Miscelánea Matemática 48 (2009) 7–28 SMM

Cópulas y dependencia de variables aleatorias: Una introducción

Arturo Erdely

Deparatmento de Matemáticas Aplicadas y Sistemas
Universidad Autónoma Metropolitana-Cuajimalpa
México

aerdely@correo.cua.uam.mx

Resumen

Se presenta una introducción a conceptos y resultados básicos sobre cópulas, y su utilidad para estudiar y medir dependencia de variables aleatorias, como consecuencia del Teorema de Sklar (1959).

1. Motivación

Sean f y g funciones de densidad de probabilidades, esto es $f, g \geq 0$ y $\int_{\mathbb{R}} f(x)\,dx = 1$, $\int_{\mathbb{R}} g(y)\,dy = 1$, y sean F y G sus correspondientes funciones de distribución de probabilidades, respectivamente, esto es $F(x) = \int_{-\infty}^{x} f(t)\,dt$ y $G(y) = \int_{-\infty}^{y} g(t)\,dt$. Si definimos la familia paramétrica de funciones bivariadas $h_\theta : \mathbb{R}^2 \to \mathbb{R}$ de la siguiente manera

$$h_\theta(x,y) := f(x)g(y)\big[\,1 + \theta(1-2F(x))(1-2G(y))\,\big]\,, \tag{1}$$

es inmediato verificar que $\{h_\theta\,:\,-1 \leq \theta \leq 1\}$ es una familia de funciones de densidad de probabilidad conjunta con densidades marginales f y g, esto es $h_\theta \geq 0$ y

$$\iint_{\mathbb{R}^2} h_\theta(x,y)\,dxdy = 1\ ,\ \int_{\mathbb{R}} h_\theta(x,y)\,dy = f(x)\ ,\ \int_{\mathbb{R}} h_\theta(x,y)\,dx = g(y)\,.$$

Lo anterior es un ejemplo típico para demostrar que el sólo conocimiento de las distribuciones marginales de un vector aleatorio, en general, no es suficiente para determinar la distribución conjunta, aún cuando las distribuciones marginales se pueden obtener a partir de la distribución conjunta. A h_θ, definida en (1), se le conoce como la familia *Farlie-Gumbel-Morgenstern* de distribuciones y fue estudiada por Morgenstern (1956), Gumbel (1958, 1960), y Farlie (1960). De acuerdo a Kotz y Seeger (1991):

> En la última década [1980's] se ha vuelto cada vez más importante el considerar la dependencia de variables aleatorias más allá de una simple antítesis del concepto de independencia, siendo este último un concepto fundamental en la teoría matemática de la probabilidad. Como consecuencia, se han desarrollado diversos métodos para imponer ciertos tipos de dependencia a variables aleatorias con distribuciones marginales dadas. La mayoría de dichos métodos están basados en un conocido resultado de Hoeffding (1940) y Fréchet (1951) que establece que para cualquier par de variables aleatorias X, Y, con funciones de distribución $F_1(x)$ y $F_2(y)$, respectivamente, que la clase $\Psi(F_1, F_2) = \{H(x,y) \,|\, H$ es función de distribución conjunta con marginales $F_1(x)$ y $F_2(y)\}$, contiene una cota superior, H^*, y una cota inferior H_*. Éstas son cotas respecto al orden parcial $\prec$, que denota dominancia estocástica, esto es, si $H, H' \in \Psi(F_1, F_2)$ entonces $H \prec H'$ si y sólo si $H(x,y) \leq H'(x,y)$ para todo (x,y). Más aún, las conocidas cotas de Fréchet tienen expresiones generales en términos de F_1 y F_2, como sigue
>
> $$H_*(x,y) = \max\{F_1(x) + F_2(y) - 1, 0\}\,, \tag{2}$$
>
> $$H^*(x,y) = \min\{F_1(x), F_2(y)\}\,. \tag{3}$$
>
> [...] Diversos subconjuntos parametrizados de $\Psi(F_1, F_2)$ que están linealmente ordenados con respecto a $\prec$ [esto es que para cualesquiera $H, H' \in \Psi(F_1, F_2)$ se cumple alguna de las siguientes: $H \prec H'$, $H' \prec H$, o $H = H'$] han aparecido en la literatura [...] tal es el caso de la familia Farlie-Gumbel-Morgenstern, entre otras [...] y fueron construidos de acuerdo al punto de vista de que la forma de imponer dependencia es incrementar (o decrementar) en todos lados la función de distribución conjunta $F_1(x)F_2(y)$, correspondiente a la independencia, sin alterar las distribuciones marginales, creando así una nueva función de distribución conjunta más cercana en valores a H^* (o H_*) [...]

Fréchet (1951, 1957) y Féron (1956) hicieron importantes contribuciones en relación al problema de determinar la relación entre una función de distribución de probabilidades multivariada y sus distribuciones marginales de menor dimensión. Una solución concreta a este problema surgió como resultado de la colaboración entre Abe Sklar y

Berthold Schweizer. De acuerdo al mismo Schweizer (1991) su colaboración comenzó en el contexto de espacios métricos probabilísticos en 1957. Para 1958 tenían ya algunos resultados importantes que enviaron a Maurice Fréchet, y así dio inicio un intercambio de correspondencia entre ellos. En una de sus cartas, Fréchet les propuso estudiar el problema de determinar la relación entre la función de distribución de probabilidades mutlivariada y sus distribuciones marginales de menor dimensión.

Abe Sklar dio respuesta al problema propuesto por Fréchet para el caso de distribuciones marginales unidimensionales. Por ejemplo, sea (X, Y) un vector aleatorio con función de distribución conjunta $H(x, y)$, entonces las funciones de distribución marginal de X, Y están dadas por $F(x) := H(x, \infty)$ y $G(y) := H(\infty, y)$, respectivamente. Sklar (1959) demostró mediante el ahora célebre teorema que lleva su apellido, que existe una función C, que denominó *cópula*, que establece la relación funcional entre la distribución conjunta y sus marginales unidimensionales

$$H(x, y) = C\big(F(x), G(y)\big)\,. \tag{4}$$

Antes de revisar la definición formal de cópula, podemos proporcionar una motivación utilizando propiedades básicas de las funciones de distribución de probabilidades bivariadas:

$$H(\infty, \infty) = 1\,, \quad H(x, \infty) = F(x)\,, \quad H(\infty, y) = G(y)\,,$$
$$H(x, -\infty) = 0 = H(-\infty, y)\,;$$

y para cualesquiera números reales x_1, x_2, y_1, y_2 en donde $x_1 \leq x_2$ y $y_1 \leq y_2$ tenemos que

$$H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1) = \\ \mathbb{P}\big[\, x_1 < X \leq x_2\,,\ y_1 < Y \leq y_2 \,\big] \geq 0\,.$$

