $\theta = 2$.



○ Cópula de Frank

La cópula de Frank equivale a:

$$C_{\theta}(u_1, u_2) = -\frac{1}{\theta} \ln \left(1 - \frac{(1 - e^{\theta u_1})(1 - e^{\theta u_2})}{(1 - e^{\theta})} \right),$$

donde $\theta \in]-\infty, 0[\cup]0, +\infty[$. Al igual que las cópulas Gaussiana y t-Student, la cópula de Frank admite dependencia positiva o negativa. El valor de la τ de Kendall para la cópula de Frank es:

$$\tau(\theta) = 1 - \frac{4}{\theta} + \frac{4}{\theta^2} \int_0^{\theta} \frac{t}{e^t - 1} dt.$$

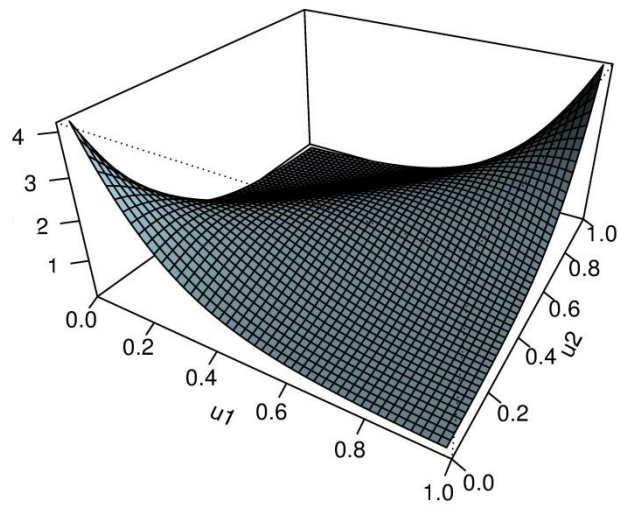
La cópula de Frank no posee dependencia en las colas.

La densidad de la cópula de Frank es:

$$c_{\theta}(u_1, u_2) = -\theta e^{\theta} \left(\frac{1 + e^{(u_1+u_2)\theta}}{e^{\theta u_1} e^{\theta u_2} + e^{2\theta}} \right).$$

En la Figura 3 se muestra la densidad de la cópula de Frank, se observa como su comportamiento en las colas es similar a la cópula Gaussiana, aunque muestra más dependencia en los valores centrales.

Figura 3: Densidades de las cópula de Frank con $\theta = 5$.



7. Medidas de riesgo

Las medidas de riesgo permiten cuantificar las pérdidas obtenidas en una situación de riesgo. En éste caso, dicho cálculo se realizará a través de dos medidas ampliamente conocidas como lo son el *Value at Risk* y el *Tail Value at Risk*.

- **Valor en Riesgo (*Value at Risk*)**

Sea $L = L_1 + L_2$ la variable aleatoria (en el caso bivalente) que representa la pérdida total en una situación de riesgo y sea F_L su función de distribución, entonces se define el Valor en Riesgo (*Value at Risk*) como:

$$\text{VaR}_\alpha(L) = \inf\{l \in \mathbb{R}: F_L(l) \geq \alpha\} = F_L^{-1}(\alpha),$$

donde α es lo que se denomina nivel de confianza que coincide con la probabilidad cercana a 1. El VaR es simplemente el cuantil $(1 - \alpha)$ de la variable pérdida, así que la probabilidad de tener una pérdida superior a la determinada por el VaR será $(1 - \alpha)$.

- ***Tail Value at Risk***

El *Tail Value at Risk* se define como:

$$\text{TVaR}_\alpha(L) = \frac{1}{1 - \alpha} \int_\alpha^1 \text{VaR}_u(L) du$$

Y que en el caso de que L sea continua la expresión anterior se reescribe como:

$$\text{TVaR}_\alpha(L) = E[L|L > \text{VaR}_\alpha(L)]$$

8. Ejemplo con datos financieros

A continuación se presenta un ejemplo de cómo realizar el cálculo del riesgo de una cartera a partir del uso de cópulas paramétricas bivariantes y el método de simulación de Monte Carlo. En primer lugar se expondrá el proceso de estimación de los parámetros de las cópulas, luego se realizará la estimación de las diferentes distribuciones marginales y finalmente se calculará el VaR y TVaR.

