



**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO**

FACULTAD DE CIENCIAS

**"SOBRE COPULAS Y DEPENDENCIA DE
VARIABLES ALEATORIAS"**

T E S I S

QUE PARA OBTENER EL TITULO DE:

A C T U A R I O

P R E S E N T A :

NELSON OMAR MURIEL TORRERO



FACULTAD DE CIENCIAS
UNAM

DIVISION DE ESTUDIOS PROFESIONALES
DIRECTOR DE TESIS:
DRA. M. ASUNCION BEGORA FERNANDEZ FERNANDEZ



2004
FACULTAD DE CIENCIAS
SECCION ESCOLAR



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

ACT. MAURICIO AGUILAR GONZÁLEZ
Jefe de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito: Sobre cúpulas y dependencia de variables aleatorias.

realizado por Nelson Omar Muriel Torrero

con número de cuenta 9852000-8 , quien cubrió los créditos de la carrera de:

Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

A t e n t a m e n t e

Director de Tesis

Propietario Dra. María Asunción Begoña Fernández Fernández *Begoña Fernández*

Propietario Dra. Ana Meda Guardiola *Ana Meda*

Propietario Dr. José María González Barrios Murguía *José M. González Barrios*

Suplente Dr. Alberto Contreras Cristán *Alberto Contreras*

Suplente M. en C. Arturo Erdely Ruíz *Arturo Erdely Ruíz*

**Consejo Departamental de
Matemáticas**

José Antonio Flores Díaz
 M. en C. José Antonio Flores Díaz

A mi mamá

Que gusto poder agradecer a quienes siento me han brindado ayuda y apoyo en el proceso de esta tesis. Quiero comenzar, como es de suponerse, con mi familia. A mi madre, por el apoyo que siempre me ha dado, siempre suficiente o un poco más. A mi abuela, indispensable soporte. A mi hermano, fuente de fuerzas e inspiraciones. A mi mujer, paciente y tan dispuesta a ayudarme. A mi hijo, que sin saberlo o quererlo también me impulsa.

Habiendo dicho esto, quiero agradecer a Begoña por su muy grande apoyo en la realización de este trabajo. Por su tiempo y sus comentarios, por su compromiso, su disciplina y por su cariño. A tí, hermano Gerardo, por impulsarme, por compartirme tus ideas, tu trabajo, tus sonrisas. Gracias. A mi muy querida Coco, sin quien este trabajo hubiese demorado muchísimo tiempo más entre ventanilla y ventanilla. A Ana, por su muy amistoso apoyo.

A Beto, José María y Arturo, los otros tres sinodales de esta tesis por sus observaciones que mejoraron mucho el contenido y presentación del presente trabajo. A Guillermo, Gerónimo y María Emilia, tres personas que contribuyeron mucho en mi formación académica. A mis compañeros y alumnos que colaboraron en mis estudios.

A todos ustedes, muchas gracias.

Contenido

Introducción	1
1 Cópulas	5
1.1 Introducción	5
1.2 La función de distribución conjunta y sus marginales.	6
1.3 Cópulas	10
1.4 Teorema de Sklar.	18
1.5 Cópulas y dependencia bivariada	23
1.6 Reordenamientos de M	32
1.7 Modos de dependencia	34
1.8 Cópulas n -dimensionales	47
1.9 Conclusiones	49
2 Cópulas Arquimedianas	51
2.1 Introducción	51
2.2 Construcción de Cópulas Arquimedianas	52
2.3 Propiedades	60
2.4 Conclusiones	71
3 Procesos de Markov	73
3.1 Introducción	73
3.2 Un producto para cópulas bidimensionales	74
3.3 El producto $*$ y las ecuaciones de Chapman-Kolmogorov	78
3.4 Ejemplos	81
3.5 El producto $*$ y la propiedad de Markov	86
3.6 Estructura de $(\mathcal{C}, *)$	90
3.7 Conclusiones	94

4 Cópulas Empíricas	95
4.1 Introducción	95
4.2 Simulación	96
4.3 Cópulas Empíricas	103
4.4 Cópulas Arquimedianas	113
4.5 Conclusiones	125
Conclusiones Generales	128

Introducción

En los cursos elementales de probabilidad se estudia la función de distribución correspondiente a variables y vectores aleatorios. Un resultado importante que involucra estas funciones es el siguiente:

Teorema 1. *Dado un vector aleatorio (X, Y) con función de distribución conjunta F y marginales F_X y F_Y respectivamente*

$$F(x, y) = F_X(x)F_Y(y) \quad \forall(x, y)$$

si y sólo si las variables X y Y son independientes.

Esto nos lleva a preguntarnos ¿qué relación existe entre la función de distribución conjunta de un vector aleatorio y sus funciones de distribución marginales?

Para responder a esta interrogante es útil el concepto de cópula.

Definición. *Una función $C^*: [0, 1] \times [0, 1] \rightarrow [0, 1]$ es una cópula si*

1. *Para cada $(u, v) \in [0, 1] \times [0, 1]$ si $\min(u, v) = 0$, entonces $C^*(u, v) = 0$.*
2. *Para todo par de puntos $(u_1, v_1), (u_2, v_2) \in [0, 1] \times [0, 1]$*

$$C^*(u_1, v_1) + C^*(u_2, v_2) - C^*(u_1, v_2) - C^*(u_2, v_1) \geq 0$$

3. *Para todo $u \in [0, 1] \times [0, 1]$*

$$C^*(1, v) = v \quad C^*(u, 1) = u$$

Alternativamente, podemos definir una cópula como sigue

Definición. *Una función $C^*: [0, 1] \times [0, 1] \rightarrow [0, 1]$ es una cópula si es una función de distribución correspondiente a un vector aleatorio (U, V) donde $U \sim U(0, 1)$ y $V \sim U(0, 1)$.*

El principal resultado referente a la relación entre la función de distribución conjunta y sus distribuciones marginales es el Teorema de Sklar. En él se afirma que toda función de distribución multivariada es la composición de una cópula con las funciones marginales de distribución. En su versión de dos variables el teorema es el siguiente.

Teorema (Teorema de Sklar, (1959)). *Sea $F_{X,Y}$ una función de distribución bivariada con marginales F_X y F_Y . Entonces existe una cópula tal que*

$$F(x, y) = C(F_X(x), F_Y(y)) \quad \forall x, y \in \mathfrak{R} \quad (1)$$

Además, si F_X y F_Y son continuas, entonces C es única. De otro modo, queda únicamente determinada sobre $\text{Ran}F_X \times \text{Ran}F_Y$.

Por ejemplo, para el caso en que el vector (X, Y) es de variables independientes, la cópula para la distribución conjunta es $\Pi(u, v) = uv$ como lo muestra el Teorema 1.1. La factorización de la distribución conjunta que demuestra el Teorema de Sklar sugiere que las cópulas guardan cierta relación con la dependencia de las entradas de un vector aleatorio. De hecho nos ayuda a caracterizar los aspectos de la dependencia que no se relacionen con las distribuciones marginales.

Por ejemplo, la relación $Y \stackrel{c.s.}{=} f(X)$ queda resumida en la cópula de la distribución del vector (X, Y) cuando f es monótona. Hay otros resultados que revelan la relación de la cópula de un vector aleatorio y ciertos aspectos de la dependencia entre sus entradas. Por ejemplo, dos medidas de concordancia muy usuales, la τ de Kendall y la ρ de Spearman son función exclusivamente de la cópula. No obstante, no todos los coeficientes de dependencia tienen esta propiedad. Tal es el caso del coeficiente de correlación que depende de las distribuciones marginales.

Existen varios métodos para construir cópulas. Podemos usar el Teorema de Sklar y dar una cópula a partir de una función de distribución multivariada. Otra alternativa es el siguiente teorema.

Teorema. *Sea $\phi: [0, 1] \rightarrow [0, \infty]$ una función continua, estrictamente decreciente tal que $\phi(1) = 0$. Entonces la función $C: [0, 1] \times [0, 1] \rightarrow [0, 1]$ definida por*

$$C(u, v) = \phi^{[-1]} \left(\phi(u) + \phi(v) \right)$$

es una cópula si y sólo si ϕ es convexa.

Las cópulas construidas mediante este procedimiento se llaman cópulas Arquimedianas y ϕ se llama generador de la cópula C . Una propiedad evidente de esta familia de cópulas es que son simétricas, es decir $C(u, v) = C(v, u)$ para todo $(u, v) \in [0, 1] \times [0, 1]$. Otra propiedad menos evidente es que son asociativas, es decir

$$C(u, C(v, w)) = C(C(u, v), w) \quad \forall u, v, w \in [0, 1]$$

El Teorema de Ling nos asegura que estas dos condiciones son suficientes para que una cópula sea Arquimediana.

Hay otras propiedades que caracterizan a una cópula Arquimediana. Una muy importante es la siguiente

Proposición. *Dado un vector aleatorio (U, V) con distribución C , donde C es una cópula Arquimediana con generador ϕ , la función*

$$K_C(t) = \mathbb{P}(C(U, V) \leq t)$$

caracteriza a ϕ mediante la ecuación

$$\phi(t) = \exp \left(\int_{t_0}^t \frac{1}{x - K_C(x)} dx \right)$$

siempre que $K(t^-) > t$ para todo $t \in (0, 1)$.

Esta propiedad resulta de gran ayuda para utilizar las cópulas Arquimedianas en problemas estadísticos.

Otro campo de estudio en que se pueden utilizar cópulas es el de los Procesos de Markov. El enfoque tradicional para la construcción de un proceso de Markov $\{X_t\}_{t \in T}$ es brindar una distribución inicial F_0 y una familia $\{P_t\}_{t \in T}$ de probabilidades de transición que satisfaga las ecuaciones de Chapman-Kolmogorov. Usando la cópulas de las variables del proceso podemos caracterizar estas ecuaciones. Gracias a ello, se plantea un nuevo modo para construir procesos de Markov que consiste en dar una familia de cópulas bivariadas $\{C_{s,t}\}_{s,t \in T}$ que satisfagan las ecuaciones de Champan-Kolmogorov y una familia $\{F_t\}_{t \in T}$ de distribuciones para las variables del proceso.

Con este método se pueden dar procesos que compartan la familia de cópulas $\{C_{s,t}\}_{s,t \in T}$ (y así algunos aspectos de la dependencia entre las variables del proceso) con el movimiento Browniano sin tener marginales Normales y, de hecho, eligiendo para cada $t \in T$ la distribución F_t que deseemos.

Se puede también demostrar que las ecuaciones de Chapman–Kolmogorov no caracterizan la propiedad de Markov (hecho bien conocido) aunque las cópulas del proceso sí.

También en problemas de datos reales podemos emplear cópulas. Tenemos un Teorema al estilo Glivenko–Cantelli en que se prueba que una cópula obtenida a partir de una muestra, y que llamamos cópula empírica, converge uniformemente, c.s., a la cópula poblacional. Este resultado es útil en el estudio de la independencia de variables aleatorias ya que a partir de la convergencia mencionada se pueden establecer criterios de aceptación o rechazo para la hipótesis de independencia.

El objetivo de esta tesis es estudiar las propiedades y usos de las cópulas en la teoría de probabilidad.

En el capítulo primero se tratan las cópulas en general, abordando sus propiedades más elementales para luego relacionarlas con varias modalidades de dependencia. Se examina fundamentalmente el caso bivariado y se extienden los resultados al caso de n variables.

El objeto de estudio del segundo capítulo es la familia de cópulas Arquimedianas. Estudiamos su construcción y propiedades y brindamos varios ejemplos de cópulas de esta familia. Obtenemos relaciones entre la familia Arquimediana de cópulas y la τ de Kendall, que en el capítulo cuarto sirven de apoyo para la construcción de métodos de ajuste en problemas de datos reales.

En el tercer capítulo examinamos los procesos de Markov y el modo de caracterizar tanto las ecuaciones de Chapman–Kolmogorov como la propiedad de Markov mediante las cópulas del proceso. Elaboramos algunas construcciones ilustrativas.

El cuarto y último capítulo nos sitúa en el campo de la estadística y la simulación. Demostramos algoritmos para la simulación de observaciones de diversas cópulas y presentamos un resumen gráfico de algunas simulaciones. Utilizamos después las cópulas empíricas para examinar la independencia de los precios del tipo de cambio dolar–peso, fix, misma que rechazamos a una distancia de 7 días. Por último ajustamos a estos precios una cópula Arquimediana. Damos el resumen de los resultados y los gráficos explicativos.

Capítulo 1

Cóputas

1.1 Introducción

El presente capítulo aborda la definición y algunas propiedades de las cóputas relacionándolas con funciones de distribución bivariadas. El tema central de este capítulo es descubrir si hay alguna relación entre las cóputas y la dependencia de vectores aleatorios.

En la sección 1.2 se observa que las funciones de distribución marginales no son suficientes para determinar de modo único la distribución conjunta. Sin embargo, se demuestra que hay casos en que usando una relación funcional entre las distribuciones marginales podemos recobrar la distribución conjunta de modo único.

Esto da la pauta para que en la sección 1.3 se defina una cóputa y se examinen sus propiedades. Se demuestra en la sección 1.4 que toda función de distribución bivariada es la composición de una cóputa con sus marginales. En razón de ello se propone que la cóputa de un vector aleatorio encierra información de la dependencia entre sus coordenadas.

En la sección 1.5 se demuestra que la independencia entre las entradas de un vector (X, Y) es un atributo de su cóputa. Se encuentran cotas del conjunto de las cóputas y se examina qué información encierran dichas cotas en términos de la dependencia de (X, Y) .

En la sección 1.6 se trabaja con una clase de cóputas que tiene una propiedad muy especial en términos de la dependencia entre variables, el ser densa con respecto a la norma uniforme en el conjunto de las cóputas de dos dimensiones.

En la sección 1.7 examinamos si existen algunos modos de dependencia que se resuman en la cópula y no dependan de las distribuciones marginales.

En la sección 1.8 se extienden la definición y propiedades de cópula al caso n -dimensional. Se presentan después las conclusiones del trabajo realizado.

Para este capítulo utilizamos los capítulos 1, 2, 3 y 5 de [1].

1.2 La función de distribución conjunta y sus marginales.

Uno de los resultados en que se hace hincapié en los cursos elementales de probabilidad es el siguiente:

Teorema 1.1. *Dado un vector aleatorio (X, Y) con función de distribución conjunta F y marginales F_X y F_Y respectivamente*

$$F(x, y) = F_X(x)F_Y(y) \quad \forall(x, y)$$

si y sólo si las variables X y Y son independientes.

Existen una infinidad de ejemplos de vectores aleatorios con funciones de distribución marginales iguales y funciones de distribución conjunta distintas. Uno de los más populares en la literatura es el siguiente:

Ejemplo 1.1. *Consideremos f_1 y f_2 funciones de densidad con funciones de distribución respectivas F_1 y F_2 . Para cada $\alpha \in [-1, 1]$, la función*

$$f_\alpha(x, y) = f_1(x)f_2(y) \left[1 + \alpha \{2F_1(x) - 1\} \{2F_2(y) - 1\} \right]$$

es una densidad bivariada con densidades marginales f_1 y f_2 .

Demostración. Demostraremos primero que f_α es una densidad.

Observemos que

$$-1 \leq 2F_1(x) - 1 \leq 1 \quad \text{y} \quad -1 \leq 2F_2(y) - 1 \leq 1$$

por lo que, como $-1 \leq \alpha \leq 1$,

$$-1 \leq \alpha \{2F_1(x) - 1\} \{2F_2(y) - 1\} \leq 1.$$

1.2. LA FUNCIÓN DE DISTRIBUCIÓN CONJUNTA Y SUS MARGINALES.7

De donde $f_\alpha(x, y) \geq 0$ para todo punto $(x, y) \in \mathfrak{R}^2$.

Ahora bien, probemos que integra la unidad.

$$\iint_{\mathfrak{R}^2} f_\alpha(x, y) dx dy = \iint_{\mathfrak{R}^2} f_1(x) f_2(y) \left[1 + \alpha \{2F_1(x) - 1\} \{2F_2(y) - 1\} \right] dx dy$$

El cálculo de esta integral es equivalente al cálculo de las integrales

$$\iint_{\mathfrak{R}^2} f_1(x) f_2(y) dx dy \quad \text{e} \quad \iint_{\mathfrak{R}^2} \alpha f_1(x) f_2(y) \{2F_1(x) - 1\} \{2F_2(y) - 1\} dx dy$$

Por ser f_1 y f_2 funciones de densidad el resultado de la primera integral es igual a 1. Resta probar que la segunda integral es nula.

Para ello, integramos con respecto a x en la expresión

$$\alpha \int_{\mathfrak{R}} f_2(y) \{2F_2(y) - 1\} dy \int_{\mathfrak{R}} f_1(x) \{2F_1(x) - 1\} dx \quad (1.1)$$

Para ello hacemos el cambio de variable $u = 2F_1(x) - 1$, $du = 2f_1(x)$ y obtenemos

$$\int_{\mathfrak{R}} f_1(x) \{2F_1(x) - 1\} dx = \frac{1}{2} \int_{-1}^1 u du = 0$$

Hemos probado que

$$\iint_{\mathfrak{R}^2} f_\alpha(x, y) dx dy = 1 \quad \text{para todo } \alpha \in [-1, 1]$$

Queda por demostrar que las densidades marginales son f_1 y f_2 . Para ello integremos $f_\alpha(x, y)$ sobre \mathfrak{R} con respecto a cada variable.

$$\int_{\mathfrak{R}} f_\alpha(x, y) dx = f_2(y) \int_{\mathfrak{R}} f_1(x) dx + \alpha f_2(y) \{2F_2(y) - 1\} \int_{\mathfrak{R}} f_2(y) \{2F_1(x) - 1\} dx$$

Dado que f_1 es densidad la primera integral es igual a 1. La integral en el segundo término la hemos calculado anteriormente y es igual a 0. Por lo tanto

$$\int f_\alpha(x, y) dx = f_2(y) \quad \forall \alpha \in [-1, 1]$$

La prueba es análoga para la otra marginal. ■

A partir del Teorema 1.1 y del ejemplo 1.1 se plantea que, en general, la función de distribución conjunta no está determinada por las funciones de distribución marginales. Esta afirmación puede dar la impresión errónea de que las distribuciones marginales no tienen importancia alguna.