Utilizando las propiedades anteriores en combinación con (4) notemos que

$$C(F(x), 1) = F(x)\,, C(1, G(y)) = G(y)\,, C(F(x), 0) = 0 = C(0, G(y))\,,$$

$$C(F(x_2), G(y_2)) - C(F(x_2), G(y_1)) \\ - C(F(x_1), G(y_2)) + C(F(x_1), G(y_1)) \geq 0\,,$$

para cualesquiera números reales x_1, x_2, y_1, y_2 tales que $x_1 \leq x_2\,,\ y_1 \leq y_2$. Si definimos $u := F(x)$ y $v := G(y)$, llegamos a la definición formal de cópula que se presenta en la siguiente sección, tal cual se encuentra, por ejemplo, en Schweizer y Sklar (2005) o en Nelsen (2006).

2. Conceptos y resultados básicos

2.1. Definición. Una *cópula bivariada* es una función $C : \mathbb{I}^2 \to \mathbb{I}$ con las siguientes propiedades:

1. Para cualesquiera u, v en $\mathbb{I} := [\,0, 1\,]$
$$C(u,0) = 0 = C(0,v)\,, \tag{5}$$
$$C(u,1) = u\,,\; C(1,v) = v\,; \tag{6}$$
2. Para cualesquiera u_1, u_2, v_1, v_2 en $\mathbb{I}$ tales que $u_1 \leq u_2$ y $v_1 \leq v_2$
$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0\,. \tag{7}$$

Aunque en la sección anterior dimos una motivación probabilística para la definición de cópula, es importante notar que, en sentido estricto, la Definición 2.1 no involucra en absoluto conceptos probabilísticos, se trata simplemente de una función real con dominio igual al cuadrado unitario, que satisface ciertas condiciones de frontera (5) y (6), y la condición (7) que es conocida como propiedad *2-creciente*. Un comentario similar es válido para la versión formal del *Teorema de Sklar*, como lo veremos más adelante.

A partir de la definición anterior se demuestra que C es monótona creciente en cada variable (sea $v_1 = 0$ o $u_1 = 0$ en (7)) y uniformemente continua (ya que la Definición 2.1 implica que C satisface la condición de Lipschitz $|C(u_2,v_2) - C(u_1,v_1)| \leq |u_2 - u_1| + |v_2 - v_1|$), para detalles de éstas y otras propiedades ver Nelsen (1995 y 2006) o Schweizer y Sklar (2005). Consecuencia inmediata de lo anterior es que las secciones horizontales, verticales y diagonal de una cópula C son todas monótonas crecientes y uniformemente continuas en $\mathbb{I} = [\,0, 1\,]$, en donde la *sección horizontal, sección vertical* y *sección diagonal* son funciones de $\mathbb{I}$ a $\mathbb{I}$ dadas por $t \mapsto C(t,a)$, $t \mapsto C(a,t)$, y $\delta_C(t) = C(t,t)$, respectivamente, con a fija en $\mathbb{I}$.

2.2. Teorema. *Sea C una cópula. Entonces para todo (u,v) en $\mathbb{I}^2$*
$$\text{máx}\,(u + v - 1, 0) \leq C(u,v) \leq \text{mín}\,(u,v)\,. \tag{8}$$

Es inmediato verificar que las cotas en (8) son en sí mismas cópulas, y son comúnmente denotadas por $M(u,v) := \text{mín}\,(u,v)$ y $W(u,v) := \text{máx}\,(u+v-1, 0)$. Así, para cualquier cópula C y cualquier (u,v) in $\mathbb{I}^2$ se tiene que
$$W(u,v) \leq C(u,v) \leq M(u,v)\,. \tag{9}$$

La doble desigualdad anterior es la versión para cópulas de las *cotas de Fréchet-Hoeffding* (2) y (3). Nos referiremos a M como la *cota superior de Fréchet-Hoeffding,* y a W como la *cota inferior de Fréchet-Hoeffding.* Una tercera cópula importante, como se verá más adelante, es la denominada *cópula producto,* definida mediante $\Pi(u,v) := uv$. Otra propiedad importante y fácil de demostrar es la siguiente:

2.3. Proposición. *Para cualesquiera cópulas bivariadas C_0 y C_1 se tiene que cualquier combinación lineal convexa de ellas es cópula, esto es que*

$$(1-\alpha)C_0 + \alpha C_1$$

es cópula, para cualquier $\alpha \in [\,0,1\,]$.

2.4. Definición. Una *función de distribución* es una función F con dominio en el conjunto extendido de los números reales $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$, tal que

1. F es monótona creciente,
2. $F(-\infty) = 0$ y $F(\infty) = 1$.

2.5. Definición. Una *función de distribución conjunta bivariada* es una función H cuyo dominio es el plano extendido $\overline{\mathbb{R}}^2 := \overline{\mathbb{R}} \times \overline{\mathbb{R}}$, tal que

1. H es 2-creciente, esto es que para cualesquiera x_1, x_2, y_1, y_2 en $\overline{\mathbb{R}}$ tales que $x_1 \leq x_2$, $y_1 \leq y_2$ se cumple
$$H(x_2,y_2) - H(x_2,y_1) - H(x_1,y_2) + H(x_1,y_1) \geq 0\,,$$
2. $H(x,-\infty) = 0 = H(-\infty,y)$ y $H(\infty,\infty) = 1$.

Respecto a las dos definiciones anteriores, una importante cita del libro de Nelsen (2006):

> Notemos que las definiciones anteriores de funciones de distribución no tienen [en estricto sentido] nada de probabilístico. No se hace referencia a variables aleatorias, ni a continuidad por la derecha o por la izquierda. Por supuesto que todas las funciones de distribución de probabilidades de una o dos variables aleatorias satisfacen alguna de las dos definiciones anteriores, así que cualquier resultado que se demuestre para dichas funciones de distribución podrá inmediatamente aplicarse, en particular, a funciones de distribución de probabilidades.

2.6. Definición. Las *marginales* de una función de distribución conjunta bivariada H son las funciones F y G obtenidas mediante $F(x) := H(x,\infty)$ y $G(y) := H(\infty,y)$.

Consecuencia inmediata de las definiciones anteriores es que las marginales de H son en sí mismas funciones de distribución.