El software utilizado será R.

a) Datos

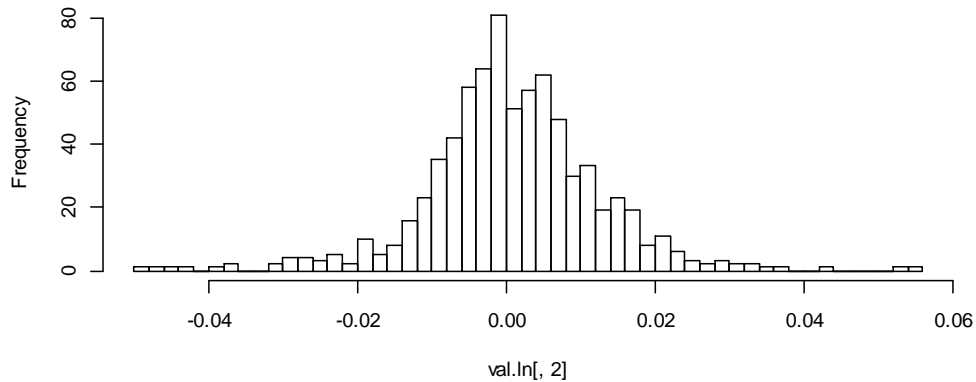
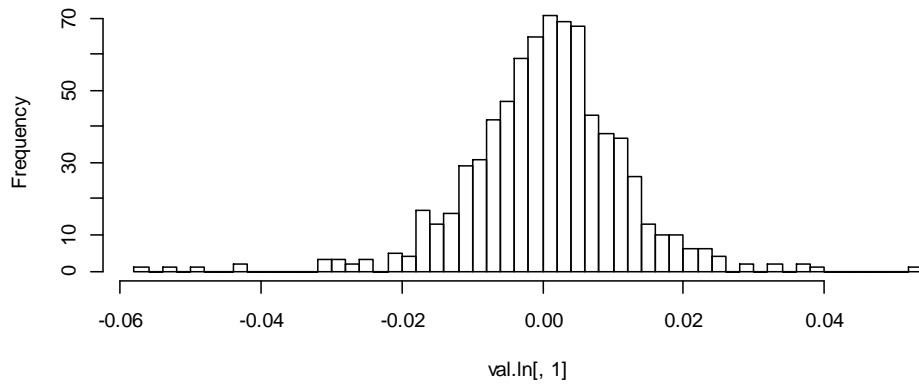
Se dispone de 753 observaciones asociadas a los log-returns de los precios diarios de las siguientes series temporales: Merck, Novartis, Pfizer, Deutsche Bank, ING, Santander, NASDAQ y S&P500. Para realizar el ejemplo solo utilizaremos Merck y Novartis.

Estadísticas descriptivas

A continuación se presentan las estadísticas descriptivas y los histogramas de nuestros datos:

Variables	Media	Desviación E.	Asimetría	Min	Max	Mediana
NVS	0.0005	0.0109	-0.4136	-0.0569	0.0532	0.0009
PFE	0.0008	0.0116	-0.0942	-0.0484	0.0547	0.0004

Histogramas de los log-returns



En general se observa que los valores de los estadísticos descriptivos son parecidos entre sí. Aunque ambas series presentan una pequeña asimetría negativa, no dejan de ser bastante simétricas, cuestión que se evidencia visual y numéricamente con la cercanía de sus medias y medianas. Así mismo, es clara la presencia de colas pesadas en los gráficos de ambas distribuciones, lo cual es característicos de éste tipo de datos financieros.

La distribución de los datos es un indicador importante para la selección del tipo de cópulas de ajuste y de las funciones de distribución marginal utilizadas en la simulación.

Pseudo observaciones

Previo al ajuste de las cópulas, será necesario transformar nuestros log-returns en pseudo observaciones. Con ello garantizaremos la obtención de variables aleatorias uniformemente distribuidas en el intervalo (0,1), como input para la posterior modelización.

b) Ajuste y estimación de los parámetros de las cópulas

Las cópulas utilizadas, serán las mencionadas en la sección 1.

El ajuste de las cópulas se realiza en función de uno o más parámetros específicos. Para ello, se propone según el tipo de cópula uno o más valores iniciales y posteriormente a través del ajuste, se realiza la estimación del parámetro o parámetros finales.

Según la naturaleza de cada cópula se utilizan diferentes parámetros iniciales. En el caso de las cópulas implícitas se considerará como parámetro inicial la correlación lineal, estimada por el método clásico. Por su parte, para las cópulas de Gumbel y Clayton se utiliza en el método de los momentos basado en la τ de Kendall. Y en el caso de la cópula de Frank se parte de un valor inicial aleatorio dado que el estimador de momentos no es directo.