Para aclarar este punto, analicemos con un poco más de cuidado este problema. Consideremos la función $C : \mathfrak{R}^2 \rightarrow \mathfrak{R}$ dada por

$$C(u, v) = \begin{cases} 0, & \text{si } u \leq 0 \text{ ó } v \leq 0, \\ uv, & \text{si } (u, v) \in (0, 1)^2, \\ u, & \text{si } u \in (0, 1), v \geq 1, \\ v, & \text{si } u \geq 1, v \in (0, 1), \\ 1, & \text{en otro caso.} \end{cases} \quad (1.2)$$

Esta función es la función de distribución conjunta de U, V , donde U, V son variables aleatorias independientes uniformes en $(0, 1)$. Sean X, Y variables aleatorias independientes con funciones de distribución F_X y F_Y respectivamente y función de distribución conjunta F . Claramente

$$F(x, y) = C(F_X(x), F_Y(y)) = F_X(x)F_Y(y) \quad \forall (x, y) \in \mathfrak{R}^2. \quad (1.3)$$

La expresión (1.3) nos dice que podemos expresar la función de distribución conjunta de (X, Y) como la composición de la función C dada por (1.2) y la función $h : \mathfrak{R}^2 \rightarrow \mathfrak{R}^2$, $h(x, y) = (F_X(x), F_Y(y))$.

Podemos ahora plantear el siguiente problema: Dado un vector aleatorio (X, Y) , ¿existe una función de distribución C con marginales uniformes tal que

$$F(x, y) = C(F_X(x), F_Y(y)) \quad (1.4)$$

para todo punto $(x, y) \in \mathfrak{R}^2$? De existir esta función C podemos decir que las distribuciones marginales determinan la distribución conjunta a través de una relación funcional. Observemos que en la expresión (1.3) sólo tiene importancia la definición de la función C en el cuadrado unitario. Consideremos el siguiente ejemplo

Ejemplo 1.2. Consideremos la distribución exponencial bivariada de Gumbel de parámetro $\theta \geq 0$

$$F_\theta(x, y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-(x+y+\theta xy)} & \text{si } x \geq 0 \text{ } y \geq 0 \\ 0 & \text{e.o.c.} \end{cases}$$

Mostremos que es función de sus marginales.

1.2. LA FUNCIÓN DE DISTRIBUCIÓN CONJUNTA Y SUS MARGINALES.9

Demostración. Denotemos a sus marginales $F(x)$ y $G(y)$ respectivamente. Tomando límites se observa que ambas son exponenciales estándar. Lo que buscamos ahora es una función $C: [0, 1]^2 \rightarrow [0, 1]$ tal que $F_\theta(x, y) = C(F(x), G(y))$ en todo par de puntos x, y . De manera equivalente, gracias a que las funciones marginales son continuas e invertibles, necesitamos que $C(u, v) = F_\theta(F^{-1}(u), G^{-1}(v))$ para cualquier pareja u, v en el intervalo unitario¹.

Conocemos las inversas $F^{-1}(u) = -\ln(1-u)$ y $G^{-1}(v) = -\ln(1-v)$ de modo que, substituyendo, tenemos

$$C(u, v) = u + v - 1 + (1-u)(1-v)e^{-\theta \ln(1-u) \ln(1-v)}$$

Es decir

$$F_\theta(x, y) = F(x) + G(y) - 1 + (1-F(x))(1-G(y))e^{-\theta \ln(1-F(x)) \ln(1-G(y))}$$

■

Pensando en lo hecho en este desarrollo, queda claro que si tenemos una distribución bivariada de marginales continuas e invertibles (estrictamente crecientes) entonces podremos recobrar la función bivariada con las marginales mediante la expresión

$$F(x, y) = C(F_X(x), G_Y(y)) \quad \text{tomando} \quad C(u, v) = F(F_X^{-1}(u), G_Y^{-1}(v))$$

Entre otras, la función C que obtuvimos en el ejemplo 1.2 tiene las siguientes propiedades.

Su dominio, $\text{dom}C$, es $\text{Ran}F_X^{-1} \times \text{Ran}G_Y^{-1}$, un subconjunto del cuadrado unitario, de hecho $\{0, 1\}^2 \subseteq \text{dom}C \subseteq [0, 1]^2$.

En su frontera se tiene que

$$F(-\infty, y) = C(0, G_Y(y)) \Rightarrow \forall v \in \text{Ran}G_Y^{-1} \quad C(0, v) = 0 \quad (1.5)$$

De la misma manera se deberá satisfacer que si $u \in \text{Ran}G_X^{-1}$ entonces $C(u, 0) = 0$.

Por otra parte

$$F(\infty, G_Y(y)) = C(1, G_Y(y)) \Rightarrow \forall v \in \text{Ran}G_Y^{-1} \quad C(1, v) = v \quad (1.6)$$

¹Haciendo $u = F(x), v = G(y)$ cubrimos todo el intervalo unitario por continuidad de F y G

Por simetría del argumento, tendremos que sobre $RanF_X^{-1} C(u, 1) = u$. Esto en lo que respecta a la frontera de $domC$. Por otro lado, recordemos que C está definida a partir de una distribución, que es una función bicreciente, es decir, si $x_1 \leq x_2$ y $y_1 \leq y_2$ entonces

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

Transcribiendo esto en términos de C , tendremos que

$$\begin{aligned} C(F_X^{-1}(x_1), G_Y^{-1}(y_1)) + C(F_X^{-1}(x_2), G_Y^{-1}(y_2)) \\ - C(F_X^{-1}(x_1), G_Y^{-1}(y_2)) - C(F_X^{-1}(x_2), G_Y^{-1}(y_1)) \geq 0 \end{aligned}$$

cuando se satisfaga que $x_1 \leq x_2$ y que $y_1 \leq y_2$. Usando el hecho de que las inversas de funciones crecientes son también crecientes, podemos decir que si $u_1 \leq u_2$ y $v_1 \leq v_2$ para puntos en $domC$ entonces

$$C(u_1, v_1) + C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) \geq 0 \quad (1.7)$$

1.3 Cópulas

Hasta ahora tenemos una respuesta parcial al problema de hallar una relación funcional entre las funciones de distribución marginales y la función de distribución conjunta. Busquemos generalizar esta respuesta. Con base en la sección anterior toda función $C : \mathfrak{R}^2 \rightarrow [0, 1]$ que sea candidata para establecer una relación del tipo (1.4) deberá tener ciertas restricciones. A partir de ellas definiremos los conceptos de *subcópula* y *cópula*.

Definición 1. Una función $C^* : S_1 \times S_2 \rightarrow [0, 1]$ se llama *subcópula* si

1. $S_1 \cup S_2 \subseteq [0, 1]$ y $\{0, 1\} \subseteq S_1 \cap S_2$
2. Para cada $(u, v) \in S_1 \times S_2$ si $\min(u, v) = 0$, entonces $C^*(u, v) = 0$.
Llamaremos a C^* *anclada por* satisfice esta propiedad
3. Para todo par de puntos $(u_1, v_1), (u_2, v_2) \in S_1 \times S_2$

$$C^*(u_1, v_1) + C^*(u_2, v_2) - C^*(u_1, v_2) - C^*(u_2, v_1) \geq 0$$

Llamaremos a C^* *bicreciente por* satisfacer esta propiedad.

4. Para todo $u \in S_1, v \in S_2$

$$C^*(1, v) = v \quad C^*(u, 1) = u$$

Definición 2. Una cópula es una subcópula cuyo dominio es $[0, 1] \times [0, 1]$.

Lema 1. Una cópula es creciente en sus entradas

Demostración. Por ser C bicreciente las funciones $C(t, v_2) - C(t, v_1)$ y $C(u_1, t) - C(u_2, t)$ son crecientes para $u_1 \leq u_2$ y $v_1 \leq v_2$. Basta evaluar en $t = 0$, $t = u$ en la primera expresión para notar que $C(u, v_2) - C(u, v_1) - (C(0, v_1) - C(0, v_2)) \geq 0$ de donde se sigue que $C(u, v_2) - C(u, v_1) \geq 0$ siempre que $v_1 \leq v_2$. Haciendo lo mismo para la otra variable, se termina la prueba. ■

Lema 2. Sea C una cópula. Entonces C es Lipchitz-continua, esto es, para cualquier pareja de puntos $(x_1, y_1), (x_2, y_2)$ del intervalo unitario

$$|C(x_2, y_2) - C(x_1, y_1)| \leq |x_2 - x_1| + |y_2 - y_1|$$

Demostración. Por la desigualdad del triángulo, tenemos

$$|C(x_2, y_2) - C(x_1, y_1)| \leq |C(x_2, y_2) - C(x_2, y_1)| + |C(x_2, y_1) - C(x_1, y_1)| \quad (1.8)$$

Analicemos ahora los sumandos del lado derecho. Por ser creciente C en sus entradas, si $y_2 \geq y_1$ entonces

$$C(1, y_2) - C(1, y_1) \geq C(x_2, y_2) - C(x_2, y_1) \geq 0$$

Por lo tanto

$$|C(1, y_2) - C(1, y_1)| \geq |C(x_2, y_2) - C(x_2, y_1)|$$

Del mismo modo, si $y_1 \geq y_2$ tenemos

$$|C(1, y_1) - C(1, y_2)| \geq |C(x_2, y_2) - C(x_2, y_1)|$$

Como $C(1, y_2) = y_2$ y $C(1, y_1) = y_1$, las desigualdades anteriores aseguran que

$$|y_2 - y_1| \geq |C(x_2, y_2) - C(x_2, y_1)|$$

Basta aplicar el mismo razonamiento al segundo sumando de la expresión (1.8) para finalizar la prueba. ■

Corolario 1. *Una cópula es una función de distribución correspondiente a v.a. uniformes restringidas al cuadrado unitario.*

Demostración. Verifiquemos los límites

$$\begin{aligned}\lim_{u \rightarrow 0} C(u, v) &= C(0, v) = 0 \\ \lim_{v \rightarrow 0} C(u, v) &= C(u, 0) = 0 \quad \text{por continuidad} \\ \lim_{v \rightarrow 1} C(u, v) &= C(u, 1) = u \\ \lim_{u \rightarrow 1} C(u, v) &= C(1, v) = v\end{aligned}$$

Esto y el hecho de que C es bicreciente y continua en sus entradas termina la prueba. ■

Lema 3. *Sea C una cópula. Para todo $v \in [0, 1]$ la derivada parcial $\partial C / \partial u$ existe casi dondequiera (c.d.q.) y además*

$$0 \leq \frac{\partial}{\partial u} C(u, v) \leq 1 \quad (1.9)$$

Analogamente, para todo $u \in [0, 1]$ la derivada parcial $\partial C / \partial v$ existe c.d.q. y

$$0 \leq \frac{\partial}{\partial v} C(u, v) \leq 1 \quad (1.10)$$

Más aún, las funciones $u \mapsto \partial C(u, v) / \partial v$ y $v \mapsto \partial C(u, v) / \partial u$ están definidas y son crecientes c.d.q. en $[0, 1]$.

Demostración. Siendo las asignaciones

$$\begin{aligned}v &\mapsto C(u, v) \quad u \in I \\ u &\mapsto C(u, v) \quad v \in I\end{aligned}$$

crecientes, son derivables c.d.q., es decir, tiene sentido hablar de las derivadas parciales

$$\frac{\partial}{\partial u} C(u, v) \quad \frac{\partial}{\partial v} C(u, v)$$

Además, dado que la función

$$u \mapsto C(u, v_2) - C(u, v_1) \quad 0 \leq v_1 \leq v_2 \leq 1$$

es creciente (por ser C bicreciente) se tiene que es derivable de derivada no negativa c.d.q. de donde

$$u \mapsto \partial C(u, v)/\partial v \quad v \mapsto \partial C(u, v)/\partial u$$

están definidas y son crecientes c.d.q. en $[0, 1]$.

La desigualdad (1.9) se obtiene notando que

$$C(u, 0) - C(u + h, 0) = 0$$

con lo que $\partial C(u, 0)/\partial u = 0$. Del mismo modo

$$\frac{C(u, 1) - C(u + h, 1)}{h} = 1$$

con lo que $\partial C(u, 1)/\partial u = 1$. La desigualdad (1.10) se obtiene procediendo análogamente. ■

Lema 4. *Toda cópula $C: [0, 1] * \times [0, 1] \rightarrow [0, 1]$ es absolutamente continua en cada variable. Esto es, dado $\varepsilon > 0$ y una colección de subintervalos $\{[a_i, b_i]\}_{i \geq 1}$ del $[0, 1]$ existe $\delta > 0$ tal que si*

$$\sum_i (b_i - a_i) < \delta \quad \text{entonces} \quad \sum_i |C(b_i, v) - C(a_i, v)| < \varepsilon$$

para toda $v \in [0, 1]$.

Demostración. Sea $\varepsilon > 0$ y sea $\{[a_i, b_i]\}_{i \geq 1}$ una colección de subintervalos de $[0, 1]$ tal que $\sum_i (b_i - a_i) < \varepsilon$. Por el Lema 2 tenemos que

$$\sum_i |C(b_i, v) - C(a_i, v)| \leq \sum_i (|b_i - a_i| + |v - v|)$$

y como $|v - v| = 0$ y $\sum_i (b_i - a_i) < \varepsilon$ concluimos que

$$\sum_i |C(b_i, v) - C(a_i, v)| < \varepsilon$$

de modo que $\delta = \varepsilon$ es suficiente. ■

Corolario 2. *Sea $C: [0, 1] * \times [0, 1] \rightarrow [0, 1]$ una cópula. Entonces*

$$\int_0^u \frac{\partial C(x, v)}{\partial x} dx = C(u, v)$$

para toda $u \in [0, 1]$ y para toda $v \in [0, 1]$.

Demostración. Se sigue inmediatamente de la continuidad absoluta de $C_v(x) = C(x, v)$ que

$$\int_0^u \frac{\partial C(x, v)}{\partial x} dx = C(u, v) - C(0, v)$$

Como $C(0, v) = 0$, obtenemos que

$$\int_0^u \frac{\partial C(x, v)}{\partial x} dx = C(u, v)$$

como se quería demostrar. ■

Las propiedades que definen una cópula nos permiten crear nuevas cópulas a partir de cópulas conocidas. Dos métodos útiles se demuestran en los siguientes lemas.

Lema 5. Sean α y $\beta \in [0, 1]$ $\cdot \exists \cdot \alpha + \beta = 1$. Sean C_1 y C_2 cópulas, entonces la combinación convexa

$$\alpha C_1 + \beta C_2 = C_{\alpha, \beta} \text{ es una cópula}$$

Demostración. Para empezar, el dominio de $C_{\alpha, \beta}$ sigue siendo el mismo, es decir $[0, 1] \times [0, 1]$. Para un rectángulo $B = [u_1, u_2] \times [v_1, v_2]$, definamos

$$V_C(B) = C(u_1, v_1) + C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1)$$

Probar que C es bicreciente es entonces equivalente a que para cada B la cantidad $V_C(B)$ sea no-negativa. En el caso presente tenemos

$$V_{C_{\alpha, \beta}}([u_1, u_2] \times [v_1, v_2]) = \alpha V_{C_1}([u_1, u_2] \times [v_1, v_2]) + \beta V_{C_2}([u_1, u_2] \times [v_1, v_2]) \geq 0$$

Con lo cual $C_{\alpha, \beta}$ es bicreciente. Ser anclada es, de nuevo hereditario; y para verificar que sus marginales son uniformes, sean $u \in [0, 1], v \in [0, 1]$ entonces

$$\begin{cases} C_{\alpha, \beta}(u, 1) = \alpha C_1(u, 1) + \beta C_2(u, 1) = (\alpha + \beta)u = u & \text{en el primer caso} \\ C_{\alpha, \beta}(1, v) = \alpha C_1(1, v) + \beta C_2(1, v) = (\alpha + \beta)v = v & \text{en el segundo caso} \end{cases}$$

De modo que las combinaciones convexas de cópulas son, de nuevo, cópulas. De hecho, dichas combinaciones pueden estar conformadas por cualquier cantidad finita de términos. Es decir, dada una familia de cópulas $\{C_j\}_{j=1}^n$ y escalares $\alpha_i \in \mathbb{R}^+$ tales que $\sum \alpha_i = 1$ tendremos que

$$C(u, v) = \sum_{i=1}^n \alpha_i C_i(u, v)$$

es una cópula. ■

Lema 6. *El conjunto \mathcal{C} de las cópulas bivariadas es compacto relativo a la norma uniforme*

$$\|C\| := \sup \{ C(u, v) \quad t.q. \quad u, v \in [0, 1] \}$$

Demostración. Por el Teorema de Arzelà–Ascoli la compacidad es equivalente a que \mathcal{C} sea acotado, cerrado y uniformemente equicontinuo. Verifiquemos estas condiciones.

Para toda cópula C tenemos que $C(u, v) \leq 1$ con lo cual

$$\|C(u, v)\| \leq 1 \quad \forall C \in \mathcal{C}$$

de modo que \mathcal{C} es acotada. Para probar la cerradura de esta familia, tomemos una sucesión $\{C_n\}_{n=1}^{\infty} \subseteq \mathcal{C}$. Supongamos que C_n converge a C respecto a la norma uniforme, esto es

$$\|C_n - C\| \xrightarrow{n \rightarrow \infty} 0$$

Verifiquemos que C es una cópula. Dado $\varepsilon > 0$ sea $N = N(\varepsilon)$ el correspondiente a la definición de convergencia. Para $u \in [0, 1]$ tenemos

$$|C(u, 0)| = |C(u, 0) - C_n(u, 0)| < \varepsilon \quad n \geq N$$

pues $C_n(u, 0) = 0$ para todo natural n . De este modo, siendo ε arbitrario, tenemos que $C(u, 0) = 0$ para todo $u \in [0, 1]$. Análogamente para toda $v \in [0, 1]$ tenemos que $C(0, v) = 0$ con lo cual C es anclada.

Consideremos ahora $u \in [0, 1]$ y veamos que $C(u, 1) = u$.

$$|C(u, 1) - u| \leq |C(u, 1) - C_n(u, 1)| + |C_n(u, 1) - u| < \varepsilon$$

si $n \geq N$ puesto que $C_n(u, 1) = u$ para toda n . Es así que $C(u, 1) = u$. Análogamente para todo $v \in [0, 1]$ tenemos que $C(1, v) = v$ con lo cual C satisface las condiciones de frontera de una cópula.

Para probar que C es bicreciente es suficiente notar que, por la convergencia uniforme, para todo rectángulo $B \subseteq [0, 1] \times [0, 1]$ tenemos

$$V_C(B) = \lim_{n \rightarrow \infty} V_{C_n}(B) \geq 0$$

Queda probado que \mathcal{C} es un conjunto cerrado. Queda por demostrar que es equicontinuo. Para ello, recordemos que por el Lema 2 para toda cópula C

$$|C(u_1, v_1) - C(u_2, v_2)| \leq |u_1 - u_2| + |v_1 - v_2|$$

de donde es evidente la equicontinuidad.

De este modo \mathcal{C} es un conjunto compacto relativo a la norma uniforme. ■

Corolario 3. Sea $\{C_n\}_{n \in \mathbb{N}}$ una familia de cópulas y sean $\{\alpha_n\}_{n \in \mathbb{N}}$ constantes reales no-negativas tales que $\sum \alpha_n = 1$, entonces la función

$$C(u, v) = \sum_{n=1}^{\infty} \alpha_n C_n(u, v)$$

es una cópula.