2.7. Teorema. ***(Sklar, 1959)*** *Sea H una función de distribución conjunta bivariada con marginales F y G . Entonces existe una cópula C tal que para cualesquiera x,y en $\overline{\mathbb{R}}$*

$$H(x,y) \,=\, C\big(F(x),G(y)\big)\,. \tag{10}$$

Si F y G son continuas, entonces C es única; en cualquier otro caso, C sólo está determinada de forma única sobre el conjunto $Ran\,F \times Ran\,G$. También, si C es una cópula dada, F y G funciones de distribución dadas, entonces la función H definida mediante (10) es una función de distribución conjunta bivariada con marginales F y G .

Notación: $Ran\,F := \{z \,:\, z = F(x) \text{ para algún } x\}$. Para los detalles de la demostración ver Sklar (1959), Schweizer y Sklar (2005), Nelsen (2006) o Carley y Taylor (2002).

2.8. Definición. Sea F una función de distribución. Una *cuasi-inversa* de F es cualquier función $F^{(-1)}$ con dominio $\mathbb{I} = [\,0,1\,]$ tal que

1. Si $t \in \,Ran\,F$, entonces $F^{(-1)}(t)$ es cualquier $x \in \overline{\mathbb{R}}$ tal que $F(x) = t$, esto es, para cualquier $t \in \,Ran\,F$

$$F\big(F^{(-1)}(t)\big) = t\,;$$

2. Si $t \notin \,Ran\,F$, entonces

$$F^{(-1)}(t) = \text{ínf}\{x \,|\, F(x) \geq t\} = \sup\{x \,|\, F(x) \leq t\}\,.$$

Si F es, en particular, estrictamente creciente, entonces tiene una única cuasi-inversa, que de hecho es $F^{(-1)} = F^{-1}$, en donde F^{-1} es la inversa usual.

2.9. Corolario. *Sea H una función de distribución conjunta bivariada con marginales continuas F y G , y sea C la única cópula tal que (10) se cumple. Entonces para cualquier $(u,v) \in \mathbb{I}^2$*

$$C(u,v) \,=\, H\big(F^{(-1)}(u),G^{(-1)}(v)\big)\,. \tag{11}$$

En el contexto particular de variables aleatorias, si X, Y son variables aleatorias continuas con distribuciones de probabilidad marginal F y G , respectivamente, y con función de distribución conjunta de probabilidades H, el Teorema de Sklar implica que existe una única cópula, que podemos denotar C_{XY}, tal que $H(x,y) = C_{XY}\big(F(x), G(y)\big)$. Más aún, si C es cualquier cópula y F_1 y G_1 son funciones de distribución de probabilidades (marginales), entonces $C\big(F_1(x), G_1(y)\big)$ es también una función de distribución conjunta de probabilidades. Como se menciona en Mikusiński *et al.* (1991):

> La cópula C contiene información valiosa acerca del tipo de dependencia que existe entre variables aleatorias que tengan a C como su cópula subyacente [...] Puede uno pensar en la cópula como el representante canónico de todas las distribuciones conjuntas H que pudieran corresponder a las variables aleatorias X, Y con cierto tipo particular de dependencia entre ellas.

Como observación particular importante, tenemos que si U y V son variables aleatorias continuas uniformes en el intervalo abierto $]\,0,1\,[$, entonces por el Teorema de Sklar tenemos que su función de distribución conjunta de probabilidades H restringida al cuadrado unitario $\mathbb{I}^2$ es igual a la cópula subyacente C_{UV}. Dicho de otro modo, de acuerdo a Nelsen (2006):

> [...] cada cópula C induce una medida de probabilidad en $\mathbb{I}^2$ [...] Así, a un nivel intuitivo, la C-medida de un subconjunto de $\mathbb{I}^2$ es la probabilidad de que dos variables aleatorias continuas uniformes en $]\,0,1\,[$, U y V, con función de distribución conjunta C, tomen valores que pertenezcan a dicho subconjunto.

Dada una función de distribución conjunta de probabilidades bivariada H, correspondiente a un par de variables aleatorias continuas con funciones de distribución (marginal) de probabilidades F y G, podemos "extraer" la cópula subyacente C_{XY} mediante el Corolario 2.9, y luego construir una nueva función de distribución conjunta de probabilidades bivariada H_1, con la misma cópula pero distintas funciones de distribución (marginal) de probabilidades F_1 y G_1, esto es $H_1(x,y) = C_{XY}\big(F_1(x), G_1(y)\big)$. Así, por ejemplo, el ejercicio de construir una función de distribución conjunta de probabilidades bivariada distinta a la normal bivariada, pero que tenga marginales normal estándar, es inmediato. Así también, podemos “extraer” la cópula subyacente de la distribución normal bivariada, (denominada *cópula gaussiana*) y utilizarla para construir una nueva función de distribución conjunta de probabilidades bivariada con la misma estructura de dependencia que la normal bivariada, pero con marginales continuas arbitrarias, no necesariamente normales. Para lo anterior resulta crucial la continuidad de las marginales, de otro modo es importante tomar en cuenta una advertencia de Marshall (1996):

> Las marginales F y G pueden ser insertadas en cualquier cópula, así que no tienen información directa sobre el apareamiento; al mismo tiempo, cualquier par de marginales puede ser insertado en [una cópula] C, por lo que C no tiene información directa acerca de las marginales. Siendo éste el caso, puede parecer razonable esperar que las conexiones entre las marginales de H estén sólo determinadas por C, y que cualquier pregunta en relación a dicha conexión debiera ser resuelta con el sólo conocimiento de C.
>
> Por supuesto, las cosas no son así de simples. Algunos problemas surgen del hecho de que las cópulas no son únicas cuando al menos una de las marginales es discontinua. De hecho, las marginales pueden en ocasiones llegar a jugar un papel tan importante como el de la cópula para determinar la forma en cómo serán transformadas para obtener la distribución conjunta H; en un caso extremo, marginales degeneradas pueden por sí mismas determinar la distribución conjunta bajo cualquier cópula. Pero la interacción entre la cópula y las marginales es por lo general muy relevante [...]
>
> [...] Mucha de la literatura sobre cópulas está basada en el supuesto de que las marginales de H son continuas porque es ésta una condición necesaria y suficiente para la unicidad de la cópula subyacente.