A continuación se exponen los parámetros iniciales propuestos, los parámetros finales obtenidos y el logaritmo de la verosimilitud para cada cópula ajustada:

Estimación de parámetros

Copulas	Parámetros iniciales	Parámetros finales	Log-verosimilitud
Gaussiana	0.5297	0.5068	109.4706
t-Student	0.5297 / df=4.0000	0.5016 / df= 5.4730	118.1811
Gumbel	1.4939	1.4491	97.7283
Clayton	0.9878	0.8450	108.8734
Frank	-	3.3007	96.2156

c) Bondad del ajuste

La bondad del ajuste de los modelos se valorará en función de los criterios de evaluación *Akaike's Information Criterion* (AIC) y *Schwarz's Bayesian criterion* (BIC) mencionados en la sección 5. Mientras más pequeño sea el valor obtenido, habrá menor pérdida de información y por tanto un mejor ajuste.

A continuación se exponen los diferentes valores del AIC y BIC obtenidos:

AIC - BIC

Copulas	AIC	BIC
Gaussiana	-216.9412	-212.3171
t-Student	-232.3622	-223.1140
Gumbel	-193.4566	-188.8326
Clayton	-215.7467	-211.1227
Frank	-190.4312	-185.8071

A partir de estos resultados se concluye que la cópula t-Student es la que mejor ajusta el comportamiento de nuestros datos.

d) Tail dependence

Los coeficientes de dependencia en las colas λ_l y λ_u , son medidas de dependencia entre los cuantiles extremos de las distribuciones de nuestras variables aleatorias.

Según sea la naturaleza de cada cópula tendrá sentido el cálculo de ambos o de sólo uno de ellos.

A continuación se muestran los resultados obtenidos:

Copulas	Dependencia en la cola izquierda	Dependencia en la cola derecha
Gaussiana	0.0000	0.0000
t-Student	0.1895	0.1895
Gumbel	0.0000	0.3866
Clayton	0.4403	0.0000
Frank	0.0000	0.0000

e) Simulación

La cuantificación del riesgo se realizará por el método de simulación de Monte Carlo expuesto en la sección 4, considerando las medidas de riesgo definidas en la sección 7, es decir, primero se obtendrán las pérdidas simuladas provenientes de la cópula ajustada y posteriormente se realizará la cuantificación del riesgo.

Estimación de las marginales

La estimación de las funciones de distribución marginal se realiza en función del tipo de datos, por ello, en éste caso se propone ajustar las distribuciones Normal y t-Student.

Con la finalidad de valorar el ajuste de estas distribuciones, se realiza el cálculo del AIC y BIC, obteniéndose los siguientes resultados:

AIC - BIC de las marginales

Marginales	AIC	BIC
Normales	-4665.3810	-4656.1330
	-4569.4330	-4560.1840
t-Student	-4751.7000	-4737.8280
	-4653.1850	-4639.3130

Se observa que el mejor resultado se logra cuando se ajustan funciones de distribución del tipo t-Student.

Simulación de los factores de riesgo

Los factores de riesgo L_1 y L_2 , son cuantiles asociados a los valores aleatorios provenientes de las cópulas ajustadas y cuyo comportamiento viene determinado por las funciones de distribución marginales $F_1(l_1)$ y $F_2(l_2)$.

Una vez que han sido calculados los factores de riesgo, se define el vector de pérdidas linealizadas, a partir del cual se realizará la cuantificación del riesgo. Para ello se ha de tener en cuenta que el valor de la cartera en el momento t es igual a 1, y los pesos de cada factor de riesgo o proporción de la inversión en cada acción son también iguales a 1.

Cuantificación del riesgo

El cálculo de las medidas de riesgo se realizará considerando los siguientes niveles de confianza: 99%, 99.5% y 99.9%.