Demostración. Probaremos que C es límite puntual de cópulas. Sea, para cada $m \in \mathbb{N}$

$$C_m = \sum_{n=1}^m \alpha_n C_n + \left(1 - \sum_{n=1}^m \alpha_n\right) C_{m+1}$$

Claramente C_m es una combinación convexa de cópulas, y es, por lo tanto, una cópula. Además, tomando límite conforme $m \rightarrow \infty$ obtenemos

$$\lim_{m \rightarrow \infty} C_m(u, v) = \sum_{n=1}^{\infty} \alpha_n C_n(u, v)$$

para toda $u, v \in [0, 1]$.

Como el conjunto de las cópulas bidimensionales es compacto y convexo, la convergencia puntual implica la uniforme con lo cual C es límite uniforme de cópulas y es entonces una cópula. ■

Lema 7. Sea $\{C_\theta\}$ una familia de cópulas y sea $F(\theta)$ una función de distribución. La función

$$C(u, v) = \int_{\mathbb{R}} C_\theta(u, v) dF(\theta) \quad (1.11)$$

es una cópula.

Demostración. Demostremos primero que C satisface las condiciones de frontera. Dado que para cada θ , $C_\theta(0, v) = 0$, $C_\theta(1, v) = v$ y que $\int_{\mathbb{R}} dF(\theta) = 1$ se sigue que

$$\begin{aligned} C(0, v) &= \int_{\mathbb{R}} C_\theta(0, v) dF(\theta) = \int_{\mathbb{R}} 0 dF(\theta) = 0 \\ C(1, v) &= \int_{\mathbb{R}} C_\theta(1, v) dF(\theta) = \int_{\mathbb{R}} v dF(\theta) = v \end{aligned}$$

Las condiciones de frontera para la otra variable se siguen análogamente. Ahora bien, para probar que es bicreciente veamos que si $u_1 \leq u_2, v_1 \leq v_2$ entonces

$$\begin{aligned} & C(u_1, v_1) + C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) \\ &= \int_{\mathfrak{R}} C_{\theta}(u_1, v_1) + C_{\theta}(u_2, v_2) - C_{\theta}(u_1, v_2) - C_{\theta}(u_2, v_1) dF(\theta) \geq 0 \end{aligned}$$

puesto que $C_{\theta}(u_1, v_1) + C_{\theta}(u_2, v_2) - C_{\theta}(u_1, v_2) - C_{\theta}(u_2, v_1) \geq 0$ para cada θ . ■

Consideremos la variable X_{uv} que toma el valor $C_i(u, v)$ con probabilidad p_i de modo que $\sum_i p_i = 1$. El corolario 3 nos dice que la asignación $(u, v) \mapsto \mathbb{E}(X_{uv})$ es una cópula. El lema 7 es análogo cuando la variable X_{uv} toma los valores $C_{\theta}(u, v)$ siguiendo una distribución $F(\theta)$.

Proposición 1. *Dadas dos variables aleatorias X y Y cuyas funciones de distribución univariadas F_X y F_Y respectivamente sean continuas, y dada una cópula C , la función*

$$F(x, y) = C(F_X(x), F_Y(y)) \tag{1.12}$$

es de distribución, y sus marginales son F_X y F_Y .

Demostración. Como C toma valores de I^2 y los mapea en I , la composición $C(F_X(x), F_Y(y))$ tiene sentido y además

$$\forall x, y \in \mathfrak{R} \quad C(F_X(x), F_Y(y)) \in [0, 1]$$

Como para todo $u, v \in [0, 1]$ $C(u, 0) = C(0, v)$ y por el lema 2 implica que:

$$\lim_{x \rightarrow -\infty} F(x, y) = \lim_{x \rightarrow -\infty} C(F_X(x), F_Y(y)) = C(0, F_Y(y)) = 0$$

Análogamente $\lim_{y \rightarrow -\infty} F(x, y) = 0$. Ahora bien, si $x_1 \leq x_2$ $y_1 \leq y_2$ son puntos en \mathfrak{R} , entonces

$$\begin{aligned} & F(x_1, y_1) + F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) = \\ & C(F_X(x_1), F_Y(y_1)) + C(F_X(x_2), F_Y(y_2)) \\ & \quad - C(F_X(x_1), F_Y(y_2)) - C(F_X(x_2), F_Y(y_1)) \end{aligned}$$

Como F_X y F_Y son crecientes y C es bicreciente lo anterior resulta no-negativo. De este modo

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

Siendo F una composición de una función continua con funciones continuas por la derecha, se sigue que es continua por la derecha en cada variable.

Queda probado que $F(x, y) = C(F_X(x), F_Y(y))$ es una función de distribución y dado que C tiene marginales uniformes,

$$\lim_{x \rightarrow \infty} F(x, y) = \lim_{x \rightarrow \infty} C(F_X(x), F_Y(y)) = C(1, F_Y(y)) = F_Y(y) \quad \forall y \quad (1.13)$$

$$\lim_{y \rightarrow \infty} F(x, y) = \lim_{y \rightarrow \infty} C(F_X(x), F_Y(y)) = C(F_X(x), 1) = F_X(x) \quad \forall x. \quad (1.14)$$

De modo que dicha función de distribución tiene como marginales a las funciones F_X y F_Y . ■

Hasta ahora hemos probado que dada una función de distribución bivariada $F(x, y)$ de marginales F_X , F_Y ,

$$F(x, y) = C(F_X(x), F_Y(y)) \Leftrightarrow C \text{ es una cópula (o subcópula)}$$

Nuestra pregunta ahora es ¿Cuáles distribuciones satisfacen una ecuación de este estilo? El teorema de Sklar afirma que cualquiera, y afirma así que $F(x, y)$ es una distribución bivariada si y sólo si existe una cópula C tal que para cualquier punto (x, y) se satisface que $F(x, y) = C(F_X(x), F_Y(y))$.

1.4 Teorema de Sklar.

Teorema 1.2 (Teorema de Sklar, (1959)). *Sea $F_{X,Y}$ una función de distribución bivariada con marginales F_X y F_Y . Entonces existe una cópula tal que*

$$F(x, y) = C(F_X(x), F_Y(y)) \quad \forall x, y \in \mathfrak{R} \quad (1.15)$$

Además, si F_X y F_Y son continuas, entonces C es única. De otro modo, queda únicamente determinada sobre $\text{Ran}F_X \times \text{Ran}F_Y$.

Demostración. El método de la prueba es el siguiente: Para las variables X, Y se dará de modo explícito una subcópula que asegure que $C(F(x, y)) = F_X(x)F_Y(y)$ sobre $RanF_X \times RanF_Y$. Después extenderemos esta subcópula hacia una cópula en $[0, 1] \times [0, 1]$.

Pimera parte: Construcción de una subcópula

Por ser $F_{X,Y}$ una función de distribución bivariada con marginales F_X, F_Y tenemos que para cualesquiera puntos $(x_1, y_1), (x_2, y_2) \in \mathfrak{R}^2$

$$|F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_1)| \leq |F_X(x_2) - F_X(x_1)| + |F_Y(y_2) - F_Y(y_1)|$$

De lo anterior, si $F_X(x_1) = F_X(x_2)$ y $F_Y(y_2) = F_Y(y_1)$ entonces también $F_{X,Y}(x_2, y_2) = F_{X,Y}(x_1, y_1)$. Siendo así, la asignación

$$((F_X(x), F_Y(y)) \xrightarrow{C^*} F_{X,Y}(x, y)) \quad (1.16)$$

es una función real con dominio $RanF_X \times RanF_Y$. De nuevo, por ser $F_{X,Y}$ función de distribución se tiene, directamente, que C^* es una subcópula.

Seunda Parte: De una subcópula a una cópula

Por continuidad de la subcópula, podemos extender hacia las cerraduras $\overline{S_1}$ y $\overline{S_2}$ tomando, simplemente, límites. La función así obtenida, $C^{**} : \overline{S_1} \times \overline{S_2} \rightarrow [0, 1]$, será también una subcópula.

Tomemos cualquier elemento $(x, y) \in [0, 1] \times [0, 1]$. Sean

$$x_1 = \sup\{z \in S_1 | z \leq x\} \quad (1.17)$$

$$x_2 = \inf\{z \in S_1 | z \geq x\} \quad (1.18)$$

$$y_1 = \sup\{w \in S_2 | w \leq y\} \quad (1.19)$$

$$y_2 = \inf\{w \in S_2 | w \geq y\} \quad (1.20)$$

Observación 1. Si $x \in \overline{S_1} \Rightarrow x_1 = x = x_2$ y análogamente $y \in \overline{S_2} \Rightarrow y_1 = y = y_2$.

Recordemos que el único trabajo que tenemos que realizar es *extender el dominio de C^{**} preservando sus propiedades*. Sabiendo que cualquier combinación convexa de cópulas es una cópula, es natural pensar en la interpolación

bilineal como una posibilidad. Así, sean $\{\alpha_i\}_{i=1}^2$ definidas como sigue

$$\alpha_1 = \begin{cases} (x - x_1)/(x_2 - x_1), & \text{si } x_1 < x_2, \\ 1, & \text{si } x_1 = x_2. \end{cases} \quad (1.21)$$

$$\alpha_2 = \begin{cases} (y - y_1)/(y_2 - y_1), & \text{si } y_1 < y_2, \\ 1, & \text{si } y_1 = y_2. \end{cases} \quad (1.22)$$

Sea

$$C(x, y) = (1 - \alpha_1)(1 - \alpha_2)C^{**}(x_1, y_1) + (1 - \alpha_1)\alpha_2C^{**}(x_1, y_2) \\ + \alpha_1(1 - \alpha_2)C^{**}(x_2, y_1) + \alpha_1\alpha_2C^{**}(x_2, y_2). \quad (1.23)$$

Claramente $DomC = [0, 1] \times [0, 1]$ y además, por la observación (1) C extiende a C^{**} . Como $C^{**}(0, v) = C^{**}(u, 0) = 0$, se concluye lo mismo para C . Por último, gracias a que C^{**} tiene marginales uniformes se sigue que C las tiene también (pues es una combinación convexa). Lo único que resta es notar que la función C es bicreciente. Para ello, tomemos una caja $[u_1, u_2] \times [v_1, v_2] \subseteq [0, 1] \times [0, 1]$.

Consideremos los puntos correspondientes a ella según (1.17) denotados por x_{11}, x_{21} para u_1 ; y_{11}, y_{21} para v_1 y análogamente para u_2, v_2 . Tomemos además los escalares $\alpha_{11}, \alpha_{21}, \alpha_{12}, \alpha_{22}$ definidos mediante (1.21) para los puntos (u_1, v_1) y (u_2, v_2) respectivamente. El caso más simple es aquel en que no hay puntos de \bar{S}_1 en el intervalo (u_1, u_2) y $\bar{S}_2 \cap (v_1, v_2) = \emptyset$ en cuyo caso $x_{11} = x_{12}$, $x_{21} = x_{22}$, $y_{11} = y_{12}$, $y_{21} = y_{22}$, con lo cual, evaluando y simplificando

$$V_C([u_1, u_2] \times [v_1, v_2]) = (\alpha_{12} - \alpha_{11})(\alpha_{22} - \alpha_{21})V_C([x_{11}, x_{21}] \times [y_{11}, y_{21}]).$$

Como $u_1 \leq u_2$ y $v_1 \leq v_2$ tendemos que $\alpha_{11} \leq \alpha_{12}$ y que $\alpha_{21} \leq \alpha_{22}$, de modo que $V_C([u_1, u_2] \times [v_1, v_2]) \geq 0$. Por otro lado, el caso más complicado es aquél en que $\bar{S}_1 \cap (u_1, u_2) \neq \emptyset$ y $\bar{S}_2 \cap (v_1, v_2) \neq \emptyset$. En este caso, $u_1 < x_{21} \leq x_{12} < u_2$, $v_1 < y_{21} \leq y_{12} < v_2$. Tras evaluar y simplificar obtenemos

$$V_C([u_1, u_2] \times [v_1, v_2]) = (1 - \alpha_{11})\alpha_{22}V_C([x_{11}, x_{21}] \times [y_{12}, y_{22}]) \\ + \alpha_{22}V_C([x_{21}, x_{12}] \times [y_{12}, y_{22}]) + \alpha_{12}\alpha_{22}V_C([x_{12}, x_{22}] \times [y_{12}, y_{22}]) \\ + (1 - \alpha_{11})V_C([x_{11}, x_{21}] \times [y_{21}, y_{12}]) + V_C([x_{21}, x_{12}] \times [y_{21}, y_{12}]) \\ + \alpha_{12}V_C([x_{12}, x_{22}] \times [y_{21}, y_{12}]) + (1 - \alpha_{11})(1 - \alpha_{21})V_C([x_{11}, x_{21}] \times [y_{11}, y_{21}]) \\ + (1 - \alpha_{21})V_C([x_{21}, x_{12}] \times [y_{11}, y_{21}]) + \alpha_{12}(1 - \alpha_{21})V_C([x_{12}, x_{22}] \times [y_{11}, y_{21}])$$

Siendo una combinación no negativa de cantidades no negativas, se concluye que $V_C([u_1, u_2] \times [v_1, v_2]) \geq 0$. El resto de los casos son similares. Queda probado que C es bicreciente.

En cuanto a la unicidad, si F_X y F_Y son continuas, entonces, por el teorema del valor intermedio, $RanF = RanG = [0, 1]$ de modo que la subcópula de la parte inicial de la prueba es, de hecho, una cópula que resulta, según la construcción, única.

Si acaso F_X o F_Y es(son) discontinua(s) entonces la cópula no es única (pues viene, como se ve en la segunda parte de la demostración, de alguna combinación convexa). Sin embargo, dado que sobre $RanF \times RanG$ el método para hallarla es el de la parte primera de la demostración, sobre este conjunto hay unicidad.

Esto completa la prueba. ■

Esta prueba nos brinda un método para construir cópulas a partir de funciones de distribución. Para abordarlo definamos la función quasi-inversa de una distribución F como sigue

$$F^{-1}(t) = \begin{cases} x, & \text{si } F(x) = t, \\ \inf\{x|F(x) \geq t\} = \sup\{x|F(x) \leq t\}, & \text{e.o.c.} \end{cases}$$

Observación 2. En esta definición, si para algún $t \in [0, 1]$ fijo se tiene que el conjunto $\{x \text{ t.q. } F(x) = t\}$ tiene más de un elemento, se elige uno único al asignar $F^{-1}(t) = x$.

Corolario 4. Si F_{XY} es una función de distribución conjunta y F_X, F_Y son sus marginales y C^* la subcópula del teorema de Sklar, entonces, para cada $(u, v) \in DomC^*$ se tiene que

$$C^*(u, v) = F_{XY}(F_X^{-1}(u), F_Y^{-1}(v))$$

Demostración. Basta tomar el cambio de variable $u = F_X(x), v = F_Y(y)$. Observemos que si las marginales son continuas, entonces la igualdad se verifica para la cópula de X, Y . ■

Ejemplo 1.3. Podemos formar una cópula tomando la función de distribución normal bivariada de parámetros $(\mu = (0, 0), \sigma^2 = (1, 1), \rho)$ de densidad

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{(-x^2 - 2\rho xy + y^2)}{2(1-\rho^2)}\right)$$

para $x, y \in \mathfrak{R}$.

Denotemos por ϕ a la distribución normal estándar. Entonces, como las marginales de f son normales estándar, la función

$$G(u, v) = \int_{-\infty}^{\phi^{-1}(u)} \int_{-\infty}^{\phi^{-1}(v)} f(s, t) ds dt$$

es una cópula de acuerdo con el corolario 4. Dicha cópula se llama Gaussiana.

Obsérvese que una pareja (X, Y) con función de distribución $G(\phi(x), \phi(y))$ es conjuntamente normal con coeficiente de correlación ρ . ■

Lema 8. *Dada una variable aleatoria continua X con distribución F , la distribución de la composición $F(X)$ es uniforme en $[0, 1]$*

Demostración. Como F es una función de distribución, $F(X)$ toma valores en $[0, 1]$. Ahora bien, para cada $x \in [0, 1]$,

$$\mathbb{P}(F(X) \leq x) = \mathbb{P}(X \leq F^{-1}(x)) = F(F^{-1}(x)) = x$$

con lo que $F(X) \sim U(0, 1)$. ■

Proposición 2. *Dadas dos variables aleatorias continuas, la cópula C que describe su distribución (según el teorema de Sklar) es la distribución de las variables aleatorias $F_X(X), F_Y(Y)$ restringida al cuadrado unitario I^2 .*

Demostración. Como $F_X(X) \sim U(0, 1)$ y $F_Y(Y) \sim U(0, 1)$, sean $u, v \in I$, entonces

$$\begin{aligned} \mathbb{P}[F_X(X) \leq u, F_Y(Y) \leq v] &= \mathbb{P}[X \leq F_X^{-1}(u), Y \leq F_Y^{-1}(v)] \\ &= F(F_X^{-1}(u), F_Y^{-1}(v)) \end{aligned}$$

Y por el teorema de Sklar esto es $C(u, v)$. ■

Usemos ahora el teorema de Sklar para entender más a fondo lo que es una cópula. La función de distribución de un vector (X, Y) explica completamente el comportamiento conjunto de las v.a. X y Y . ¿Qué información se encierra en esta distribución? En primer lugar, la información marginal, es decir, el comportamiento individual de las variables X y Y . En segundo lugar, la relación que guardan las variables sin tomar en cuenta sus marginales.

Un modo de intuir esto es pensando en la densidad conjunta para las variables X, Y . Si C es derivable

$$\begin{aligned} f(x, y) &= \frac{\partial^2 F(x, y)}{\partial x \partial y} \\ &= \frac{\partial^2 C(F_X(x), F_Y(y))}{\partial F_X \partial F_Y} \cdot \frac{\partial F_X(x)}{\partial x} \cdot \frac{\partial F_Y(y)}{\partial y} \\ &= f_X(x) \cdot f_Y(y) \cdot c(F_X(x), F_Y(y)) \quad \forall (x, y) \in \mathfrak{R}^2 \end{aligned} \quad (1.24)$$

donde $c(u, v) = \frac{\partial^2}{\partial u \partial v} C(u, v)$.

Este cálculo expresa que siempre que la densidad conjunta de la pareja X, Y exista, ésta es el producto de la densidad en caso de independencia, $f_X(x)f_Y(y)$, multiplicada por una función $c(u, v)$ relacionada con la cópula de (X, Y) .