Pero aún en el contexto de variables aleatorias continuas, es tentador considerar al Teorema de Sklar como una fuente infinita de construcción de distribuciones multivariadas de todo tipo con tan sólo escoger e insertar marginales continuas en cualquier cópula que se desee. De alguna manera, Marshall y Olkin (1967) hacen énfasis en la importancia de hacer trabajo adicional en el momento de construir una distribución multivariada:

[. . .] La familia de soluciones

$$H(x,y) = F(x)G(y)\{1-\alpha[1-F(x)][1-G(y)]\}\,, \quad |\alpha| \leq 1\,, \tag{12}$$

atribuida a Morgenstern (1956) ha sido estudiada por Gumbel (1960) cuando F y G son [funciones de distribución de probabilidades] exponenciales. Gumbel también estudió la distribución bivariada

$$H(x,y) = 1 - e^{x} - e^{y} + e^{-x-y-\delta xy}\,, \quad 0 \leq \delta \leq 1\,, \tag{13}$$

que también tiene marginales exponenciales. Sin embargo, [hasta el momento] no tenemos noticia de modelo alguno o base alguna para determinar cómo este tipo de distribuciones pudieran surgir en la práctica.

Por cierto, con marginales continuas F y G, y aplicando el Teorema de Sklar en (12), de inmediato identificamos la familia paramétrica subyacente de cópulas

$$C_\theta(u,v) = uv\big[\,1-\theta(1-u)(1-v)\,\big]\,, \quad |\theta| \leq 1\,, \tag{14}$$

conocida como la familia *Farlie-Gumbel-Morgenstern* de cópulas. Más aún, el lector puede verificar que (12), junto con el supuesto de que F y G sean absolutamente continuas, tiene por función de densidad de probabilidades conjunta a (1).

Sería maravilloso tener un extenso catálogo que describa la interpretación probabilística que tiene cada cópula, como de algún modo lo exigen Marshall y Olkin (1967). En la actualidad se cuenta con muchísimas más familias de cópulas, que interpretaciones probabilísticas de las mismas. El mismo Roger B. Nelsen, autor del célebre libro *An Introduction to Copulas* comentó en una ocasión a quien este artículo escribe, que cada mes sin falta le llegan mensajes electrónicos de todo el mundo informándole acerca de la construcción de una nueva familia de cópulas, pero como el mismo Nelsen me mencionó, si eso no va acompañado de una intepretación probabilística, será una cópula más que se agrega al catálogo de cópulas de aplicación desconocida.

Se cuenta con la interpretación probabilística de algunas cópulas. Por ejemplo, Fréchet (1951) demostró que $M(u,v) = \text{mín}(u,v)$ es la cópula de un vector (X,Y) de variables aleatorias continuas si y sólo si X,Y son casi seguramente funciones creciente una de la otra, y que $W(u,v) = \text{máx}(u+v-1,0)$ es la cópula para (X,Y) si y sólo si X,Y

son casi seguramente funciones decrecientes una de la otra. La cópula producto $\Pi(u,v) = uv$ está asociada a variables aleatorias independientes, y comentaremos al respecto en la siguiente sección.

La mayoría de las definiciones y resultados para el caso bivariado se extienden a dimensiones mayores a 2, para detalles consúltese Schweizer y Sklar (2005) y Nelsen (2006), por lo que sólo haremos algunos comentarios acerca de problemas que no surgen en dimensión 2.

Las extensiones de las cópulas bivariadas M , Π , y W a dimensión $n \geq 3$ están dadas por

$$\begin{aligned} M^{(n)}(u_1,\ldots,u_n) &:= \min(u_1,\ldots,u_n)\,; \\ \Pi^{(n)}(u_1,\ldots,u_n) &:= u_1 u_2 \cdots u_n\,; \\ W^{(n)}(u_1,\ldots,u_n) &:= \max(u_1+u_2+\cdots+u_n-n+1,0)\,. \end{aligned} \tag{15}$$

$M^{(n)}$ y $\Pi^{(n)}$ son cópulas multivariadas (o n-cópulas) para todo $n \geq 2$, pero $W^{(n)}$ no es n-cópula para cualquier $n > 2$. Aún así, seguimos teniendo una versión n-dimensional de las cotas de Fréchet-Hoeffding (9): Si C es cualquier n-cópula, entonces para cualquier $\mathbf{u} = (u_1,\ldots,u_n) \in \mathbb{I}^n := [\,0,1\,]^n$

$$W^{(n)}(\mathbf{u}) \leq C(\mathbf{u}) \leq M^{(n)}(\mathbf{u})\,. \tag{16}$$

Aunque la cota inferior de Fréchet-Hoeffding $W^{(n)}$ nunca es cópula para $n > 2$, la desigualdad anterior no puede ser mejorada, ver Nelsen (2006), ya que:

2.10. Teorema. *Para cualquier $n \geq 3$ y cualquier $\mathbf{u} = (u_1,\ldots,u_n) \in \mathbb{I}^n$ fijo, existe una n-cópula $C_{\mathbf{u}}$, dependiente de $\mathbf{u}$, de modo que $C_{\mathbf{u}}(\mathbf{u}) = W^{(n)}(\mathbf{u})$.*

Con $n = 2$ sólo se tienen marginales univariadas, las cuales, en virtud de la Definición 2.1, son funciones identidad, esto es $C(u,1) = u$ y $C(1,v) = v$. Para n-cópulas, $n \geq 3$, el concepto de m-marginales se vuelve necesario, $2 \leq m \leq n$, definiendo como m-marginal a la función obtenida fijando $n - m$ de los argumentos de C iguales a 1. Una n-cópula tiene así $\binom{n}{m}$ m-marginales. Es inmediato verificar que cada m-marginal de C es una m-cópula. Sin embargo, dadas $\binom{n}{m}$ m-cópulas, no se puede garantizar que exista una n-cópula que las tenga como sus m-marginales; en caso de que así fuera, se dice entonces que las m-cópulas son *compatibles.* El determinar cuándo ciertas m-cópulas dadas pueden o no ser m-marginales de cópulas de mayor dimensión es conocido como el *problema de compatibilidad*. Sklar (1996) da algunos ejemplos de *incompatibilidad*. Existen resultados sobre condiciones necesarias y suficientes para la compatibilidad de 3-cópulas: Dall'Aglio (1959, 1960, 1972), Quesada-Molina y Rodríguez-Lallena (1994), Joe (1997). Para algunos casos en dimensiones superiores, ver Joe (1996, 1997).