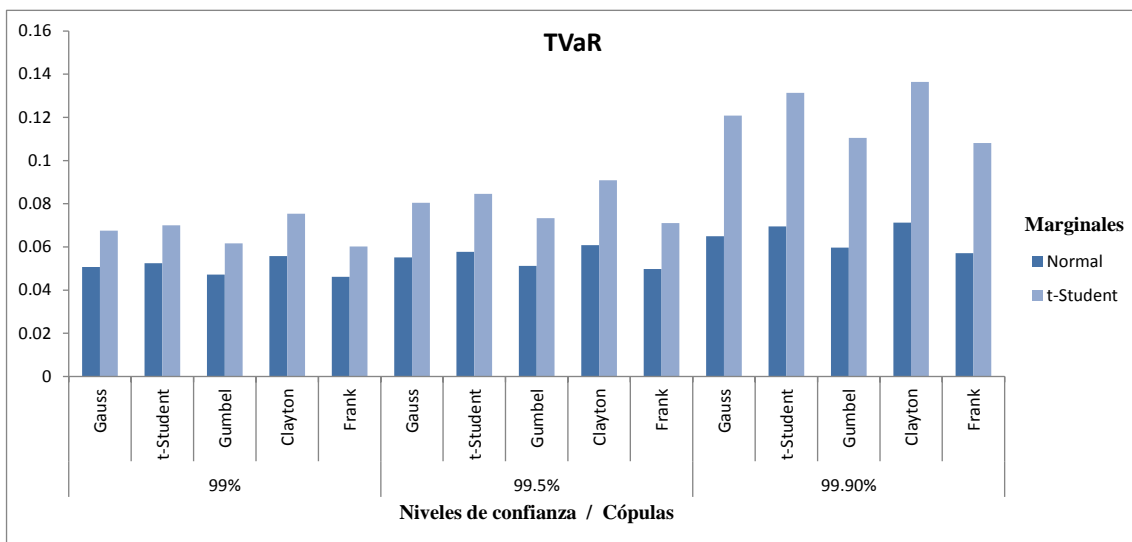
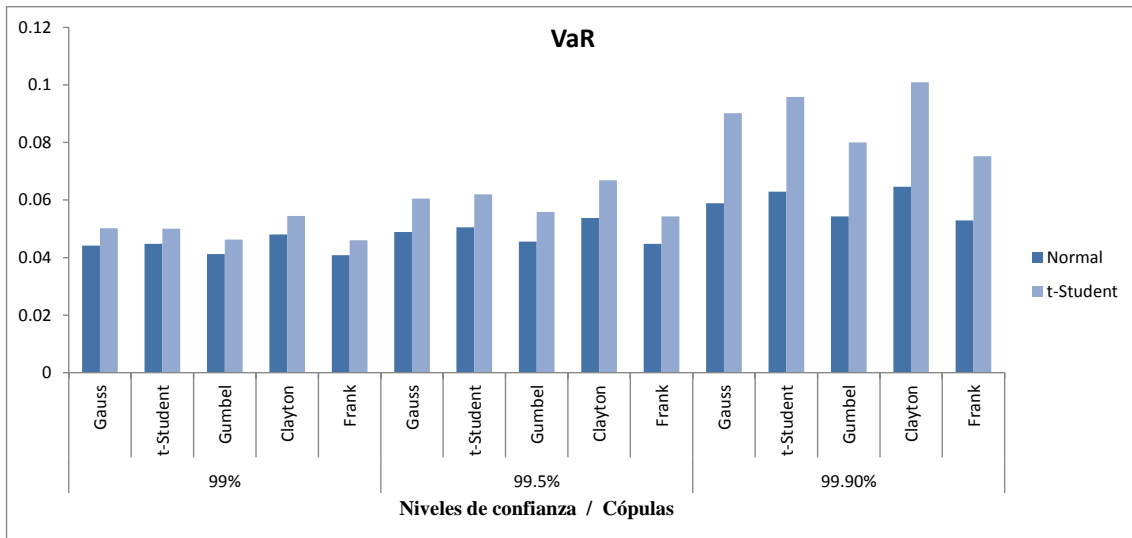
A continuación se exponen los resultados obtenidos:

Valor en Riesgo (Value at Risk)

Copulas	Marginales	VaR		
		99 %	99.5 %	99.9 %
Gaussiana	Normal	0.0441	0.0488	0.0588
	t-Student	0.0502	0.0605	0.0901
t-Student	Normal	0.0447	0.0505	0.0629
	t-Student	0.0500	0.0620	0.0958
Gumbel	Normal	0.0412	0.0455	0.0543
	t-Student	0.0463	0.0558	0.0800
Clayton	Normal	0.0480	0.0538	0.0646
	t-Student	0.0545	0.0668	0.1009
Frank	Normal	0.0408	0.0448	0.0529
	t-Student	0.0460	0.0543	0.0752

Tail Value at Risk

Copulas	Marginales	TVaR		
		99 %	99.5 %	99.9 %
Gaussiana	Normal	0.0507	0.0551	0.0649
	t-Student	0.0675	0.0805	0.1208
t-Student	Normal	0.0525	0.0577	0.0695
	t-Student	0.0700	0.0846	0.1314
Gumbel	Normal	0.0472	0.0512	0.0597
	t-Student	0.0617	0.0733	0.1105
Clayton	Normal	0.0557	0.0608	0.0712
	t-Student	0.0754	0.0909	0.1364
Frank	Normal	0.0462	0.0498	0.0571
	t-Student	0.0602	0.0710	0.1081



Se observa que el valor asociado al cálculo de ambas medidas de riesgo es mucho mayor cuando se consideran marginales del tipo t-Student. Así mismo, en el cuantil 0.999 es donde se evidencia una diferencia superior respecto a los resultados obtenidos en los otros dos cuantiles.