1.5 Cópulas y dependencia bivariada

En la sección anterior se demuestra el teorema de Sklar y con él se intuye que existen aspectos de la dependencia entre las entradas de un vector aleatorio que se resumen en su cópula. De ser así, ¿cuáles de estos aspectos podemos identificar plenamente? Esto es, sabiendo que la cópula de (X, Y) es C , ¿qué conocimiento tenemos sobre la relación de dependencia que guardan sus entradas? Esta pregunta es muy amplia para responderla totalmente. Estudiaremos ahora algunos casos.

Proposición 3. *La función definida como*

$$\Pi(u, v) = uv \quad \forall (u, v) \in [0, 1] \times [0, 1]$$

es una cópula.

Demostración. Es claro que su dominio es el adecuado. Además, que sea biceciente se sigue de que, dados $u_1 \leq u_2 \quad v_1 \leq v_2$

$$u_1 v_1 + u_2 v_2 - u_1 v_2 - u_2 v_1 = u_2(v_2 - v_1) - u_1(v_2 - v_1) = (u_2 - u_1)(v_2 - v_1) \geq 0$$

Es anclada pues $u \cdot 0 = 0 \cdot v = 0$ para cualesquiera u, v . Sus marginales son uniformes pues $u \cdot 1 = u$ y $1 \cdot v = v$ para $u, v \in \mathfrak{R}$. ■

Para mostrar su importancia, tomemos dos variables aleatorias continuas, X, Y y demostremos que son independientes si y sólo si su cópula (esto es, la correspondiente a su distribución según el teorema de Sklar) es la cópula producto.

La definición misma de independencia entre dos variables aleatorias implica que

$$X \text{ es independiente de } Y \Leftrightarrow \forall(x, y) \quad F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y)$$

Queda claro entonces que, utilizando el teorema de Sklar, tenemos

$$X \text{ es independiente de } Y \Leftrightarrow \forall(x, y) \quad C\left(F_X(x), F_Y(y)\right) = F_X(x) \cdot F_Y(y)$$

Utilizando la transformación $F_X(x) = u, F_Y(y) = v$ y la continuidad de las distribuciones F_X, F_Y es claro que lo anterior es equivalente a que

$$\forall u, v \quad C(u, v) = u \cdot v$$

■

Es también importante observar que el conjunto de las cópulas contiene cotas, tanto inferior como superior. Esto es, existen dos cópulas que denotaremos respectivamente W y M tales que $W \leq C \leq M$ para toda C cópula pensando este orden como resultado de la comparación de las evaluaciones en todo punto como se prueba en el siguiente teorema.

Teorema 1.3. *Sea C una cópula. Entonces, para cada $(u, v) \in \text{Dom}(C)$*

$$W(u, v) := \max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v) := M(u, v) \quad (1.25)$$

Además, $W(u, v)$ y $M(u, v)$ son cópulas. Dichas cópulas reciben el nombre de cotas -superior e inferior- de Fréchet-Hoeffding.

Demostración. Sean $u, v \in [0, 1]$, entonces, por ser C bicreciente

$$C(1, 1) + C(u, v) - C(1, v) - C(u, 1) \geq 0 \Rightarrow C(u, v) \geq u + v - 1$$

Análogamente

$$C(u, v) = C(u, v) + C(0, 0) - C(0, v) - C(u, 0) \geq 0$$

Con lo que obtenemos

$$\forall u, v \in [0, 1]: \quad C(u, v) \geq \max(u + v - 1, 0)$$

Dado que C es una distribución, es creciente en sus argumentos,

$$C(u, v) \leq C(u, 1) \quad \text{y} \quad C(u, v) \leq C(1, u) \Rightarrow C(u, v) \leq \min(u, v)$$

De donde (1.25) es válida. Queda sólo por demostrar que estas cotas son cópulas.

Partiendo de su definición, tenemos:

$$i) \quad \text{Dom}(W) = \text{Dom}(M) = [0, 1] \times [0, 1]$$

$$ii) \quad \text{Si } u \leq 1, v \leq 1 \text{ entonces } W(0, v) = \max(v - 1, 0) = 0 \quad \text{y} \quad W(u, 0) = \max(u - 1, 0) = 0.$$

$$\text{Para } M, \text{ tomemos } u \geq 0, v \geq 0. \text{ Entonces } M(0, v) = \min(0, v) = 0 = \min(u, 0) = M(u, 0)$$

Queda demostrar que son bicrecientes. Para ello, consideremos $u_1 \leq u_2$ y $v_1 \leq v_2$. Para M tendemos

$$\begin{aligned} V_M([u_1, v_1] \times [u_2, v_2]) &= M(u_1, v_1) + M(u_2, v_2) - M(u_1, v_2) - M(u_2, v_1) \\ &= u_1 \wedge v_1 + u_2 \wedge v_2 - u_1 \wedge v_2 - u_2 \wedge v_1 \\ &= \begin{cases} u_2 - u_1 & \text{si } v_1 \leq u_1 \leq u_2 \leq v_2 \\ v_2 - u_1 & \text{si } v_1 \leq u_1 \leq v_2 \leq u_2 \\ v_2 - v_1 & \text{si } u_1 \leq v_1 \leq v_2 \leq u_2 \\ u_2 - v_1 & \text{si } u_1 \leq v_1 \leq u_2 \leq v_2 \end{cases} \end{aligned}$$

De donde M es bicreciente. Para probar que W lo es observemos que $W(u, v) = 0$ en todo punto (u, v) tal que $u + v - 1 \leq 0$ es decir $u \leq 1 - v$. Así, W se anula bajo la contradiagonal $u = 1 - v$. De este modo, si los cuatro puntos u_1, u_2, v_1, v_2 están en esta región tenemos

$$W(u_1, v_1) + W(u_2, v_2) - W(u_1, v_2) - W(u_2, v_1) = 0$$

Ahora bien, si todos los puntos están en el triángulo $(1, 1), (1, 0), (0, 1)$ entonces

$$\begin{aligned} V_W([u_1, v_1] \times [u_2, v_2]) &= W(u_1, v_1) + W(u_2, v_2) - W(u_1, v_2) - W(u_2, v_1) \\ &= (u_1 + v_1 - 1) + (u_2 + v_2 - 1) - (u_1 + v_2 - 1) - (u_2 + v_1 - 1) \\ &= 0 \end{aligned}$$

Si tan sólo un vértice está en el triángulo superior (debería ser el (u_2, v_2)) entonces

$$V_W([u_1, v_1] \times [u_2, v_2]) = u_2 + v_2 - 1 \geq 0$$

Si son dos los vértices en la región (por lo tanto (u_2, v_2) y (u_2, v_1) o (u_2, v_2) y (u_1, v_2)) entonces

$$V_W([u_1, v_1] \times [u_2, v_2]) = \begin{cases} v_2 - v_1 & \text{en un caso,} \\ u_2 - u_1 & \text{en el otro.} \end{cases}$$

Así, W es bicreciente. ■

Una aplicación importante de estas cópulas y el teorema de Sklar es que podemos acotar cualquier función de distribución a partir de sus marginales, en particular,

$$\max(F(x) + G(y) - 1, 0) \leq H(x, y) \leq \min(F(x), G(y))$$

Una pregunta natural es la de cuándo dichas cotas se alcanzan y qué reflejan de la pareja (X, Y) .

Para ello, consideremos el vector $(X, Y) \stackrel{c.s.}{=} (X, X)$, entonces

$$\begin{aligned} F_{X,Y}(x, y) &= \mathbb{P}(X \leq x; Y \leq y) = \mathbb{P}(X \leq x, X \leq y) = \mathbb{P}(X \leq \min(x, y)) \Rightarrow \\ &= \mathbb{P}(X \leq \min(x, y)) = \min\left(\mathbb{P}(X \leq x); \mathbb{P}(X \leq y)\right) = C(F_X(x), F_X(y)) \end{aligned}$$

Es decir, $C(u, v) = \min(u, v) = M(u, v)$

Por otro lado, si X es una variable aleatoria uniforme en $(0, 1)$ y $(X, Y) \stackrel{c.s.}{=} (X, 1 - X)$ entonces

$$\begin{aligned} F_{X,Y} &= \mathbb{P}(X \leq x; Y \leq y) = \mathbb{P}(X \leq x, 1 - X \leq y) = \mathbb{P}(1 - y \leq X \leq x) \\ &= \begin{cases} 0 & \text{si } 1 - y \geq x \\ x + y - 1 & \text{si } 1 - y \leq x \end{cases} \end{aligned}$$

Es decir, $F_{X,Y}(x, y) = \max(x + y - 1, 0) = W(F_X(x), F_Y(y))$.

De lo anterior, podemos preguntarnos si estos son los únicos casos en que las cotas de Fréchet se alcanzan o bajo que condiciones se obtienen, en otras palabras, si la cópula es una cota de Fréchet, ¿qué tipo de relación guardan las variables aleatorias en cuestión.

Teorema 1.4. Sean X, Y variables aleatorias con función de distribución continua. Entonces

$$\begin{aligned} C_{X,Y} = M &\Leftrightarrow \min(\mathbb{P}(X \leq x, Y > y), \mathbb{P}(X > x, Y \leq y)) = 0 \\ C_{X,Y} = W &\Leftrightarrow \min(\mathbb{P}(X \leq x, Y \leq y), \mathbb{P}(X > x, Y > y)) = 0 \end{aligned} \quad (1.26)$$

Demostración. Por la regla de la probabilidad total,

$$\begin{aligned} F(x) = \mathbb{P}(X \leq x) &= \mathbb{P}(X \leq x, Y \leq y) + \mathbb{P}(X \leq x, Y > y) \\ &= H(x, y) + \mathbb{P}(X \leq x, Y > y) \end{aligned} \quad (1.27)$$

Y haciendo lo mismo para la otra variable, tenemos

$$\begin{aligned} H(x, y) &= \min(F(x), G(y)) \Leftrightarrow \\ H(x, y) &= \min(H(x, y) + \mathbb{P}(X \leq x, Y > y), H(x, y) + \mathbb{P}(X > x, Y \leq y)) \\ &= H(x, y) + \min(\mathbb{P}(X \leq x, Y > y), \mathbb{P}(X > x, Y \leq y)) \end{aligned}$$

De donde

$$\begin{aligned} H(x, y) &= \min(F(x), G(y)) \quad \text{si y sólo si} \\ &\quad \min(\mathbb{P}(X \leq x, Y > y), \mathbb{P}(X > x, Y \leq y)) = 0 \end{aligned}$$

para todo punto (x, y) .

Ahora bien, podemos hacer lo mismo con la cota inferior. En este caso, tenemos

$$\begin{aligned} H(x, y) &= \max(F(x) + G(y) - 1) \\ &= \frac{1}{2} \left(F(x) + G(y) - 1 + |F(x) + G(y) - 1| \right) \end{aligned}$$

De donde, usando las relaciones obtenidas en (1.27) tenemos

$$\begin{aligned} 2H(x, y) + \mathbb{P}(X \leq x, Y > y) + \mathbb{P}(X > x, Y \leq y) \\ = \pm(1 - \mathbb{P}(X \leq x, Y > y) - \mathbb{P}(X > x, Y \leq y)) \end{aligned}$$

Y así $H(x, y) = 0$ o $H(x, y) = 1 - \mathbb{P}(X \leq x, Y > y) - \mathbb{P}(X > x, Y \leq y)$. Es decir,

$$H(x, y) = \max(F(x) + G(y) - 1, 0) \quad \text{si y sólo si} \\ \min(\mathbb{P}(X \leq x, Y \leq y), \mathbb{P}(X > x, Y > y)) = 0$$

■

Podemos ahondar un poco más recordando que toda cópula es una función de distribución con marginales uniformes y como tal, cuenta con un soporte.

Definición 3. *El soporte es el complemento de la unión de los conjuntos abiertos de probabilidad 0, es decir, si \mathcal{S}_F denota el soporte de la distribución F entonces, nombrando τ al conjunto de abiertos y considerando*

$$\mathcal{A} = \{ A \in \tau \quad \text{t.q.} \quad \mathbb{P}(A) = 0 \} \quad \text{tenemos} \quad \mathcal{S}_F = \left(\bigcup_A A \right)^c$$

Siendo de este modo, para cualquier evento A

$$\mathbb{P}(A) = \mathbb{P}(A \cap \mathcal{S}_F)$$

puesto que $\mathbb{P}(A \cap \mathcal{S}_F^c) = 0$. De este modo, el soporte es el conjunto en que se acumula la probabilidad en el sentido que si algún evento no le intersecta, entonces es \mathbb{P} -nulo.

Ejemplo 1.4. *Consideremos la distribución*

$$M(u, v) = \begin{cases} 0, & \text{si } \max(u, v) \leq 0, \\ \min(u, v), & \text{si } (u, v) \in (0, 1)^2, \\ u, & \text{si } u \in (0, 1), v \geq 1, \\ v, & \text{si } u \geq 1, v \in (0, 1), \\ 1, & \text{en otro caso.} \end{cases}$$

Su soporte es el conjunto $\{(x, y) \in [0, 1] \quad \text{t.q.} \quad x = y\}$.

Demostración. Sea $R = (u_1, v_1) \times (u_2, v_2)$ un rectángulo abierto de \mathbb{R}^2 . Si dicho rectángulo está contenido en $\mathbb{R}^2 \setminus [0, 1]^2$ entonces su M -medida es cero pues en este caso $M(u_2, v_2) = M(u_2, v_1)$ y también $M(u_1, v_1) = M(u_1, v_2)$. Si $R \cap [0, 1]^2 \neq \emptyset$ y $R \cap \mathbb{R}^2 \setminus [0, 1]^2 \neq \emptyset$ podemos expresar a R como la unión

de los rectángulos disjuntos $R_1 = R \cap [0, 1]^2$ y $R_2 = R \setminus R_1$ y por lo anterior la medida de R coincide con la de R_1 . Es así que basta considerar rectángulos completamente contenidos en $[0, 1]^2$. Sea $S = \{(x, y) \in [0, 1]^2 \text{ t.q. } x = y\}$. Si $R \cap S = \emptyset$ entonces para todo $(x, y) \in R$ o bien $x > y$ o bien $x < y$. Siendo así $\min(u_1, v_1) = \min(u_1, v_2)$ y también $\min(u_2, v_2) = \min(u_2, v_1)$ de modo que la medida de R es nula.

Si, por el contrario, $R \cap S \neq \emptyset$ entonces la M -medida de R es la misma que la del rectángulo $R' = (u_1, u_2) \times (u_1, u_2)$. Evaluando según la definición de M esta es $u_2 - u_1 > 0$ siempre que $u_2 > u_1$.

Para concluir la prueba, consideremos A un conjunto abierto de M -medida cero. Entonces $(x, y) \in A$ si y sólo si existe R un rectángulo abierto tal que $x \in R \subseteq A$. Por la contención tenemos que R tiene M -medida cero y con ello $R \cap S = \emptyset$. Se concluye que $A \cap S = \emptyset$ de donde S es el soporte de M . ■

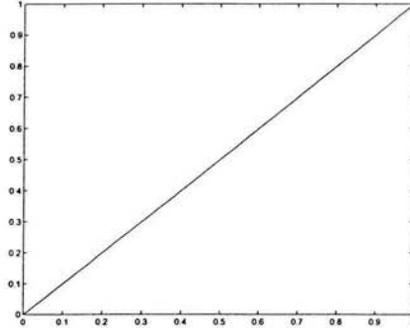


Figura 1.1: Soporte de la distribución $M(u, v)$

Una propiedad importante del soporte que usaremos para interpretar las cotas de Fréchet es el siguiente lema.

Lema 9. Si X, Y son v.a. continuas con distribución $F_{X, Y}$ entonces $Y \stackrel{c.s.}{=} g(X)$ si y sólo si el soporte de $F_{X, Y}$ es la gráfica de dicha función, es decir, $\mathcal{S}_F = \{(x, y) \text{ t.q. } y = g(x)\}$.

Demostración. Supongamos primero que $Y \stackrel{c.s.}{=} g(X)$, y sea A un evento tal que $A \cap \{(x, g(x)) \text{ t.q. } x \in \mathfrak{R}\} = \emptyset$. En este caso $\mathbb{P}(X \in A) =$

0 puesto que $A \subseteq \{(x, g(x)) \text{ t.q. } x \in \mathfrak{R}\}^c$ de modo que $0 \leq \mathbb{P}(A) \leq \mathbb{P}(\{(x, g(x)) \text{ t.q. } x \in \mathfrak{R}\}^c) = 0$.

Queda por demostrar que si el soporte es la gráfica de una cierta función, entonces X y Y se relacionan vía esa misma función. Para ello basta notar que

$$\mathbb{P}[Y = g(X)] = \iint_{\{y=g(x)\}} dF_{X,Y}(x, y)$$

y así si $\mathcal{S}_F = \{(x, y) \text{ t.q. } y = g(x)\}$ se sigue inmediatamente que $Y \stackrel{c.s.}{=} g(X)$. ■

Volviendo a las cotas de Fréchet tenemos los siguientes teoremas.

Teorema 1.5. Sean X, Y dos v.a. continuas no degeneradas y sea C su cópula según el teorema de Sklar. Entonces $C = M$ si y sólo si existe g una función creciente tal que

$$Y \stackrel{c.s.}{=} g(X)$$

Demostración. Según el Teorema 1.26, si $(x, y) \in \mathfrak{R}^2$, entonces $\min(\mathbb{P}(X \leq x, Y > y), \mathbb{P}(X > x, Y \leq y)) = 0$. Es decir, o bien $\mathbb{P}(X \leq x, Y > y) = 0$ o bien $\mathbb{P}(X > x, Y \leq y) = 0$. Obsérvese que esto es equivalente a que

$$\begin{aligned} \{(u, v) \text{ t.q. } u \leq x, v > y\} \cap \mathcal{S} &= \emptyset \quad \text{o bien} \\ \{(u, v) \text{ t.q. } u > x, v \leq y\} \cap \mathcal{S} &= \emptyset \end{aligned}$$

Es decir, $(u, v) \in \mathcal{S}$ si y sólo si $\forall (x, y) \in \mathfrak{R}^2$

$$\begin{aligned} u \leq x \Rightarrow v \leq y \quad \text{o bien} \\ v \leq y \Rightarrow u \leq x \end{aligned} \tag{1.28}$$

Demostraremos ahora que lo anterior implica que el conjunto \mathcal{S}_F satisface que

$$\forall (x, y), (u, v) \in \mathcal{S}_F \quad \text{si } x < u \Rightarrow y \leq v$$

Para ello, supongamos lo contrario. Entonces, existen puntos $(u_1, v_1), (u_2, v_2) \in \mathcal{S}$ con $u_1 < u_2, v_1 > v_2$. Para el punto medio entre ellos (x, y) de \mathfrak{R}^2 las relaciones (1.28) no se satisfacen.

Esto quiere decir que los puntos (x, y) que conforman el soporte de la distribución corresponden a la gráfica de una función creciente. ■

Tenemos un resultado análogo para la cota inferior W .