3. El problema de medir dependencia de variables aleatorias

Este problema amerita todo un artículo aparte, pero al menos mencionaremos algunas ideas fundamentales. Comenzamos con una cita de Drouet y Kotz (2001):

> El concepto de dependencia aparece por todas partes en nuestra Tierra y sus habitantes de manera profunda. Son innumerables los ejemplos de fenómenos meteorológicos interdependientes en la naturaleza, o de interdependencia en aspectos médicos, sociales, políticos y económicos de nuestra existencia. Más aún, la dependencia es obviamente no determinística, sino de naturaleza estocástica.
>
> Es por lo anterior que resulta sorprendente que conceptos y medidas de dependencia no hayan recibido suficiente atención en la literatura estadística, al menos hasta 1966 cuando el trabajo pionero de E.L. Lehmann entró en escena. El concepto de correlación (y sus modificaciones) introducido por F. Galton en 1885 ha dominado la estadística durante unos 70 años del siglo XX, sirviendo prácticamente como la única medida de dependencia generalmente aceptada, a pesar de resultar en ocasiones claramente una medida inapropiada. La última parte del siglo XX ha sido testigo de un rápido resurgimiento en investigaciones sobre dependencia desde los puntos de vista probabilístico y estadístico. El primer libro, hasta donde sabemos, que ha sido dedicado a conceptos de dependencia apareció bajo la autoría de Harry Joe (1997). Más aún, pareciera ser que no hay departamento de matemáticas o estadística en Estados Unidos o Europa que ofrezca cursos especialmente dedicados a estudiar conceptos y medidas de dependencia.

Recordemos que la *correlación* (lineal o de Pearson) entre dos variables aleatorias X, Y se define mediante

$$r(X,Y) := \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}}\,,$$

suponiendo existencia de segundos momentos no nulos. Antes de citar algunos comentarios en relación a la correlación así definida, citamos una propiedad de cópulas de Nelsen (2006):

3.1. Teorema. *Sean X, Y variables aleatorias continuas con cópula C_{XY}. Si α y β son funciones estrictamente crecientes en $\mathrm{Ran}\, X$ y $\mathrm{Ran}\, Y$, respectivamente, entonces $C_{\alpha(X)\beta(Y)} = C_{XY}$. Esto es, C_{XY} es invariante bajo transformaciones estrictamente crecientes de X, Y.*

Este resultado es también válido para dimensiones mayores a 2, ver Schweizer y Sklar (2005). Este teorema en conexión con el concepto de correlación mereció el siguiente comentario de Embrechts *et al.* (2003):

> Las cópulas proveen una manera natural de estudiar y medir dependencia entre variables aleatorias. Como consecuencia directa del Teorema 3.1, las propiedades de las cópulas son invariantes bajo transformaciones estrictamente crecientes de las variables aleatorias involucradas. La correlación lineal (o de Pearson) es frecuentemente utilizada en la práctica como medida de dependencia. Sin embargo, como la correlación lineal no es una medida basada en la cópula subyacente, en ocasiones conduce a resultados aberrantes y por tanto no debiera tomarse como la medida canónica de dependencia [...] La popularidad de la correlación lineal se explica por la facilidad con la que puede ser calculada y por ser una medida escalar de dependencia que de forma natural surge en el contexto de las distribuciones elípticas (con conocidos miembros como lo son las distribuciones normal y t multivariadas). El problema es que a muchas variables aleatorias no les corresponde una distribución elíptica, y utilizar en tal caso la correlación lineal usualmente conduce a conclusiones verdaderamente alejadas de la realidad. Y aún cuando resulte razonable utilizar como distribución conjunta a un miembro de la familia elíptica, hay situaciones en las que utilizar la correlación lineal [...] no tiene sentido; por ejemplo, en el caso de distribuciones de colas pesadas: en tales casos ni siquiera está definida la correlación lineal ante la presencia de segundos momentos infinitos.

Más aún, encontramos en Embrechts *et al.* (1999) una lista de observaciones a tomar en cuenta al utilizar correlación lineal:

> La correlación lineal es un campo minado para el desprevenido. No hace falta esforzarse mucho para encontrar en la literatura sobre administración de riesgos financieros enorme confusión y malinterpretación. Esto es preocupante ya que la correlación lineal es un concepto técnico fundamental en finanzas [...] :
>
> 1. La correlación lineal es simplemente una medida escalar de dependencia; no puede decirnos todo lo que quisiéramos saber en relación a la estructura de dependencia de los riesgos.
> 2. Los valores posibles de la correlación lineal dependen de las distribuciones marginales de los riesgos. Todos los valores entre -1 y $+1$ no son necesariamente alcanzables.
> 3. Riesgos con una dependencia positiva perfecta no tienen necesariamente una correlación lineal igual a $+1$; riesgos con una dependencia negativa perfecta no tienen necesariamente una correlación lineal igual a -1.
> 4. Correlación lineal igual a cero no implica independencia de los riesgos.
> 5. La correlación lineal no es invariante bajo transformaciones crecientes de los riesgos. Por ejemplo, $\log X$ y $\log Y$ por lo general no tienen la misma correlación lineal que X, Y.
> 6. La correlación sólo está definida cuando las varianzas de los riesgos son finitas, así que no es una medida de dependencia adecuada para riesgos de colas pesadas en donde las varianzas son infinitas.

Entrar a fondo en una discusión sobre cómo medir dependencia será motivo de un artículo posterior, así que por el momento sólo citaremos algunas propuestas de propiedades deseables para medidas de dependencia en el caso de variables aleatorias continuas.

La primera tiene que ver con una propuesta de Drouet y Kotz (2001): una medida de dependencia debiera ser independendiente de las distribuciones marginales, es decir, sólo basada en la cópula subyacente. De hecho, esta propuesta coincide con lo ya citado de Embrechts *et al.* (1999 y 2003). Esta propuesta surge como consecuencia del Teorema de Sklar ya que en el caso de variables aleatorias continuas, la estructura de dependencia queda determinada de manera única por la cópula subyacente. De aceptar esta propuesta entonces la correlación lineal (o de Pearson) quedaría excluida como medida de dependencia por dos razones: primero, depende de la existencia de segundos momentos finitos de las distribuciones marginales; segundo, porque aún bajo la existencia de segundos momentos finitos, como consecuencia del Corolario 2.9 se tiene que al tratar de calcular la correlación lineal en términos de la cópula subyacente se obtiene

$$r(X,Y) = \frac{1}{\sqrt{\mathbb{V}(X)\mathbb{V}(Y)}} \iint_{\mathbb{I}^2} \big[\, C(u,v) - uv \,\big]\, dF^{-1}(u) dG^{-1}(v)\,. \quad (17)$$

Esto es, la correlación lineal no es solo función de la cópula sino también del comportamiento marginal de las variables aleatorias involucradas, lo que implica que, con una misma estructura de dependencia, es posible tener distintos valores de correlación lineal tan solo modificando una o ambas marginales.