Conclusiones

La cuantificación del riesgo mediante cópulas resulta una alternativa flexible en términos de la modelización de la dependencia entre variables. Ésta se ve notablemente afectada por el tipo de cópula y marginal utilizada, por lo cual el resultado obtenido dependerá de la selección realizada.

En términos de ajuste, la cópula y marginales del tipo t-Student son las que mejor modelan el comportamiento conjunto e individual de nuestros datos.

Referencias

- [1].Aas, K., Czado C., Frigessic, A. y Bakkend, H. (2009) “Pair-copula constructions of multiple dependence”. *Insurance: Mathematics and Economics*, 44(2), 182-198.
- [2].Bahraoui, Z., Bolancé, C. y Pérez-Marín, A.M. (2014). “Testing extreme value copulas to estimate the quantile”. *Statistics and Operations Research Transactions, SORT*, 38, 89-102.
- [3].Bargès, M., Cossette, H., Marceau, É. (2009) “TVaR-based capital allocation with copulas”. *Insurance: Mathematics and Economics*, 45, 348- 361.
- [4].Bolancé, C., Bahraoui, Z. y Artís, M. (2014). “Quantifying the risk using copulae with nonparametric marginals”. *Insurance: Mathematics and Economics*, 58, 46-56.
- [5].Genest, C. y Rivest, L. (1993) “Statistical inference procedure for bivariate archimedean copulas” *Journal of the American Statistical Association*, 88, 1034-1043.
- [6].Gronneberg, S. y Hjort, N. (2014) “The copula information criteria”. *Scandinavian Journal of Statistics*, 41, 436-459.
- [7].Nelsen, R. (2006) *An Introduction to Copulas*. Springer, USA.
- [8].Sklar, A. (1959) “Fonctions de répartition á n dimensions et leurs marges”. *Publications de l’istitut de Statistique de l’université de Paris*, 8, 229-231.
- [9].McNeil A. J., Frey R., Embrechts P. (2006) *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press.

UB-Riskcenter Working Paper Series

List of Published Working Papers

- [WP 2014/01]. Bolancé, C., Guillén, M. and Pitt, D. (2014) “Non-parametric models for univariate claim severity distributions – an approach using R”, UB Riskcenter Working Papers Series 2014-01.
- [WP 2014/02]. Mari del Cristo, L. and Gómez-Puig, M. (2014) “Dollarization and the relationship between EMBI and fundamentals in Latin American countries”, UB Riskcenter Working Papers Series 2014-02.
- [WP 2014/03]. Gómez-Puig, M. and Sosvilla-Rivero, S. (2014) “Causality and contagion in EMU sovereign debt markets”, UB Riskcenter Working Papers Series 2014-03.
- [WP 2014/04]. Gómez-Puig, M., Sosvilla-Rivero, S. and Ramos-Herrera M.C. “An update on EMU sovereign yield spread drivers in time of crisis: A panel data analysis”, UB Riskcenter Working Papers Series 2014-04.
- [WP 2014/05]. Alemany, R., Bolancé, C. and Guillén, M. (2014) “Accounting for severity of risk when pricing insurance products”, UB Riskcenter Working Papers Series 2014-05.
- [WP 2014/06]. Guelman, L., Guillén, M. and Pérez-Marín, A.M. (2014) “Optimal personalized treatment rules for marketing interventions: A review of methods, a new proposal, and an insurance case study.”, UB Riskcenter Working Papers Series 2014-06.
- [WP 2014/07]. Piulachs, X., Alemany, R. and Guillén, M. (2014) “A joint longitudinal and survival model with health care usage for insured elderly”, UB Riskcenter Working Papers Series 2014-07.
- [WP 2014/08]. Abad, P, Chuliá, H. (2014) “European government bond market integration in turbulent times”, UB Riskcenter Working Papers Series 2014-08.
- [WP 2014/09]. Belles-Sampera, J., Guillén, M. and Santolino, M. (2014) “The use of flexible quantile-based measures in risk assessment”, UB Riskcenter Working Papers Series 2014-09.
- [WP 2015/01]. Bolancé C., Guillén, M. and Padilla, A. (2015) “Estimación del riesgo mediante el ajuste de cópulas”, UB Riskcenter Working Papers Series 2015-01.