Teorema 1.6. Sean X, Y dos v.a. continuas no degeneradas y sea C su cópula según el teorema de Sklar. Entonces $C = W$ si y sólo si existe h una función decreciente tal que

$$Y \stackrel{c.s.}{=} h(X)$$

Demostración. En este caso, el Teorema 1.26 es equivalente con que

$$\begin{aligned} \{(u, v) \text{ t.q. } u \leq x, v \leq y\} \cap \mathcal{S} &= \emptyset \quad \text{o bien} \\ \{(u, v) \text{ t.q. } u > x, v > y\} \cap \mathcal{S} &= \emptyset \end{aligned}$$

De donde $(u, v) \in \mathcal{S}$ si y sólo si $\forall (x, y) \in \mathbb{R}^2$

$$\begin{aligned} u \leq x &\Rightarrow v > y \quad \text{o bien} \\ v > y &\Rightarrow u \leq x \end{aligned} \tag{1.29}$$

Lo anterior implica que para todo $(u, v), (x, y) \in \mathcal{S}$ si $x < u \Rightarrow y \geq v$. Para demostrarlo supongamos lo contrario. Podemos entonces hallar puntos $(u_1, v_1), (u_2, v_2) \in \mathcal{S}$ tales que $u_1 < u_2$ y $v_1 < v_2$. Para el punto medio entre ellos, (x, y) , las relaciones (1.29) no se ven satisfechas. ■

Observación 3. Si U, V son variables aleatorias uniformes entonces la cópula de (U, V) es M si y sólo si $U \stackrel{c.s.}{=} V$. Tomando X, Y variables aleatorias continuas sabemos que $F_X(X), F_Y(Y)$ son variables aleatorias uniformes. Además, por la proposición 2 sabemos que si (X, Y) tiene cópula C entonces C es la cópula de $(F_X(X), F_Y(Y))$. Por lo tanto, dadas dos variables aleatorias X, Y , su cópula es M si y sólo si $F_X(X) \stackrel{c.s.}{=} F_Y(Y)$.

Análogamente la cópula de (X, Y) es W si y sólo si $F_X(X) + F_Y(Y) \stackrel{c.s.}{=} 1$. Con estos razonamientos podemos enunciar de otro modo el teorema anterior.

Teorema 1.7. Sean X, Y dos v.a. continuas no degeneradas y sea C su cópula según el teorema de Sklar. Entonces si F_Y es invertible

$$\begin{aligned} C &= M \text{ si y sólo si } Y = F_Y^{-1}(F_X(X)) \\ C &= W \text{ si y sólo si } Y = F_Y^{-1}(1 - F_X(X)) \end{aligned}$$

■

1.6 Reordenamientos de M

Un reordenamiento de M es una cópula cuyo soporte es una alteración específica del soporte de M . El procedimiento será el siguiente. Tomaremos una partición $\mathcal{P} = \{P_i\}_{i=1}^n$ de intervalos cerrados de $[0, 1]$. Dividimos el $[0, 1]^2$ con los rectángulos $\mathcal{Q} = P_i \times [0, 1]$, entonces para cada uno de ellos hay una parte del soporte de M , a saber, la intersección de $P_i \times [0, 1]$ con la gráfica de la función identidad. Permutamos estos rectángulos transformándolos en $\pi(\mathcal{Q}) = \{\pi(P_i) \times [0, 1]\}_{i=1}^n$. Con ello hemos cambiado el soporte de lugar. Tomamos ahora una función $w: \pi(\mathcal{P}) \rightarrow \{-1, 1\}$ de modo que $w(\pi(P_i)) = 1$ si dejamos la recta correspondiente del soporte como está y $w(\pi(P_i)) = -1$ si hacemos negativa su pendiente (respetando su magnitud). Denotaremos a cada reordenamiento como $M(n, \{P_i\}, \pi, w)$.

Ejemplo 1.5. *Examinemos el soporte para $M(2, \{[0, 1/2], [1/2, 1]\}, (2, 1), (-1, 1))$.*

En este caso se determina, gracias a la partición y la permutación que sobre el primer rectángulo, el $[0, 1/2] \times [0, 1]$ se pone el trozo de segmento que estaba originalmente en el rectángulo $[1/2, 1] \times [0, 1]$ y como en este caso $w([0, 1/2]) = -1$ se cambia su pendiente. Del mismo modo en $[1/2, 1] \times [0, 1]$ yace ahora lo que antes estuvo en $[0, 1/2] \times [0, 1]$ y conserva su pendiente ($w([1/2, 1]) = 1$). Es así que el soporte queda determinado por la función

$$S(x) = \begin{cases} 1 - x, & \text{si } x \in [0, 1/2], \\ x - 1/2, & \text{si } x \in [1/2, 1]. \end{cases}$$

■

Esto significa que si $(X, Y) \sim M(2, \{[0, 1/2], [1/2, 1]\}, (2, 1), (-1, 1))$, entonces $Y \stackrel{c.s.}{=} S(X)$ con lo que las variables X, Y están determinísticamente relacionadas.

Una propiedad importante de la familia de reordenamientos de M es que podemos aproximar uniformemente la cópula Π con ella. Formalmente hablando, tenemos el siguiente teorema.

Teorema 1.8. *Sea $\varepsilon > 0$. Entonces existe un reordenamiento de M que denotaremos M_ε tal que*

$$\|\Pi - M_\varepsilon\| < \varepsilon$$

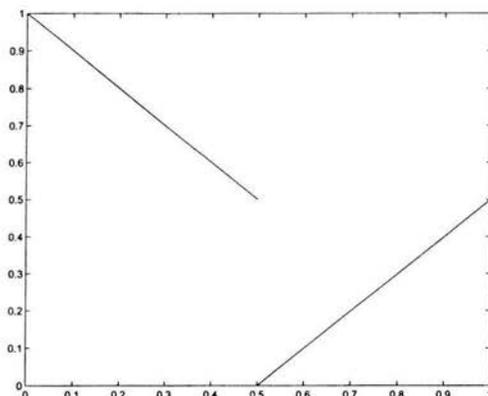


Figura 1.2: Soporte del reordenamiento de M

Demostración. Para efectuar la prueba generamos una partición del intervalo unitario y usándola repartimos la masa de M de modo uniforme sobre los rectángulos que ella determina.

Sea $\varepsilon > 0$ y para ella, $m \in \mathbb{N}$ tal que $m \geq 4/\varepsilon$. Entonces, por el lema 2, para cualquier cópula

$$|C(u, v) - C(u', v')| < \varepsilon/2 \quad , \text{ si, } |u - u'| < 1/m, |v - v'| < 1/m$$

Consideremos $n = m^2$ y denotemos por I_m e I_n las particiones regulares de orden m y n respectivamente del intervalo $[0, 1]$. El soporte de M , la identidad, pasa por todos los cuadrados de la forma $[i/r, (i+1)/r]^2$, $r = m, n$ dejando en ellos una masa de $1/r$. La permutación de los componentes de I_n que usaremos es aquella que reparta la masa de probabilidad del cuadrado $[i/m, i+1/m]^2$ —a saber, $1/m$ — en el rectángulo $[0, 1] \times [i/m, i+1/m]$ logrando que en cada uno de los cuadrados que se determinan en I_m^2 haya probabilidad $1/n$.

La permutación $\pi[m(j-1) + k] = m(k-1) + j$ para $j, k = 1, 2, \dots, m$ logra el efecto deseado. Para terminar de determinar un reordenamiento de M digamos que $w \equiv 1$ y nombremos a $M(n, I_n, \pi, w) = M_\varepsilon$. Observemos que si $p, q \in \{1, 2, \dots, m\}$ entonces $M_\varepsilon(p/m, q/m) = \Pi(p/m, q/m)$. Esto se

debe a que

$$V_{M_\varepsilon}([0, p/m] \times [0, q/m]) = \sum_{\substack{1 \leq i \leq p \\ 1 \leq j \leq q}} V_{M_\varepsilon}([i/n, i+1/n] \times [j/n, j+1/n]) = pq/n$$

Con lo cual

$$\begin{aligned} M_\varepsilon(p/m, q/m) &= V_{M_\varepsilon}([0, p/m] \times [0, q/m]) \\ &= V_\Pi([0, p/m] \times [0, q/m]) = \Pi(p/m, q/m) \end{aligned}$$

Siendo así y dado $(u, v) \in I^2$ tomemos p, q tales que $|u - p/m| < 1/m$ y $|v - q/m| < 1/m$. Entonces

$$\begin{aligned} |C(u, v) - M_\varepsilon(u, v)| &\leq |C(u, v) - C(p/m, q/m)| + |C(p/m, q/m) \\ &\quad - M_\varepsilon(p/m, q/m)| + |M_\varepsilon(p/m, q/m) - M_\varepsilon(u, v)| \\ &\leq \varepsilon/2 + \varepsilon/2 = \varepsilon \end{aligned}$$

■

Esto no es todo. Si tomamos ahora en vez de Π una cópula cualquiera, podemos hacer un reordenamiento de M que le aproxime uniformemente. La prueba de este hecho utiliza la misma permutación pero con una partición distinta. Esto significa que los reordenamientos de M contienen un conjunto que es denso en \mathcal{C} , el conjunto de las cópulas bidimensionales. Volveremos a hablar de esta propiedad más adelante en términos estadísticos.

1.7 Modos de dependencia

Hemos examinado ciertas propiedades de la dependencia que se resumen en la cópula. Los resultados presentados sobre las cotas de Fréchet tienen cierta debilidad en el sentido que nos explican una relación *determinista* y de tipo *funcional* entre las variables en cuestión. ¿Qué pasa cuando la relación es más sofisticada, menos explícita? ¿Podemos “leer” en la cópula este hecho?

Veremos ahora algunas nociones de dependencia y su relación con las cópulas.

Definición 4. Sea X una v.a. con distribución F . Entonces X es simétrica alrededor de $\theta \in \mathfrak{R}$ si y sólo si

$$F(\theta + x) = 1 - F(\theta - x) \quad \forall x \in \mathfrak{R}$$

Los ejemplos de este tipo de variables son muchos, comenzando por las v.a. con distribución normal que tienen simetría alrededor de la media. Para extender este concepto a dos dimensiones una alternativa es la siguiente.

Definición 5. Sea (X, Y) un vector aleatorio, entonces (X, Y) es simétrico alrededor del punto (θ, γ) si y sólo si

$$(X - \theta, Y - \gamma) \stackrel{D}{=} (\theta - X, \gamma - Y)$$

Esta definición es equivalente a que la distribución conjunta del vector (X, Y) satisfaga, para cada $(x, y) \in \mathfrak{R}^2$

$$F_{X,Y}(x + \theta, y + \gamma) = 1 - F_X(\theta - x) - F_Y(\gamma - y) + F_{X,Y}(\theta - x, \gamma - y) \quad (1.30)$$

Debido a que si $(x, y) \in \mathfrak{R}^2$ entonces

$$\begin{aligned} F_{X,Y}(x + \theta, y + \gamma) &= \mathbb{P}(X - \theta \leq x, Y - \gamma \leq y) \\ &= \mathbb{P}(\theta - X \leq x, \gamma - Y \leq y) \\ &= \mathbb{P}(X \geq \theta - x, Y \geq \gamma - y) \\ &= 1 - \mathbb{P}(X \leq \theta - x) - \mathbb{P}(Y \leq \gamma - y) + \mathbb{P}(X \leq \theta - x, Y \leq \gamma - y) \\ &= 1 - F_X(\theta - x) - F_Y(\gamma - y) + F_{X,Y}(\theta - x, \gamma - y) \end{aligned}$$

Esta definición es una extensión de la simetría en una variable en el sentido que si antes los dos rayos $(-\infty, -\theta]$ y $[\theta, \infty)$ tenían la misma probabilidad para cada $x \in \mathfrak{R}$, ahora los semiplanos

$$\begin{aligned} \{ (u, v) \quad t.q. \quad u \geq \theta - x, v \geq \gamma - y \} \\ \{ (u, v) \quad t.q. \quad u \leq \theta + x, v \leq \gamma + y \} \end{aligned} \quad (1.31)$$

tienen la misma probabilidad.

Siempre que las v.a. sean simétricas, podemos expresar la simetría del vector (X, Y) en términos sólo de su cópula utilizando la relación (1.30) como sigue.

Proposición 4. Sean X, Y variables aleatorias no degeneradas con cópula C . Supóngase que ambas son simétricas alrededor θ y γ respectivamente. Entonces (X, Y) es simétrico alrededor de (θ, γ) si y sólo si para todo $u, v \in I$

$$C(u, v) = u + v - 1 + C(1 - u, 1 - v)$$

Demostración. Sean F_X, F_Y las distribuciones de X y de Y , y sea $F_{X,Y}$ su distribución conjunta. Entonces, para $(x, y) \in \mathfrak{R}^2$

$$\begin{aligned} F_{X,Y}(x + \theta, y + \gamma) &= 1 - F_X(\theta - x) - F_Y(\gamma - y) + F_{X,Y}(\theta - x, \gamma - y) \\ \Leftrightarrow C(F_X(x + \theta), F_Y(y + \gamma)) &= 1 - F_X(\theta - x) - F_Y(\gamma - y) + C(F_X(\theta - x), F_Y(\gamma - y)) \end{aligned}$$

Utilizando ahora la simetría de X y Y en esta expresión se concluye que

$$\begin{aligned} C(F_X(x + \theta), F_Y(y + \gamma)) &= F_X(\theta + x) - (1 - F_Y(\gamma + y)) \\ &\quad + C(1 - F_X(\theta + x), 1 - F_Y(\gamma + y)) \end{aligned}$$

Con lo cual queda probado que para cada $(u, v) \in [0, 1]^2$

$$C(u, v) = u + v - 1 + C(1 - u, 1 - v)$$

■

Observación 4. En este desarrollo observamos que la simetría de la cópula se hereda hacia la distribución conjunta si las marginales también son simétricas. Sin embargo, si las marginales no son simétricas, la conjunta no necesariamente lo es.

Como otro ejemplo en que la cópula resume un modo de dependencia, consideremos dos variables aleatorias (X, Y) y veamos que la cópula de las variables $(\min(X, Y), \max(X, Y))$ es independiente de las distribuciones de X y de Y .

Ejemplo 1.6. Sean U_1 , y U_2 variables aleatorias independientes uniformemente distribuidas sobre el intervalo unitario. Considérense $Y_1 = \max(U_1, U_2)$ y $Y_2 = \min(U_1, U_2)$ sus estadísticos de orden. Entonces, la cópula $C = C_{Y_1, Y_2}$ es

$$C(u, v) = \begin{cases} 2(1 - \sqrt{1 - u}) \cdot (\sqrt{v}) & \text{si } (1 - u) - \sqrt{1 - u} \leq v/2 \\ v & \text{en otro caso} \end{cases} \quad (1.32)$$

Demostración. La densidad conjunta para estadísticos de orden está dada por

$$f_{Y_1, Y_2}(y_1, y_2) = 2! f(y_1) \cdot f(y_2), \quad \text{si } y_2 \geq y_1,$$

con lo cual, la distribución conjunta de Y_1, Y_2 en caso de v.a. uniformes es,

$$\begin{aligned} F_{Y_1, Y_2}(y_1, y_2) &= \mathbb{P}(Y_1 \leq y_1, Y_2 \leq y_2) = 2 \int_0^{y_1} \int_0^{y_2} \chi_{(0,1)}(s) \chi_{(s,1)}(t) dt ds \\ &= 2 \int_0^{y_1} \chi_{(0,1)}(s) \int_s^{y_2} dt ds = 2 \int_0^{y_1} (y_2 - s) ds \\ &= 2(y_2 y_1 - y_1^2/2) \end{aligned}$$

Es decir

$$F_{Y_1, Y_2}(y_1, y_2) = \begin{cases} 2y_2 y_1 - y_1^2, & \text{si } 0 \leq y_2 \leq y_1 \leq 1, \\ y_2^2, & \text{si } y_2 \leq y_1 \leq 1 \text{ o } y_2 \leq 1 \leq y_1, \\ 2y_1 - y_1^2, & \text{si } y_2 \geq 1, 0 \leq y_1 \leq 1, \\ 1, & \text{si } y_1, y_2 \geq 1 \end{cases}$$

Dado que las distribuciones de Y_1 y Y_2 son

$$F_{Y_1}(y_1) = 2y_1 - y_1^2 \quad F_{Y_2}(y_2) = [F(y_2)]^2$$

podemos utilizar sus inversas

$$F_{Y_1}^{-1}(u) = 1 - \sqrt{1-u} \quad F_{Y_2}^{-1}(v) = \sqrt{v}$$

para encontrar que la cópula $C(u, v) = F_{Y_1, Y_2}(F_{Y_1}^{-1}(u), F_{Y_2}^{-1}(v))$ está dada por la expresión

$$C(u, v) = \begin{cases} 2(1 - \sqrt{1-u})(\sqrt{v}) - v, & \text{si } 1 - \sqrt{1-u} \leq \sqrt{v}, \\ v, & \text{en otro caso.} \end{cases}$$

■

Esta demostración puede realizarse aún cuando las distribuciones marginales no sean uniformes. Esto nos dice que la relación entre (X, Y) de ser el máximo y mínimo de v.a. independientes idénticamente distribuidas no depende más que de la cópula.

Ejemplo 1.7. Si X, Y son v.a. con distribución F continua, entonces la cópula de sus estadísticos de orden está dada por (1.32).

Demostración. Como en el ejemplo anterior, tenemos que

$$F_{Y_1, Y_2}(y_1, y_2) = 2F(y_2)F(y_1) - [F(y_1)]^2$$

y teniendo en cuenta la distribución para Y_1 y Y_2

$$F_{Y_1}^{-1}(u) = F^{-1}(1 - \sqrt{1-u}) \quad F_{Y_2}(v) = F^{-1}(\sqrt{v})$$

con lo cuál se concluye que la cópula es

$$C(u, v) = \begin{cases} 2(1 - \sqrt{1-u})(\sqrt{v}) - v, & \text{si } F^{-1}(1 - \sqrt{1-u}) \leq F^{-1}(\sqrt{v}), \\ v, & \text{en otro caso.} \end{cases}$$

Siendo la función F^{-1} creciente (inversa de creciente), se concluye que $F^{-1}(1 - \sqrt{1-u}) \leq F^{-1}(\sqrt{v})$ si y sólo si $1 - \sqrt{1-u} \leq \sqrt{v}$ de donde el resultado es cierto. ■

Los pasados ejemplos muestran que hay relaciones que no son deterministas ni funcionales entre dos variables que se resumen en su cópula. Este no es el único caso, también es cierto que hay relaciones que dependen de las funciones de distribución marginales como muestra el siguiente ejemplo.