La segunda propuesta de propiedad deseable de una medida de dependencia para variables aleatorias continuas tiene que ver con pedir que dicha medida reporte el valor cero si y sólo si las variables involucradas son independientes. Esta propuesta aparece en Rényi (1959) y en Schweizer y Wolff (1981). Esta condición no se cumple en el caso de la correlación lineal y de las *medidas de concordancia,* como lo son la correlación de Kendall (1938), la correlación de Spearman (1904), el coeficiente de Gini (1914) y el de Blomqvist (1950), mismas que, a pesar de que pueden calcularse solamente en términos de la cópula subyacente, no caracterizan la independencia.

¿Cómo construir entonces medidas de dependencia que caractericen la independencia de variables aleatorias continuas? Consecuencia inmediata del Teorema de Sklar es la siguiente: para un vector (X,Y) de variables aleatorias continuas la cópula subyacente es la cópula producto $\Pi(u,v) = uv$ si y sólo si X,Y son independientes. Es por ello que Schweizer y Sklar (2005) sugieren medir de algún modo qué tan lejos/cerca está la cópula correspondiente respecto de la cópula Π que representa la independencia:

> [...] el hecho de que la superficie de Π se encuentre entre las superficies de W y M (las cópulas para el caso extremo de dependencia estrictamente monótona) es natural utilizar cualquier medida de distancia entre las superficies como medida de dependencia para parejas de variables aleatorias.

Schweizer y Wolff (1981) proponen en concreto la siguiente definición de medida de dependencia:

3.2. Definición. Sean X, Y cualesquiera variables aleatorias continuas con cópula subyacente C. Una medida numérica de asociación para este par de variables aleatorias es *medida de dependencia,* denotada mediante $\mu_{X,Y}$ o μ_C, si satisface las siguiente propiedades:

1. $\mu_{X,Y}$ está definida para cualquier par de variables aleatorias continuas;

2. $0 \leq \mu_{X,Y} \leq 1$;

3. $\mu_{X,Y} = \mu_{Y,X}$;

4. X, Y son independientes si y sólo si $\mu_{X,Y} = \mu_\Pi = 0$;

5. $\mu_{XY} = 1$ si y sólo si X, Y son cada una casi seguramente función estrictamente monótona de la otra; es decir, si C_{XY} es igual a M o W.

6. Si α y β son casi seguramente funciones estrictamente monótonas en $Ran\, X$ y $Ran\, Y$, respectivamente, entonces $\mu_{\alpha(X),\beta(Y)} = \mu_{X,Y}$;

7. Si $\{(X_n, Y_n)\}$ es una sucesión de variables aleatorias continuas con cópulas subyacentes C_n, respectivamente, y si $\{C_n\}$ converge puntualmente a C, entonces $\lim_{n\to\infty} \mu_{C_n} = \mu_C$.

Agregan además Schweizer y Wolff (1981) que

> [. . .] cualquier medida de distancia, adecuadamente estandarizada, entre las superficies $z = C(u, v)$ y $z = uv$, por ejemplo, cualquier distancia L_p, debiera proporcionar una medida simétrica y no paramétrica de dependencia.

Para cualquier $p \in [1, \infty[$ la distancia L_p entre una cópula C y Π está dada por

$$\mho_C(p) := \left(k_p \iint_{\mathbb{I}^2} |C(u, v) - uv|^p du dv\right)^{1/p}, \tag{18}$$

en donde k_p es una constante escogida de modo que (18) asigne el valor $\mho_C(p) = 1$ cuando C sea igual a M o W. De Nelsen (2006) se tiene el siguiente:

3.3. Teorema. *Para cada $p \in [\,1,\infty\,[$ la distancia L_p calculada mediante $\mho_C(p)$ es una medida de dependencia, en donde*

$$k_p = \frac{\Gamma(2p+3)}{2\big[\,\Gamma(p+1)\,\big]^{\,2}}\,. \tag{19}$$

Casos particulares estudiados son: $\mho_C(1)$ mejor conocida como la medida de dependencia de Schweizer y Wolff (1981), y $[\,\mho_C(2)\,]^2$ mejor conocido como el índice de dependencia de Hoeffding (1940). En el caso particular de la medida de dependencia de Schweizer y Wolff dada por

$$\sigma_C := 12\iint_{\mathbb{I}^2} |\,C(u,v) - uv\,|\,dudv\,, \tag{20}$$

podemos interpretarla como una medida del "valor absoluto promedio" entre la distribución conjunta de un par de variables aleatorias continuas (representada por su cópula C) y la distribución conjunta de un par de variables aleatorias continuas independientes (representada por $\Pi(u,v) = uv$).

Schweizer y Wolff (1981) también estudiaron las propiedades de la distancia L_∞ entre C y Π dada por

$$\delta_{X,Y} = \delta_C := 4\sup_{(u,v)\,\in\,\mathbb{I}^2} |\,C(u,v) - uv\,|\,, \tag{21}$$

y demostraron que esta medida satisface todas excepto la propiedad 5 en la Definición 3.2. Al respecto, y en un contexto mucho más general (dimensión igual o mayor a 2 y con variables aleatorias no necesariamente continuas) Fernández y González-Barrios (2004) proponen una medida de dependencia multidimensional que entre otras cosas prescinde de la propiedad 5

$$\delta_{X_1,\ldots,X_n} := \sup_{(x_1,\ldots,x_n)\,\in\,\mathbb{R}^n} \left| F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) - \prod_{i\,=\,1}^{n} F_{X_i}(x_i) \right|, \tag{22}$$

en donde $F_{X_1,\ldots,X_n}$ es la función de distribución conjunta de las variables aleatorias $X_1,\ldots,X_n$ y F_{X_i} es la función de distribución marginal de X_i, y demostraron el siguiente:

3.4. Teorema. ***(Fernández y González-Barrios, 2004)*** *Para cualesquiera variables aleatorias $X_1,\ldots,X_n$:*

a) $\delta_{X_1,\ldots,X_n} = \delta_{X_{\omega(1)},\ldots,X_{\omega(n)}}$ *para cada permutación* ω *de* $\{1,2,\ldots,n\}$.

b) $\delta_{X_1,\dots,X_n} = 0$ *si y sólo si* $X_1,\dots,X_n$ *son independientes.*

c) *Para cada* $x_i \in \mathbb{R}$ *e* $i \in \{1,2,\dots,n\}$

$$-\left(\frac{n-1}{n}\right)^n \leq F_{X_1,\dots,X_n}(x_1,\dots,x_n) - \prod_{i=1}^{n} F_{X_i}(x_i) \leq \left(\frac{1}{n}\right)^{\frac{1}{n-1}}\left(1-\frac{1}{n}\right) < 1\,.$$

Y en consecuencia $0 \leq \delta_{X_1,\dots,X_n} \leq 1$. *Más aún, las cotas son alcanzables.*

d) $0 \leq \delta_{X_1,X_2} \leq \delta_{X_1,X_2,X_3} \leq \cdots \leq \delta_{X_1,\dots,X_{n-1}} \leq \delta_{X_1,\dots,X_n}$.