Ejemplo 1.8. Sean $X \sim N(0, 1)$, $Y \sim N(\mu, \sigma^2)$. Entonces la cópula correspondiente al vector $(X, X + Y)$ depende de σ^2 .

Demostración. Por el teorema de cambio de variable la distribución conjunta de $(X, X + Y)$ es normal $N\left(\begin{pmatrix} 0 \\ \mu \end{pmatrix}, \begin{pmatrix} 1 & \\ & \sigma^2 \end{pmatrix}\right)$ y coeficiente de correlación $1/(1 + \sigma^2)$. Esto significa que la cópula de $(X, X + Y)$ es la cópula Gaussiana con parámetros $(0, \mu, 1, \sigma \cdot 1/(1 + \sigma^2))$, una función de σ^2 .

De hecho, gracias a que el coeficiente de correlación es $1/(1 + \sigma^2)$ podemos observar que conforme $\sigma \rightarrow \infty$ la distribución conjunta tiende a la de v.a. independientes, es decir $C_\sigma \rightarrow \Pi$. Por otra parte, si $\sigma \rightarrow 0$ entonces la distribución tiende a la del vector $(X, X + \mu)$, de modo que $C_\sigma \rightarrow M$. ■

Otros modos de dependencia que están completamente resumidos en la cópula se tratan en las siguientes páginas.

Definición 6. Sean X, Y variables aleatorias. Diremos que X y Y son positivamente dependientes por cuadrante, $PQD(X, Y)$ (del inglés “positively quadrant dependent”), cuando

$$\mathbb{P}(X \leq x, Y \leq y) \geq \mathbb{P}(X \leq x) \cdot \mathbb{P}(Y \leq y) \quad \forall x, y \in \mathfrak{R} \quad (1.33)$$

Equivalentemente, si $\mathbb{P}(X \leq x) > 0$

$$\mathbb{P}(Y \leq y | X \leq x) \geq \mathbb{P}(Y \leq y) \quad \forall x, y \in \mathfrak{R} \quad (1.34)$$

Observación 5. Esta definición puede postularse con las funciones de supervivencia $\mathbb{P}(X > x, Y > y)$, $\mathbb{P}(X > x)$, $\mathbb{P}(Y > y)$.

Para ello, observemos que

$$\begin{aligned} \mathbb{P}(X > x, Y > y) &\geq \mathbb{P}(X > x)\mathbb{P}(Y > y) \quad \text{si y sólo si} \\ 1 - \mathbb{P}(X \leq x) - \mathbb{P}(Y \leq y) + \mathbb{P}(X \leq x, Y \leq y) &\geq \\ &= (1 - \mathbb{P}(X \leq x))(1 - \mathbb{P}(Y \leq y)) \end{aligned} \quad (1.35)$$

lo que, cancelando términos, equivale con que $PQD(X, Y)$.

Para entender el sentido de esta definición más a fondo, agreguemos algunos términos a la relación (1.33) poniendo

$$\begin{aligned} F_{XY}(x, y) &\geq F_X(x)G_Y(y) \Leftrightarrow \\ F_{XY}(x, y) - F_X(x)F_{XY}(x, y) - G_Y(y)F_{XY}(x, y) + (F_{XY}(x, y))^2 &\geq \\ F_X(x)G_Y(y) - F_X(x)F_{XY}(x, y) - G_Y(y)F_{XY}(x, y) + (F_{XY}(x, y))^2 & \end{aligned}$$

Factorizando $F_{XY}(x, y)$ en el término de la izquierda y reduciendo el término de la derecha obtenemos que $PQD(X, Y)$ si y sólo si para cualquier punto $(x, y) \in \mathfrak{R}^2$

$$\begin{aligned} F_{XY}(x, y)(1 - F_X(x) - G_Y(y) + H(x, y)) &\geq \\ (F_X(x) - H(x, y))(G_Y(y) - H(x, y)) & \end{aligned} \quad (1.36)$$

Equivalentemente,

$$\mathbb{P}(X \leq x, Y \leq y)\mathbb{P}(X > x, Y > y) \geq \mathbb{P}(X \leq x, Y > y)\mathbb{P}(X > x, Y \leq y)$$

lo cual indica que los valores que toma la variable se concentran en los semiplanos $\{(a, b) \text{ t.q. } a \leq x, b \leq y\}$ $\{(a, b) \text{ t.q. } a > x, b > y\}$.

Podemos entender la dependencia por cuadrante en términos de la covarianza entre las variables X, Y . Para ello, es de utilidad calcular la covarianza como función de las distribuciones conjunta y marginales según el siguiente lema.

Lema 10 (Hoeffding (1940)). Sean X, Y variables aleatorias continuas con función de distribución conjunta F_{XY} de marginales F_X, F_Y respectivamente. Entonces

$$\text{Cov}(X, Y) = \iint_{\mathbb{R}^2} (F_{XY}(x, y) - F_X(x)F_Y(y)) dx dy$$

Con este resultado probaremos que la dependencia por cuadrante es equivalente con que el coeficiente de correlación de Pearson sea no-negativo para cierta familia de funciones aplicadas a X, Y como afirma el siguiente teorema.

Teorema 1.9 (Lehman 1966). Sean X, Y variables aleatorias continuas con distribución conjunta F_{XY} de marginales F_X, F_Y respectivamente. Entonces $PQD(X, Y)$ si y sólo si $\text{Cov}(f(X), g(Y)) \geq 0$ para toda pareja de funciones f, g crecientes para las cuales $\mathbb{E}(f(X)), \mathbb{E}(g(Y)), \mathbb{E}(f(X)g(Y))$ existen y sean finitas.

Demostración. Supongamos primero que $PQD(X, Y)$. Por el lema de Hoeffding, tenemos que

$$\text{Cov}(X, Y) = \iint_{\mathbb{R}^2} (F_{XY}(x, y) - F_X(x)F_Y(y)) dx dy \geq 0$$

Además

$$\begin{aligned} \mathbb{P}(f(X) \leq x, g(Y) \leq y) &= \mathbb{P}(X \leq f^{-1}(x), Y \leq g^{-1}(y)) \\ &\geq \mathbb{P}(X \leq f^{-1}(x))\mathbb{P}(Y \leq g^{-1}(y)) = \mathbb{P}(f(X) \leq x)\mathbb{P}(g(Y) \leq y) \end{aligned}$$

de modo que $PQD(f(X), g(Y))$ y así $\text{Cov}(f(X), g(Y)) \geq 0$.

Por otro lado, si para toda pareja de funciones creciente para las cuales $\mathbb{E}(f(X)), \mathbb{E}(g(Y)), \mathbb{E}(f(X)g(Y))$ existan y sean finitas tenemos que $\text{Cov}(f(X), g(Y)) \geq 0$ entonces en particular para $f_x(X) = \mathbb{1}(\cdot \leq x)$, $g_y(Y) = \mathbb{1}(\cdot \leq y)$ de modo que

$$0 \leq \text{Cov}(f_x(X), g_y(Y)) = F_{XY}(x, y) - F_X(x)F_Y(y) \quad \text{para todo } x, y \in \mathfrak{R}$$

con lo cuál $PQD(X, Y)$. ■

Este tipo de dependencia se puede resumir como propiedad de la cópula partiendo de la definición misma y usando el teorema de Sklar como sigue:

$$F_{X,Y}(x, y) \geq F_X(x) \cdot F_Y(y) \Leftrightarrow C(F_X(x), F_Y(y)) \geq F_X(x) \cdot F_Y(y)$$

para toda pareja $(x, y) \in \mathfrak{R}^2$. Es decir

Teorema 1.10. Sean X y Y variables aleatorias con cópula C . Entonces $PQD(X, Y)$ si y sólo si

$$C(u, v) \geq u \cdot v \quad \forall u, v \in [0, 1]$$

■

Ahora bien, utilizando la ecuación (1.36) podemos afirmar que

Teorema 1.11. Dada una pareja de variables aleatorias con cópula C , $PQD(X, Y)$ si y sólo si para cada $u, v \in [0, 1]$

$$C(u, v)(1 - u - v + C(u, v)) \geq (u - C(u, v))(v - C(u, v))$$

■

Dadas X, Y , variables aleatorias, son negativamente dependientes por cuadrante, $NQD(X, Y)$ (del inglés “negatively quadrant dependent”) si y sólo si para cada punto $(x, y) \in \mathfrak{R}^2$: $F_{XY}(x, y) \leq F_X(x)F_Y(y)$. Para esta relación tendremos los análogos a los teoremas 1.10 y 1.11.

Una relación de dependencia similar a la dependencia por cuadrante es la monotonía de colas. Recordemos que la dependencia por cuadrante se definió cuando $\mathbb{P}(X \leq x) > 0$ según (1.34) pidiendo que

$$\mathbb{P}(Y \leq y | X \leq x) \geq \mathbb{P}(Y \leq y) \quad \forall (x, y) \in \mathfrak{R}^2$$

Esto podemos verlo como

$$\mathbb{P}(Y \leq y | X \leq x) \geq \mathbb{P}(Y \leq y | X \leq \infty) \quad \forall (x, y) \in \mathfrak{R}^2$$

Una condición un tanto más fuerte es pedir que la función

$$I_y(x) = \mathbb{P}(Y \leq y | X \leq x) \quad \text{sea monótona para cada } y \in \mathfrak{R}.$$

Definición 7. Sean X, Y variables aleatorias y para cada toda pareja $x, y \in \mathfrak{R}$ definamos

$$I_y(x) = \mathbb{P}(Y \leq y | X \leq x) \quad H_y(x) = \mathbb{P}(Y > y | X > x)$$

Entonces

1. Y es de cola izquierda decreciente en X cuando $I_y(x)$ es decreciente para todo y . Denotamos este hecho como $LTD(Y|X)$ (por su nombre en inglés “left tail decreasing”).
2. Y es de cola derecha creciente en X cuando $H_y(x)$ es creciente para todo y . Denotamos este hecho como $RTI(Y|X)$ (por su nombre en inglés “right tail increasing”).

Para intuir un poco el sentido de esta definición, pensemos que X, Y son los tiempos de vida de dos componentes c_1, c_2 respectivamente. Entonces $RTI(Y|X)$ significa que conforme el tiempo de vida del componente c_1 avanza, aumenta la probabilidad de que c_2 tenga una vida larga. Análogamente, $LTD(Y|X)$ nos dice que la probabilidad de que el tiempo de vida de Y sea corto decrece conforme aumenta el tiempo de vida de X .

Estas dos relaciones implican la dependencia por cuadrante positiva. Por ejemplo, si $RTI(Y|X)$ entonces

$$\mathbb{P}(Y > y|X > x) \geq P(Y > y|X > -\infty) = \mathbb{P}(Y > y)$$

de donde $\mathbb{P}(Y > y, X > x) \geq \mathbb{P}(Y > y)\mathbb{P}(X > x)$, es decir, $PQD(X, Y)$ según la ecuación (1.35). Del mismo modo, si $LTD(Y|X)$ entonces $\mathbb{P}(Y \leq y|X \leq x) \geq \mathbb{P}(Y \leq y, X \leq \infty) = \mathbb{P}(Y \leq y)$, es decir, $PQD(X, Y)$ por definición.

Con el siguiente ejemplo demostraremos que la monotonía de colas y la dependencia por cuadrante no son equivalentes.

Ejemplo 1.9. Sean U, V variables aleatorias uniformemente distribuidas en $(0, 1)$ con distribución conjunta $C(u, v) = \min(u, v, (u^2 + v^2)/2)$. Entonces $PQD(U, V)$ pero no hay monotonía de colas.

Demostración. Gracias a que $u \geq uv, v \geq uv$ y también $(u^2 + v^2)/2 \geq uv$ se tiene que en todo punto $C(u, v) \geq \Pi(u, v)$, de modo que $PQD(U, V)$. Ahora bien

$$\begin{aligned} \mathbb{P}(U \leq 1/2|V \leq 1/2) &= \mathbb{P}(U \leq 1/2, V \leq 1/2)/\mathbb{P}(V \leq 1/2) \\ &= 2C(1/2, 1/2) = 2(\min(1/2, 1/2, 1/4)) = 1/2 \end{aligned}$$

Análogamente

$$\begin{aligned} \mathbb{P}(U \leq 1/2|V \leq \sqrt{3}/2) &= (2/\sqrt{3})C(1/2, \sqrt{3}/2) \\ &= (2/\sqrt{3})(1/2) = \sqrt{3}/3 \approx .577 \end{aligned}$$

De modo que U no es de cola izquierda decreciente en V . Además

$$\begin{aligned}\mathbb{P}(U > 1/2, |V > 1 - \sqrt{(3)/2}) &= (1 - 1/2 - (1 - \sqrt{(3)/2}) - \\ &\quad C(1/2, 1 - \sqrt{(3)/2})) / (1 - \sqrt{(3)/2}) \\ &= \sqrt{3}/3 \approx .577\end{aligned}$$

Análogamente

$$\mathbb{P}(U > 1/2 | V > 1/2) = C(1/2, 1/2) / (1/2) = 1/2$$

De modo que U no es de cola derecha creciente en V . Por simetría de C se tiene que V no es de cola izquierda decreciente en U y que V no es de cola derecha creciente en U . ■

La monotonía de colas también está resumida en la cópula como se demuestra en el siguiente teorema .

Teorema 1.12. Sean X, Y variables aleatorias continuas con cópula C . Entonces

1. $LTD(Y|X)$ si y sólo si para todo $v \in [0, 1]$, $C(u, v)/u$ es decreciente como función de u ,
2. $RTI(Y|X)$ si y sólo si para todo $v \in [0, 1]$, $1 - u - v + C(u, v)/(1 - u)$ es creciente en u .

Demostración. Para cada $x \in \mathfrak{R}$, por definición de $I_y(x)$ y el teorema de Sklar tenemos

$$I_y(x) = \frac{F_{XY}(x, y)}{F_X(x)} = \frac{C(F_X(x), F_Y(y))}{F_X(x)}$$

Ahora bien, $LTD(Y|X)$ si y sólo si $I_y(x)$ es decreciente en x . Haciendo el cambio de variable $u = F_X(x), v = F_Y(y)$ y siendo F_X, F_Y funciones crecientes y continuas, se concluye que $I(x)$ es decreciente en x si y sólo si $C(u, v)/u$ lo es en u .

Análogamente

$$\begin{aligned}H_y(x) &= \frac{1 - F_X(x) - F_Y(y) + F_{XY}(x, y)}{1 - F_X(x)} \\ &= \frac{1 - u - v + C(u, v)}{1 - u}\end{aligned}$$

De modo que $H_y(x)$ es creciente si y sólo si $1 - u - v + C(u, v)/(1 - u)$ lo es. ■

Otro aspecto importante en la relación que guardan dos variables aleatorias es la concordancia.

Dada una muestra aleatoria bivariada $(X_1, Y_1), \dots, (X_n, Y_n)$ diremos que las parejas $(X_1, Y_1), (X_2, Y_2)$ son concordantes cuando $(X_1 - X_2)(Y_1 - Y_2) > 0$ y discordantes si $(X_1 - X_2)(Y_1 - Y_2) < 0$. La τ de Kendall, utilizada frecuentemente en estadística para pruebas de hipótesis, está definida para la muestra como la diferencia entre la probabilidad de concordancia y la de discordancia para un elemento elegido aleatoriamente de la muestra, es decir

$$\tau_n = \frac{\#\text{pares concordantes} - \#\text{pares discordantes}}{\binom{n}{2}}$$

Si consideramos ahora dos vectores $(X_1, Y_1), (X_2, Y_2)$ independientes e idénticamente distribuidos, la versión poblacional de la τ es

$$\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0] \quad (1.37)$$

Para demostrar el papel que tienen las cópulas en este modo de medir la concordancia supongamos que los vectores $(X_1, Y_1), (X_2, Y_2)$ son independientes con funciones de distribución marginales idénticas y funciones de distribución conjunta posiblemente distintas y definamos su función de concordancia y discordancia como

$$Q = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0].$$

Demostremos que Q es una función de las cópulas de (X_1, Y_1) y de (X_2, Y_2) .

Teorema 1.13. *Si $(X_1, Y_1), (X_2, Y_2)$ son vectores aleatorios continuos independientes con funciones de distribución marginales idénticas (F para X_1, X_2 y G para Y_1, Y_2) y cópulas C_1, C_2 respectivamente entonces*

$$Q = Q(C_1, C_2) = 4 \iint_{[0,1]^2} C_2(u, v) dC_1(u, v) - 1 \quad (1.38)$$

Demostración. Como los vectores son continuos,

$$\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] + \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0] = 1$$

de modo que

$$Q = 2\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - 1 \quad (1.39)$$

Por otra parte

$$\begin{aligned} \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] &= \mathbb{P}(X_1 > X_2, Y_1 > Y_2) \\ &\quad + \mathbb{P}(X_1 < X_2, Y_1 < Y_2) \end{aligned} \quad (1.40)$$

y esto puede ser evaluado con las distribuciones conjuntas como sigue

$$\begin{aligned} \mathbb{P}(X_1 > X_2, Y_1 > Y_2) &= \mathbb{P}(X_2 \leq X_1, Y_2 \leq Y_1) \\ &= \iint_{\mathbb{R}^2} \mathbb{P}(X_2 \leq X_1, Y_2 \leq Y_1 | X_1 = x, Y_1 = y) dC_1(F(x), G(y)) \\ &= \iint_{\mathbb{R}^2} \mathbb{P}(X_2 \leq x, Y_2 \leq y) dC_1(F(x), G(y)) \\ &= \iint_{\mathbb{R}^2} C_2(F(x), G(y)) dC_1(F(x), G(y)) \end{aligned}$$

De modo que empleando las transformaciones $u = F(x), v = G(y)$ obtenemos

$$\mathbb{P}(X_1 > X_2, Y_1 > Y_2) = \iint_{[0,1]^2} C_2(u, v) dC_1(u, v)$$

Análogamente

$$\begin{aligned} \mathbb{P}(X_1 < X_2, Y_1 < Y_2) &= \iint_{\mathbb{R}^2} \mathbb{P}(X_2 \geq x, Y_2 \geq y) dC_1(F(x), G(y)) \\ &= \iint_{\mathbb{R}^2} (1 - F(x) - G(y) + C_2(F(x), G(y))) dC_1(F(x), G(y)) \\ &= \iint_{[0,1]^2} (1 - u - v + C_2(u, v)) dC_1(u, v) \end{aligned}$$

Pero siendo C_1 la distribución de variables aleatorias uniformes (U, V) en $[0, 1]$ tenemos que $\mathbb{E}(U) = \mathbb{E}(V) = 1/2$, y así

$$\begin{aligned} \mathbb{P}(X_1 < X_2, Y_1 < Y_2) &= 1 - 1/2 - 1/2 + \iint_{[0,1]^2} C_2(u, v) dC_1(u, v) \\ &= \iint_{[0,1]^2} C_2(u, v) dC_1(u, v) \end{aligned}$$

Por ello, usando (1.40),

$$\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] = 2 \iint_{[0,1]^2} C_2(u, v) dC_1(u, v)$$

Sustituyendo esto en (1.39) obtenemos (1.38). ■

Corolario 5. Si C_1, C_2 y Q están definidas como en el Teorema 1.13 entonces

1. Q es simétrica en sus argumentos, i.e., $Q(C_2, C_1) = Q(C_1, C_2)$,
2. Q es creciente en cada uno de sus argumentos, i.e. si $C_1 \leq D_1$ y $C_2 \leq D_2$, entonces $Q(C_1, C_2) \leq Q(D_1, C_2) \leq Q(D_1, D_2)$.