En el caso particular de variables aleatorias continuas, el teorema anterior y el Teorema de Sklar nos conducen a

$$\delta_C := \sup_{(u_1,\dots,u_n)\in \mathbb{I}^n} \left| C(u_1,\dots,u_n) - \prod_{i=1}^{n} u_i \right|, \tag{23}$$

en donde C es la cópula subyacente de las variables aleatorias continuas $X_1,\dots,X_n$. También Wolff (1980) trabajó con versiones multivariadas de (20).

4. En resumen

Las cópulas y el Teorema de Sklar han ayudado al estudio de la dependencia de variables aleatorias, particularmente las continuas, ampliando considerablemente el catálogo de distribuciones multivariadas. Esto ha tenido una consecuencia importante en la Estadística ya que permite trabajar, por un lado, la estimación de la cópula subyacente, y por otro, la estimación de las distribuciones marginales, para finalmente obtener una estimación de la distribución conjunta por medio del Teorema de Sklar. También resultan ser una herramienta útil en la discusión sobre las propiedades deseables de una medida de dependencia ya que, en el caso de variables aleatorias continuas, determinan de manera única la distribución conjunta, y por ello, contienen la información relevante sobre la estructura de dependencia.

De acuerdo a las consideraciones y propuestas de Drouet y Kotz (2001), Embrechts *et al.* (1999, 2003), Rényi (1959), Schweizer y Wolff (1981), y Fernández y González-Barrios (2004), al menos las siguientes propiedades podrían considerarse como deseables para una medida de dependencia μ para variables aleatorias continuas $X_1,\dots,X_n$:

1. μ debe estar definida para cualesquiera $X_1, \ldots, X_n$ variables aleatorias continuas, en términos de la cópula subyacente únicamente;
2. $\mu(X_1, \ldots, X_n) = \mu(\wp(X_1, \ldots, X_n))$ para cualquier permutación $\wp$ de $(X_1, \ldots, X_n)$;
3. $0 \leq \mu(X_1, \ldots, X_n) \leq k_n$, para alguna $k_n < \infty$;
4. $\mu(X_1, \ldots, X_n) = 0$ si y sólo si $X_1, \ldots, X_n$ son independientes.
5. $0 \leq \mu(X_1, X_2) \leq \mu(X_1, X_2, X_3) \leq \cdots \leq \mu(X_1, \ldots, X_n)$.

En particular, la posibilidad de considerar a la propiedad 5 como deseable, es consecuencia del trabajo desarrollado en Fernández y González-Barrios (2004). Intuitivamente, por ejemplo, si ya se tiene cierto grado de dependencia $\mu(X_1, X_2)$ entre las variables aleatorias (X_1, X_2), el considerar una tercera variable aleatoria X_3, si acaso incrementará el nivel de dependencia, ahora entre las tres variables aleatorias (X_1, X_2, X_3), pero en un caso extremo, si X_3 es independiente de (X_1, X_2) y por tanto no tiene nada que aportar a la dependencia ya existente entre (X_1, X_2), al menos se tendría un grado de dependencia igual al que ya se tenía al considerar únicamente a (X_1, X_2), esto es $\mu(X_1, X_2) \leq \mu(X_1, X_2, X_3)$.

Por último, el lector interesado en aplicaciones particulares de cópulas, podría considerar las siguientes. A manera de una guía rápida general para ajuste paramétrico de cópulas están los trabajos de Trivedi y Zimmer (2005) y De Matteis (2001), y algunos capítulos en Kurowicka y Cooke (2006). Para aplicaciones en finanzas: Cherubini *et al.* (2004), Rank (2007), Xu (2005), así como algunos capítulos en Franke *et al.* (2008), Malvergne y Sornette (2006), Cont y Tankov (2004). Para aplicaciones en teoría del riesgo y ciencias actuariales: Denuit *et al.* (2005), y algunos capítulos en Shang (2006), Panjer (2006), McNeil *et al.* (2005). Para aplicaciones en geofísica: Salvadori *et al.* (2007). Y para aplicaciones en lógica difusa y normas triangulares: Alsina *et al.* (2006), Klement y Mesiar (2005), Klement *et al.* (2000), Schweizer y Sklar (2005).

Referencias

Alsina, C., Frank, M.J., Schweizer, B. (2006). *Associative Functions: Triangular Norms and Copulas*, World Scientific Publishing Co. (Singapore).

Blomqvist, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Statist.* **21**, 593-600.

Carley, H., Taylor, M.D. (2002). A new proof of Sklar's theorem. En: *Distributions with Given Marginals and Statistical Modelling; Carles M. Cuadras, Josep Fortiana y José A. Rodríguez-Lallena editores (Kluwer)* , 29-34.

Cherubini, H., Luciano, E. and Vecchiato, W. (2004). *Copula methods in finance*, Wiley (Cornwall, UK).

Cont, R., Tankov, P. (2004). *Financial Modelling With Jump Processes*, Chapman & Hall/CRC (New York).

Dall'Aglio, C. (1959). Sulla compatibilità delle funzioni di ripartizione doppia. *Rend. Mat.* **18**, (5), 385-413.

Dall'Aglio, C. (1960). Les fonctions extrêmes de la clase de Fréchet à 3 dimensions. *Publ. Inst. Statist. Paris* **9**, 175-188.

Dall'Aglio, C. (1972). Fréchet classes and compatibility of distribution functions. *Symposia Math.* **9**, 131-150.

De Matteis, R. (2001). *Fitting copulas to data*, Diploma Thesis ETH Zürich.

Denuit, M., Dhaene, J., Goovaerts, M., Kaas, R. (2005). *Actuarial Theory for Dependent Risks*, Wiley (Sussex, UK).