Por la definición (1.37) tenemos que si C es la cópula de los vectores $(X_1, Y_1), (X_2, Y_2)$ entonces

$$\tau = Q(C, C) = 4 \iint_{[0,1]^2} C(u, v) dC(u, v) - 1 \quad (1.41)$$

Observando que C es la distribución conjunta de las variables $U = F(X_1), V = F(Y_1)$, podemos interpretar la integral en (1.41) como la esperanza de la variable $C(U, V)$, es decir

$$\tau = 4\mathbb{E}(C(U, V)) - 1 \quad (1.42)$$

De este modo, la τ de Kendall es una función exclusivamente de la cópula. Otra medida de concordancia usual es la ρ de Spearman. En este caso se define con tres vectores $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)$ aleatorios idénticamente distribuidos como

$$\rho = 3(\mathbb{P}[(X_1 - X_2)(Y_1 - Y_3) > 0]) - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_3) < 0.]$$

Observemos que en este caso la pareja (X_2, Y_3) tiene cópula Π por tener coordenadas independientes. De acuerdo con la demostración del Teorema 1.13 podemos expresar la ρ como función de la cópula de (X_1, Y_1) como sigue

Teorema 1.14. Se X, Y son variables aleatorias continuas con cópula C , entonces la ρ de Spearman para (X, Y) es

$$\rho = \rho(C) = 3Q(C, \Pi)$$



De este modo la cópula resume también dos de las medidas más usuales de concordancia entre las entradas de un vector aleatorio dado. Obsérvese que la concordancia es una medida que refleja el modo en que las observaciones de un vector aleatorio se reparten. Por ejemplo, el valor $\tau = 1$ indica que los valores grandes hacen pareja con los grandes y el valor $\tau = -1$ que los valores grandes tienden a agruparse con los pequeños.

1.8 Cópulas n -dimensionales

Del mismo modo que hemos definido una cópula en dos dimensiones como una distribución correspondiente a variables aleatorias uniformes y restringida al cuadrado unitario, diremos que una cópula en n dimensiones es una función de distribución cuyas marginales son uniformes y que está restringida a $[0, 1]^n$. Formalmente

Definición 8. Una n -Cópula es una función $C : I^n \rightarrow I$ que satisface que para cada $u = (u_1, u_2, \dots, u_n) \in I^n$,

1. $C(u) = 0$ si $\min_i(u_i) = 0$

2. $C(u) = u_j$ si $u_i = 1 \forall i \neq j$

3. $V_C(B) = \sum \text{sgn}(v)C(v) \geq 0 \quad \forall B, n\text{-caja},$

La suma tomada sobre todos los vértices de la caja $B = \otimes_k [a_k, b_k]$ y $\text{sgn}(\cdot)$ está definida como

$$\text{sgn}(b) = \begin{cases} 1 & \text{si } v_k = a_k \text{ para una cantidad par de } k\text{'s} \\ -1 & \text{si } v_k = a_k \text{ para una cantidad impar de } k\text{'s} \end{cases}$$

A las funciones $C_k(u_k) = C(1, 1, \dots, u_k, 1, \dots, 1)$ se les llama marginales unidimensionales de C . A las funciones $C_{i_1, i_2, \dots, i_j}(u_{i_1}, \dots, u_{i_j})$ que corresponden con las evaluaciones de C en que todo valor es 1 salvo los de las entradas i_1, i_2, \dots, i_j se les llama marginales j -dimensionales de C .

Las cópulas en más de dos dimensiones comparten muchas de las propiedades y las pruebas correspondientes son similares. Cabe destacar que se tiene la versión n -dimensional del teorema de Sklar

Teorema 1.15. *Si H es una función de distribución en n dimensiones con marginales F_1, F_2, \dots, F_n entonces existe una n -cópula tal que para cada $x \in \mathfrak{R}^n$*

$$H(x_1, x_2, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n)) \quad (1.43)$$

Si todas las marginales son continuas, entonces C es única, en otro caso está determinada de modo único en $\bigotimes_{i=1}^n \text{Ran}F_i$. Recíprocamente, la función definida por (1.43) es una función de distribución en n dimensiones cuyas marginales son F_1, F_2, \dots, F_n .

Del mismo modo, se tienen las cotas de Fréchet–Hoeffding, aunque en este caso no se puede afirmar que sean cópulas.

Teorema 1.16. *Si C es una n -cópula, entonces, para cada $x \in \mathfrak{R}^n$*

$$W^n(x) \leq C(x) \leq M^n(x)$$

donde $W^n(x) = \max(\sum_i x_i - n + 1, 0)$ y $M^n(x) = \min_i(x_i)$.

Proposición 5. *La función $W^n(x)$ no es una n -cópula.*

Demostración. Basta tomar la caja $[1/2, 1]^n$ y notar que las evaluaciones en cada vértice $v = (v_1, \dots, v_n)$ son nulas salvo en el caso en que se tengan $n-1$ veces el valor 1 y una vez el valor $1/2$ o n veces el valor 1, ya que en los otros casos $\max(\sum_i v_i - n + 1, 0) = 0$. Con ello, habiendo n vértices en la primera condición y sólo 1 en la segunda, se concluye que

$$V_{W^n}([1/2, 1]^n) = -n[n - 1 + 1/2 - n + 1] + 1[n - n + 1] = 1 - n/2$$

lo cual es no-negativo sólo cuando $n \leq 2$. ■

A pesar de no ser una cópula, W^n es una buena cota en el sentido siguiente que para cada $u \in [0, 1]^n$ existe una cópula C_u que coincide con W^n en este punto como se prueba en (Nelsen, [1]). Por último, diremos que las cópulas n -dimensionales también reflejan aspectos de la dependencia de variables aleatorias, por ejemplo, tenemos el análogo al teorema 1.5

Teorema 1.17. *Sean X_1, \dots, X_n variables aleatorias continuas, entonces cada una de ellas es función casi seguramente creciente de las otras si y sólo si la n -cópula de X_1, \dots, X_n es M^n .* ■

1.9 Conclusiones

El Teorema de Sklar muestra que siempre hay una relación funcional entre las distribuciones marginales y la distribución conjunta de un vector aleatorio. Como hemos visto, hay infinidad de distribuciones que comparten marginales, es decir, que sólo difieren en su cópula. Con esto en mente, uno puede tener la impresión de que la cópula resume todos los aspectos de dependencia, es decir, que basta estudiar la cópula para conocer la estructura de dependencia en un vector aleatorio. Esta visión puede reforzarse mediante el cálculo de la densidad conjunta en términos de las marginales y la cópula (1.24).

A lo largo del capítulo desarrollamos varios ejemplos en que ciertos aspectos de la dependencia se resumen en la cópula y en el ejemplo 1.8 mostramos que hay otros aspectos de la dependencia que no podemos conocer con la cópula. Siendo observadores podemos decir que las relaciones que la cópula resume son fundamentalmente referentes a la “geometría” de la distribución conjunta. Por ejemplo, la caracterización (1.31) de la simetría conjunta nos indica tan sólo en qué región de \mathbb{R}^2 se acumula mayor probabilidad. De modo similar, el Teorema 1.10 explica la dependencia por cuadrante como el hecho de que la gráfica de la distribución conjunta esté siempre sobre la del producto de las marginales.

Naturalmente hay mucho más que examinar en términos de cópulas y dependencia. A lo largo de los siguientes capítulos trataremos nuevos conceptos que nos ayuden a entender mejor el papel que juegan las cópulas en el terreno de la dependencia de vectores aleatorios.

Capítulo 2

Cóputas Arquimedianas

2.1 Introducción

Las propiedades generales de las cópulas tratadas en el capítulo anterior nos permiten construir una gran variedad de cópulas partiendo de funciones de distribución. No es este el único método para construir cópulas. En este capítulo veremos que podemos usar una función de una sola variable para este propósito. Estas cópulas son las cópulas arquimedianas. Examinaremos tanto su construcción como sus propiedades.

Partimos en la sección 2.2 de un modelo con variables latentes para motivar la construcción de las cópulas arquimedianas. Más adelante caracterizamos la formación de cópulas con este método y desarrollamos algunos ejemplos.

En la sección 2.3 examinamos el modo en que una cópula arquimediana distribuye su masa de probabilidad estudiando sus curvas de nivel y dando algunos ejemplos. Después demostramos que la τ de Kendall tiene una expresión simple en caso que se parta de una cópula arquimediana y esto abre la puerta al análisis de varias composiciones de variables aleatorias de las que se desprenden nuevas formas de caracterizar a las cópulas arquimedianas. Demostramos en esta sección que podemos conocer una cópula arquimediana de modo único a partir de una cierta distribución univariada. El capítulo termina con una tabla que muestra algunas cópulas arquimedianas paramétricas.

Para este capítulo utilizamos el capítulo 4 de [1] y el artículo [2].

2.2 Construcción de Cópulas Arquimedianas

Consideremos variables aleatorias U, V con distribuciones F_U, F_V respectivamente. Supongamos que dada γ una variable aleatoria latente, las distribuciones condicionales de U y V son

$$F_U(u|\gamma) = (F_U(u))^\gamma \quad F_V(v|\gamma) = (F_V(v))^\gamma$$

Siendo $F_U(u|\gamma), F_V(v|\gamma)$ funciones de distribución, sabemos, por el lema 7, que la función

$$H(u, v) = \int_{\mathfrak{R}} C(F_U(u|\gamma), F_V(v|\gamma)) dF(\gamma)$$

es una cópula para toda C cópula. En particular si $C \equiv \Pi$, en cuyo caso tenemos que

$$H(u, v) = \int_{\mathfrak{R}} (F_U(u))^\gamma (F_V(v))^\gamma dF(\gamma)$$

es una cópula. Marshal y Olkin (1988) demostraron que podemos expresar a H como

$$H(u, v) = \phi(\phi^{-1}(F_U(u)) + \phi^{-1}(F_V(v)))$$

donde ϕ es la función transformada de Laplace para γ . Podemos ver esta expresión como

$$\phi^{-1}(H(u, v)) = \phi^{-1}(u) + \phi^{-1}(v)$$

Siguiendo esta idea, las cópulas arquimedianas surgen de la solución de una ecuación funcional del tipo $\phi(C(u, v)) = \phi(u) + \phi(v)$ para $C(u, v)$ en caso de que exista. Claramente la solución de ella debe tener la forma

$$C(u, v) = \phi^{-1}\left(\phi(u) + \phi(v)\right) \quad (2.1)$$

lo cual nos llama a encontrar condiciones sobre la función ϕ para que esto sea posible. ϕ será definida en $[0, 1] \rightarrow [0, \infty]$ de modo que (2.1) defina una función en $[0, 1]^2 \rightarrow [0, 1]$. Pidiendo que ϕ sea inyectiva, podemos definir

la función *pseudo-inversa* $\phi^{[-1]}$ apoyándonos en la biyectividad sobre la restricción al rango como sigue

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t), & \text{si } t \in \text{Ran}(\phi) \\ K, & \text{en otro caso.} \end{cases} \quad (2.2)$$

Caso en el cual es natural pensar que si ϕ es creciente podemos utilizar el valor $K = 1$ y si es decreciente el valor $K = 0$. Si ϕ es creciente entonces la función obtenida mediante (2.1) no es una cópula. Para demostrarlo, observemos que

$$\phi(C(u, 1)) = \phi(u) + \phi(1).$$

De modo que la condición $C(u, 1) = u$ para $u \in [0, 1]$ equivale con que $\phi(1) = 0$. Esto significa, por estar definida $\phi: [0, 1] \rightarrow [0, \infty)$ que $\phi \equiv 0$, es decir que $C(u, v) = 0$ para todo punto $(u, v) \in [0, 1]^2$, y esta función no es una cópula.

Ahora bien, si la función es decreciente, tenemos que

1. La función pseudo-inversa está descrita por la regla de correspondencia

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t), & \text{si } t \in [0, \phi(0)] \\ 0, & \text{si } t \in [\phi(0), \infty) \end{cases} \quad (2.3)$$

2. $\forall u \in [0, 1]$

$$\begin{aligned} C(u, 0) &= \phi^{[-1]}(\phi(u) + \phi(v)) = 0 \\ C(u, 1) &= \phi^{[-1]}(\phi(u) + \phi(1)) \end{aligned}$$

De modo que al pedir que $\phi(1) = 0$ se satisfacen las condiciones de frontera. Para ser una cópula la función C que define la ecuación (2.1) debe ser bicreciente. La siguiente proposición nos brinda un modo de verificar esta propiedad cuando ϕ es continua.

Proposición 6. Si $\phi: [0, 1] \rightarrow [0, \infty)$ es una función continua, decreciente y $\phi(1) = 0$, entonces $C(u, v) = \phi^{[-1]}(\phi(u) + \phi(v))$ es una función bicreciente si y sólo si para cada $u_1 \leq u_2$,

$$C(u_2, v) - C(u_1, v) \leq u_2 - u_1$$

Demostración. Si suponemos que C es bicreciente entonces

$$0 \leq C([u_1, u_2] \times [v, 1]) = C(u_1, v) + u_2 - C(u_2, v) - u_1$$

Para probar el reverso, tomemos $v_1, v_2 \in [0, 1] \cdot \exists \cdot v_1 \leq v_2$ y notemos que $C(0, v_2) = 0 \leq v_1 \leq v_2 = C(1, v_2)$. Y como $C(\cdot, v_2)$ es continua por la suposición sobre ϕ (que implica que $\phi^{[-1]}$ es continua), existe $t \in [0, 1]$ tal que $C(t, v_2) = v_1$ o, dicho de otro modo, $\phi(v_2) + \phi(t) = \phi(v_1)$.

Por lo tanto

$$\begin{aligned} C(u_2, v_1) - C(u_1, v_1) &= \phi^{[-1]}(\phi(u_2) + \phi(v_1)) - \phi^{[-1]}(\phi(u_1) + \phi(v_1)) \\ &= \phi^{[-1]}(\phi(u_2) + \phi(v_2) + \phi(t)) - \phi^{[-1]}(\phi(u_1) + \phi(v_2) + \phi(t)) \\ &= C\left(C(u_2, v_2), t\right) - C\left(C(u_1, v_2), t\right) \\ &\leq C(u_2, v_2) - C(u_1, v_2), \end{aligned}$$

de modo que C es bicreciente. ■

La caracterización completa sobre la formación de cópulas arquimedianas con este procedimiento se explica en el siguiente teorema.

Teorema 2.1. *Sea $\phi: [0, 1] \rightarrow [0, \infty]$ una función continua, estrictamente decreciente tal que $\phi(1) = 0$. Entonces la función $C: [0, 1] \times [0, 1] \rightarrow [0, 1]$ definida por*

$$C(u, v) = \phi^{[-1]}\left(\phi(u) + \phi(v)\right)$$

es una cópula si y sólo si ϕ es convexa.

Demostración. Hemos ya probado que las condiciones de frontera se satisfacen. Queda por demostrar que C es bicreciente, para lo cuál usaremos la proposición 6, es decir, probaremos que

$$\phi \text{ es convexa} \Leftrightarrow C(u_2, v) - C(u_1, v) \leq u_2 - u_1 \quad \forall u_1 \leq u_2 \in [0, 1] \quad (2.4)$$

Observemos que $C(u_2, v) - C(u_1, v) \leq u_2 - u_1$ equivale con

$$u_1 + \phi^{[-1]}(\phi(u_2) + \phi(v)) \leq u_2 + \phi^{[-1]}(\phi(u_1) + \phi(v))$$

y usando la sustitución $\phi(u_i) = a_i$ y $\phi(v) = b$ esto equivale con

$$\phi^{[-1]}(a_1) + \phi^{[-1]}(a_2 + b) \leq \phi^{[-1]}(a_2) + \phi^{[-1]}(a_1 + b) \quad (2.5)$$

De este modo, buscamos probar que ϕ es convexa si y sólo si para todo $a_1 \leq a_2$, y $b \geq 0$ se verifica (2.5).

\Rightarrow]

Supongamos (2.5) válida, y para puntos $s, t \in [0, \infty]$ nombremos $a_1 = (s + t)/2$, $a_2 = s$, y $b = (s - t)/2$. Entonces

$$\phi^{[-1]} \left(\frac{s + t}{2} \right) \leq \frac{\phi^{[-1]}(s) + \phi^{[-1]}(t)}{2}$$

es decir, $\phi^{[-1]}$ es convexa al punto medio, y siendo también continua, se sigue que es convexa. Con ello se prueba que ϕ también es convexa (pues ϕ es convexa si y sólo si $\phi^{[-1]}$ lo es).

\Leftarrow]

Supongamos ahora que ϕ es convexa y consideremos $a_1 \leq a_2$ y $b \geq 0$. Definamos $\alpha = (a_1 - a_2)/(a_1 - a_2 + b)$. Dado que $a_1 = (1 - \alpha)a_2 + \alpha(a_1 + b)$ y $a_2 + b = \alpha a_2 + (1 - \alpha)(a_1 + b)$

$$\begin{aligned} \phi^{[-1]}(a_1) &\leq (1 - \alpha)\phi^{[-1]}(a_2) + \alpha\phi^{[-1]}(a_1 + b) \quad y \\ \phi^{[-1]}(a_2 + b) &\leq \alpha\phi^{[-1]}(a_2) + (1 - \alpha)\phi^{[-1]}(a_1 + b) \end{aligned}$$

Sumando estas dos desigualdades obtenemos (2.5). ■

A la función ϕ se le llama *generador*. En caso de que ϕ sea invertible de inicio, se dice que es un *generador estricto* o que C es una *cópula arquimediana estricta*.

Si bien el Teorema 2.1 nos brinda un método para construir una infinidad de cópulas arquimedianas, podemos tener la impresión de que no son cópulas muy usuales, sino que son, de algún modo, artificiales. Los siguientes ejemplos muestran que no es así.