Drouet Mari, D., Kotz, S. (2001). *Correlation and dependence*, Imperial College Press (London).

Embrechts, P., McNeil, A.J., Straumann, D. (1999). Correlation: pitfalls and alternatives. *Risk Magazine* **5**, 69-71.

Embrechts, P., Lindskog, F., McNeil, A.J. (2003). Modeling dependence with copulas and applications to risk management. En: *Handbook of Heavy-Tailed Distributions in Finance, Rachev, S. (editor) Elsevier, capítulo 8,* 329-384.

Farlie, D.J.G. (1960). The performance of some correlation coefficients for a general bivariate distribution. *Biometrika* **47**, 307-323.

Fernández, B., González-Barrios, J.M. (2004). Multidimensional dependency measures. *J. Multivariate Anal.* **89**, 351-370.

Féron, R. (1956). Sur les tableaux de corrélation dont les marges sont données. *Publ. Inst. Statist. Univ. Paris* **5**, 3-12.

Franke, J., Härdle, W.K., Hafner, C.M. (2008). *Statistics of Financial Markets*, Springer-Verlag (Berlin).

Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon* **14**, (Sect. A Ser. 3), 53-77.

Fréchet, M. (1957). Les tableaux de corrélation et les programmes linéaires. *Revue Inst. Int. Statist.* **25**, 23-40.

Gini, C. (1914). Di una misura delle relazioni tra le graduatorie di due caratteri. Apéndice de: *Mancini, A., "Le Elezioni Generali Politiche del 1913 nel Comune di Roma," Ludovico Cecchini (Rome).*

Gumbel, E.J. (1958). Distributions à plusieures variables dont les marges sont donnés. *Comptes Rendus de l'Academie des Sciences Paris* **246**, 2717.

Gumbel, E.J. (1960). Bivariate exponential distributions. *J. Amer. Statist. Assoc.* **55**, 698-707.

Hoeffding, W. (1940). Masstabinvariante Korrelationstheorie. *Schriften des Matematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin* **5**, 179-223.

Joe, H. (1996). Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. En: *Distributions with Fixed Marginals and Related Topics; Rüschendorf, Schweizer, Taylor, editores* **28**, 120-141.

Joe, H. (1997). *Multivariate models and dependence concepts*, Ed. Chapman & Hall (New York).

Kendall, M.G. (1938). A new measure of rank correlation. *Biometrika* **30**, 81-93.

Klement, E.P., Mesiar, R., Pap, E. (2000). *Triangular Norms*, Kluwer (Dordrecht).

Klement, E.P., Mesiar, R., editores (2005). *Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms*, Elsevier (Amsterdam).

Kotz, S., Seeger, J.P. (1991). A new approach to dependence in multivariate distributions. En: *Advances in Probability Distributions with Given Marginals; Dall'Aglio, Kotz, Salinetti editores (Kluwer Academic Publishers, Dordrecht),* 113-127.

Kurowicka, D., Cooke, R. (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*, Wiley (Sussex, UK).

Lehmann, E.L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37**, 1137-1153.

Malvergne, Y., Sornette, D. (2006). *Extreme Financial Risks*, Springer-Verlag (Berlin).

Marshall, A.W., Olkin, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **62**, 30-44.

Marshall, A.W. (1996). Copulas, Marginals, and Joint Distributions. En: *Distributions with Fixed Marginals and Related Topics; Rüschendorf, Schweizer, Taylor, editores* **28**, 213-222.

McNeil, A.J., Frey, R., Embrechts, P. (2005). *Quantitative Risk Management*, Princeton Series in Finance (New Jersey).

Mikusiński, P., Sherwood, H., Taylor, M.D. (1991). Probabilistic interpretations of copulas and their convex sums. En: *Advances in Probability Distributions with Given Marginals; Dall'Aglio, Kotz, Salinetti editores (Kluwer Academic Publishers, Dordrecht),* 95-112.

Morgenstern, D. (1956). Einfache Beispiele Zweidimensionaler Verteilungen. *Mitteilingsblatt für Mathematische Statistik* **8**, 234-235.

Nelsen, R.B. (1995). Copulas, Characterization, Correlation, and Counterexamples. *Mathematics Magazine* **68**, (3), 193-198.

Nelsen, R.B. (2006). *An introduction to copulas, 2nd edition*, Springer-Verlag (New York). La primera edición es de 1999.

Panjer, H. (2006). *Operational Risk. Modeling Analytics*, Wiley (New Jersey).

Quesada-Molina, J.J., Rodríguez-Lallena, J.A. (1994). Some advances in the study of the compatibility of three bivariate copulas. *J. Ital. Statist. Soc.* **3**, 397-417.

Rank, J. (2007). *Copulas. From theory to application in finance*, Riskbooks (London).

Rényi, A. (1959). On measures of dependence. *Acta. Math. Acad. Sci. Hungar.* **10**, 441-451.

Salvadori, G., De Michele, C., Kottegoda, N.T., Rosso, R. (2007). *Extremes in nature. An approach using copulas*, Springer (Dordrecht).

Schweizer, B., Wolff, E.F. (1981). On nonparametric measures of dependence for random variables. *Ann. Statist.* **9**, (4), 870-885.

Schweizer, B. (1991). Thirty years of Copulas. En: *Advances in Probability Distributions with Given Marginals; Dall'Aglio, Kotz, Salinetti editores (Kluwer Academic Publishers, Dordrecht),* 13-50.

Schweizer, B., Sklar, A. (2005). *Probabilistic Metric Spaces*, Dover (New York). Republicación de la edición de 1983, con referencias y notas históricas adicionales.

Shang, H. (2006). *Actuarial Science. Theory and Methodology*, World Scientific (Singapore).

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Inst. Statist. Univ. Paris Publ.* **8**, 229-231.

Sklar, A. (1996). Random Variables, Distribution Functions, and Copulas – A Personal Look Backward and Forward. En: *Distributions with Fixed Marginals and Related Topics; Rüschendorf, Schweizer, Taylor editores* **28**, 1-14.

Spearman, C. (1904). The proof and measurement of association between two things. *Amer. J. Psychol.* **15**, 72-101.

Trivedi, P.K., Zimmer, D.M. (2005). Copula Modeling: An Introduction for Practitioners. *Foundations and Trends in Econometrics* **1**, (1), 1-111.

Wolff, E.F. (1980). N-dimensional measures of dependence. *Stochastica* **4**, (3), 175-188.

Xu, Y. (2005). *Applications of copula-based models in portfolio optimization*, Tesis doctoral, University of Miami.