Ejemplo 2.1. *La cópula producto es una cópula arquimediana.*

Demostración. Basta tomar la función $\phi(t) = -\ln(t)$. Como $\phi(0) = \infty$ entonces ϕ es estricta, de modo que $\phi^{[-1]}(t) = \phi^{-1}(t) = e^{-t}$ lo cual trae por resultado la cópula

$$C(u, v) = \exp \left(- [(-\ln(u)) + (-\ln(v))] \right) = uv$$

■

Ejemplo 2.2. *La cota inferior de Fréchet es una cópula arquimediana.*

Demostración. En este caso, basta tomar la función $\phi(t) = 1 - t \quad \forall t \in [0, 1]$. Entonces la pseudo inversa será

$$\phi^{[-1]}(t) = \begin{cases} 1 - t, & \text{si } t \in [0, 1], \\ 0, & \text{si } t > 1 \end{cases}$$

Lo anterior quiere decir que $\phi^{[-1]}(t) = \max(1 - t, 0)$. Utilizando la ecuación conocida, obtenemos que

$$C(u, v) = \phi^{[-1]}((1 - u) + (1 - v)) = \max(u + v - 1, 0)$$

Por lo cual W es una cópula arquimediana no estricta de generador $\phi(t) = 1 - t$. ■

Con esto hemos visto que dos de las cópulas con que hemos trabajado anteriormente son arquimedias. Una pregunta válida es si existen cópulas que no sean arquimedias. Ahora veremos que así es, en particular, M , la cota superior de Fréchet, no lo es.

Ejemplo 2.3. *La cota superior de Fréchet, M , no es una cópula arquimediana.*

Demostración. Observemos que si C es una cópula arquimediana, entonces para todo $u \in [0, 1]$

$$u = C(u, 1) \geq \delta_C(u) = C(u, u) = \phi^{[-1]}(2\phi(u))$$

Supongamos que $u = \phi^{[-1]}(2\phi(u))$. Si aplicamos ϕ tenemos

$$\phi(u) = \begin{cases} 2\phi(u), & \text{si } 2\phi(u) \in [0, \phi(0)], \\ \phi(0), & \text{si } 2\phi(u) > \phi(0). \end{cases}$$

según la definición (2.3) de la función pseudo-inversa de ϕ . De este modo $\phi(u) = 0$ si $2\phi(u) \in [0, \phi(0)]$ y $\phi(u) = \phi(0)$ en otro caso. Por lo tanto si $u = \phi^{[-1]}(2\phi(u))$ entonces $u = 1$ ó $u = 0$. Así para todo $u \in (0, 1)$

$$C(u, u) = \delta_C(u) < u$$

Si la cota superior $M(u, v) = \min(u, v)$ fuese arquimediana, entonces

$$u = \min(u, u) = \delta(u) < u \quad \forall u \in (0, 1)$$

lo cual es una contradicción. Por lo tanto, la cópula M no es arquimediana. ■

Las cópulas arquimedianas son simétricas en el sentido que $C(u, v) = C(v, u)$ para toda pareja (u, v) . Esto nos puede dar la impresión errónea de que toda función de distribución con cópula arquimediana es simétrica. Esto no es así

Ejemplo 2.4. Si (X, Y) son independientes y si $X \sim N(0, 1), Y \sim N(1, 1)$ entonces la distribución de (X, Y) no es simétrica a pesar de que su cópula es arquimediana.

Por ser independientes, la simetría de la distribución conjunta equivale con

$$F_X(x)F_Y(y) = F_X(y)F_Y(x) \quad x, y \in \mathfrak{R} \quad (2.6)$$

Tomemos $x = 0, y = 1$. Como $Y \sim N(1, 1)$, entonces $F_Y(0) = F_X(-1)$, de modo que,

$$\begin{aligned} F_X(x) &= .5 & F_Y(y) &= .5 \\ F_Y(x) &= F_X(-1) & F_X(y) &\approx .85 \end{aligned}$$

De modo que $F_X(x)F_Y(y) = .25$ y $F_X(y)F_Y(x) \approx .1275$. Con ello, por (2.6), la distribución conjunta no es simétrica. ■

En este ejemplo se eligieron distribuciones simétricas. Sin embargo no simétricas con respecto al mismo punto. Esta condición tampoco es suficiente para que la distribución conjunta sea simétrica.

Ejemplo 2.5. Si $X \sim N(0, 1)$ y $Y \sim N(0, 4)$ son variables aleatorias independientes, entonces la distribución conjunta de (X, Y) no es simétrica.

Sabemos que la distribución de $Y/2$ es normal de parámetros $(0, 1)$, con ello, $F_Y(y) = F_X(y/2)$ para todo real y . Consideremos $x = 0, y = 1$. Entonces

$$\begin{aligned} F_X(x) &= .5 & F_Y(y) &= F_X(.5) \\ F_Y(x) &= .5 & F_X(y) &= F_X(1) \end{aligned}$$

Obtenemos que si F_{XY} es simétrica entonces, por (2.6), $F_X(1) = F_X(.5)$ lo cual es imposible pues la distribución normal estándar es estrictamente creciente. ■

El único caso en que la distribución correspondiente a una cópula arquimediana es simétrica es el de variables idénticamente distribuidas.

Proposición 7. *Sea C una cópula arquimediana correspondiente a la distribución F del vector aleatorio (X, Y) . Entonces F es simétrica en el sentido que $F(x, y) = F(y, x)$ para todo punto $(x, y) \in \mathfrak{R}^2$ si y sólo si X sigue la misma distribución que Y .*

Demostración. Por el teorema de Sklar tenemos que

$$F(x, y) = F(y, x) \quad \text{si y sólo si} \quad C(F_X(x), F_Y(y)) = C(F_X(y), F_Y(x))$$

para cada punto (x, y) de \mathfrak{R}^2 . Si C es arquimediana lo anterior se traduce en

$$C(F_Y(y), F_X(x)) = C(F_X(y), F_Y(x))$$

para todo punto. Conforme $x \rightarrow \infty$ obtenemos que $F_Y(y) = F_X(y)$ para cada punto $y \in \mathfrak{R}$. Análogamente, conforme $y \rightarrow \infty$ obtenemos que $F_X(x) = F_Y(x)$ para todo $x \in \mathfrak{R}$. En conclusión la distribución F con cópula arquimediana C es simétrica si y sólo si X y Y tienen la misma distribución. ■

Veremos ahora dos propiedades más de las cópulas arquimedias. Primero, que cada cópula arquimediana tiene tantos generadores como números reales, y después una razón para el nombre “arquimedias”.

Proposición 8. *Si ϕ es una función generadora para C entonces, para toda $k \neq 0$ la función $k\phi(\cdot)$ también lo es.*

Demostración. Calculemos la pseudo-inversa de $k\phi(\cdot)$ utilizando la definición (2.3) y nombrando $\psi(x) = k\phi(x)$. Tenemos

$$\psi^{[-1]}(t) = \begin{cases} \psi^{-1}(t) & \text{si } t \in \text{Ran}(\psi) \\ 0 & \text{si } t \in [\psi(0), \infty] \end{cases}$$

Siendo la condición $t \in \text{Ran}(\psi)$ equivalente con $t = k\phi(y)$ p.a. y , tenemos que $t \in \text{Ran}(\psi) \Leftrightarrow t/k \in \text{Ran}(\phi)$. Además, por las propiedades de la función inversa podemos afirmar que $\forall t \in \text{Ran}(\psi) \psi^{-1}(t) = \phi^{-1}(\frac{1}{k}t)$. Por lo que la regla de correspondencia para $\psi^{[-1]}$ es

$$\psi^{[-1]}(t) = \begin{cases} \phi^{-1}(t/k) & \text{si } t/k \in \text{Ran}(\phi) \\ 0 & \text{si } t/k \in [\phi(0), \infty) \end{cases}$$

De donde tenemos que

$$\psi^{[-1]}(\psi(u) + \psi(v)) = \phi^{[-1]} \left(\frac{\psi(u) + \psi(v)}{k} \right) = \phi^{[-1]}(\phi(u) + \phi(v))$$

con lo cual queda probado que $\forall k \neq 0$, si $\phi(\cdot)$ genera la cópula C entonces $k\phi(\cdot)$ también. ■

Para dar una explicación sobre el nombre de estas cópulas, recordemos la propiedad arquimediana de números reales que afirma que dados $a, b \in \mathfrak{R}$ existe un número natural n tal que $na \geq b$. Definiendo una operación apropiada las cópulas arquimedianas tienen una propiedad análoga.

Para una cópula C , definamos sus *potencias*, recursivamente, como $u_C^1 = u$, $u_C^{n+1} = C(u, u_C^n)$.

Proposición 9. *Si C es una cópula arquimediana estricta con generador ϕ , entonces, para cada $u, v \in [0, 1]$ $\exists n \in \mathbb{N}$ tal que $u_C^n \leq v$.*

Demostración. Recordemos que siendo ϕ un generador estricto, tenemos que $\phi^{[-1]} \equiv \phi^{-1}$.

Observemos que

$$u_C^2 = C(u, u) = \phi^{-1}(\phi(u) + \phi(u)) = \phi^{-1}(2\phi(u))$$

y del mismo modo

$$u_C^3 = C(u_C^2, u) = \phi^{-1}(\phi(\phi^{-1}(2\phi(u)) + \phi(u))) = \phi^{-1}(3\phi(u))$$

dado que $\phi(\phi^{-1}(2\phi(u))) = 2\phi(u)$. Hagamos inducción sobre n . Supongamos que para $n \in \mathbb{N}$ se tiene que $u_C^n = \phi^{-1}(n\phi(u))$ y calculemos u_C^{n+1} .

$$u_C^{n+1} = C(u_C^n, u) = \phi^{-1}(\phi(\phi^{-1}(n\phi(u))) + \phi(u)) = \phi^{-1}((n+1)\phi(u))$$

De este modo $u_C^n = \phi^{-1}(n\phi(u))$ para todo natural n . Aplicando la propiedad arquimediana sobre los números reales $\phi(u)$ y $\phi(v)$ tenemos que existe un natural k tal que $k\phi(u) \geq \phi(v)$. Por lo tanto $v = \phi^{-1}(\phi(v)) \geq \phi^{-1}(k\phi(u)) = u_C^k$. ■

2.3 Propiedades

Hemos definido una cópula arquimediana como la solución de una ecuación funcional en términos de una función $\phi: [0, 1] \rightarrow [0, \infty)$ y de su pseudo-inversa $\phi^{[-1]}: [0, \infty) \rightarrow [0, 1]$. Es de esperarse que las propiedades de la cópula C así definida guarden estrecha relación con las características de ϕ , su generador. Estudiaremos a continuación algunas propiedades referentes al modo en que C distribuye la probabilidad en $[0, 1]$ en términos de su generador.

Dada una cópula C , definimos la curva de nivel t como el conjunto de puntos para los cuáles $C(u, v) = t$, es decir,

$$L_C(t) = \{ (u, v) \text{ t.q. } C(u, v) = t \}.$$

Si C es Arquimediana, tenemos

$$L_C(t) = \{ (u, v) \text{ t.q. } \phi(t) = \phi(u) + \phi(v) \}$$

Generalmente, escribiremos esto como función de u , denotándola como la curva $v = L_t(u)$, ya que podemos resolver la expresión anterior para v obteniendo

$$v = L_t(u) = \phi^{[-1]}(\phi(t) - \phi(u))$$

Observemos que $\phi^{[-1]}$ puede ser remplazada por ϕ^{-1} gracias a que $\phi(t) - \phi(u) \in [0, \phi(0)]$.

Una de las propiedades importantes de las curvas de nivel es la relación que guardan con el soporte. Si la cópula puede ser expresada como la doble integral de alguna función (su densidad), entonces toda curva de nivel tiene probabilidad cero. De este modo, si alguna curva de nivel tiene probabilidad positiva, la cópula tiene una componente singular, esto es, no será expresable como la integral de su densidad en alguna región. Además, si existen curvas de nivel $L_t(u)_i$ tales que $\sum \mathbb{P}(L_t(u)_i) = 1$ entonces la cópula es singular en todo su dominio y su soporte será la unión de dichas curvas.

Para una cópula arquimediana general, veamos la probabilidad que acumula en una curva de nivel dada.

Teorema 2.2. *Sea C una cópula Arquimediana con generador ϕ , entonces para cada $t \in (0, 1)$ la probabilidad que C acumula en la curva de nivel t es*

$$\mathbb{P}_C(L_t(u)) = \phi(t) \left(\frac{1}{\phi'(t^-)} - \frac{1}{\phi'(t^+)} \right) \quad (2.7)$$

Esto es, si (U, V) tienen cópula C entonces

$$\mathbb{P}((U, V) \in L_t(u)) = \phi(t) \left(\frac{1}{\phi'(t^-)} - \frac{1}{\phi'(t^*)} \right)$$

Donde $\phi'(t^+)$ y $\phi'(t^-)$ denotan las derivadas laterales de ϕ en el punto t . En particular, si ϕ es derivable en t –lo cual sucede casi donde quiera– entonces dicha probabilidad es cero.

Demostración. Para $t \in (0, 1)$ definamos $w = \phi(t)$ y consideremos la partición del intervalo $[t, 1]$ inducida por la partición regular $\{0, w/n, 2w/n \dots, w\}$, dada por

$$\{t_{n-k} = \phi^{[-1]}(kw/n)\}_{k=0}^n. \quad (2.8)$$

Dada la definición de la función $\phi^{[-1]}$ expresada en (2.3) y como $w = \phi(t) \leq \phi(0)$ para todo $t \in (0, 1)$, tenemos

$$\begin{aligned} C(t_j, t_k) &= \phi^{[-1]}(\phi(t_j) + \phi(t_k)) \\ &= \phi^{[-1]}\left(\frac{n-j}{n}w + \frac{n-k}{n}w\right) = \phi^{[-1]}\left(w + \frac{n-j-k}{n}w\right) \end{aligned}$$

Esto significa que $C(t_j, t_{n-j}) = t$ para cualquier j .

Usemos los rectángulos $R_k = [t_{k-1}, t_k] \times [t_{n-k}, t_{n-k+1}]$ para cubrir la curva de nivel $L_t(u)$ mediante $S_n = \bigcup_{k=1}^n R_k$. Dado que $\phi^{[-1]}$ es convexa

$$\phi^{[-1]}((n-k/n)w) \leq \frac{\phi^{[-1]}((n-k-1/n)w) + \phi^{[-1]}((n-k+1/n)w)}{2}$$

por lo que $0 \leq t_0 - t_1 \leq t_1 - t_2 \leq \dots \leq 1 - t_{n-1}$. Ahora bien, como ϕ es continua (de nuevo, por ser convexa), se tiene que

$$\lim_{n \rightarrow \infty} 1 - t_{n-1} = 1 - \phi^{[-1]}(0) = 0.$$

De modo que que

$$\mathbb{P}_C(L_t(u)) = \lim_{n \rightarrow \infty} \mathbb{P}(S_n) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{P}(R_k) \quad (2.9)$$

Queda por calcular esta suma. Para cada k

$$\begin{aligned}\mathbb{P}_C(R_k) &= C(t_{k-1}, t_{n-k}) + C(t_k, t_{n-k+1} - t - t) \\ &= (\phi^{[-1]}(w + w/n) - \phi^{[-1]}(w)) - (\phi^{[-1]}(w) - \phi^{[-1]}(w - w/n))\end{aligned}$$

Usando estos términos para calcular (2.9) obtenemos

$$\mathbb{P}_C(L_t(u)) = \lim_{n \rightarrow \infty} w \left(\frac{\phi^{[-1]}(w + w/n) - \phi^{[-1]}(w)}{w/n} - \frac{\phi^{[-1]}(w) - \phi^{[-1]}(w - w/n)}{w/n} \right)$$

Por convexidad de $\phi^{[-1]}$ existen sus derivadas laterales, por lo que el límite anterior es (2.7). ■

Corolario 6. Si C es una cópula arquimediana no-estricta, entonces la probabilidad de la curva $L_0(t)$ es

$$\mathbb{P}_C(L_0(t)) = -\frac{\phi(0)}{\phi'(0^+)}$$

Demostración. Si C no es estricta, entonces $\phi(0)$ es finito, y además sobre y bajo la curva de nivel $L_0(t)$ se tiene que $C(u, v) = 0$ de donde tenemos para cada k que $\mathbb{P}_C(R_k) = C(t_k, t_{n-k+1})$. Sustituyendo esto en (2.9) se obtiene el resultado deseado. ■

Proposición 10. Sea C una cópula arquimediana con generador ϕ . Entonces, si ϕ es lineal por tramos en los intervalos $[0, a_1], [a_1, a_2], \dots, [a_{n-1}, 1]$, C es singular y su soporte está conformado por las curvas de nivel a_i .

Demostración. Se hará la prueba en el caso en que se tienen sólo dos tramos siendo análoga la prueba para otros casos. Considérese $a_1 < 1$ y constrúyase $\phi(x)$ de modo que una linealmente los puntos $(0, b_0), (a_1, b_1), (1, 0)$ con $b_0 > b_1 > 0$. En este caso, las derivadas direccionales serán las pendientes de las rectas que conforman a ϕ . Es así que

$$\begin{aligned}\mathbb{P}_C(L_{a_1}(u)) &= b_1 \left(\frac{a_1}{b_1 - b_0} + \frac{1 - a_1}{b_1} \right) \\ &= \frac{a_1 b_1 - (1 - a_1)(b_1 - b_0)}{b_1 - b_0} \\ &= \frac{b_1 - b_0 + a_1 b_0}{b_1 - b_0}\end{aligned}$$

Y la probabilidad de la curva de nivel 0 es

$$\mathbb{P}_C(L_0(u)) = -b_1 \frac{a_1}{b_1 - b_0} = -\frac{a_1 b_1}{b_1 - b_0}$$

Y, de este modo, $\mathbb{P}_C(L_{a_1}(u)) + \mathbb{P}_C(L_0(u)) = 1$, con lo cual C es singular y su soporte está formado por estas dos curvas de nivel. ■

El caso opuesto a la proposición anterior es aquél en que la cópula C no tenga componentes singulares, es decir, que tenga densidad. El siguiente resultado brinda una expresión para la densidad en términos de ϕ y sus dos primeras derivadas.

Proposición 11. *Sea C una cópula arquimediana absolutamente continua con generador ϕ . Entonces*

$$\frac{\partial^2}{\partial u \partial v} C(u, v) = -\frac{\phi''(C(u, v))\phi'(u)\phi'(v)}{(\phi'(C(u, v)))^3} \quad (2.10)$$

Demostración. De la forma de la cópula arquimediana y usando la regla de la cadena justificada por la existencia c.d.q. de la derivada de ϕ y de $\phi^{[-1]}$ tenemos

$$\begin{aligned} \frac{\partial}{\partial u} C(u, v) &= \phi^{[-1]'}(\phi(u) + \phi(v))\phi'(u) \\ \frac{\partial}{\partial v} \frac{\partial}{\partial u} C(u, v) &= \phi^{[-1]''}(\phi(u) + \phi(v))\phi'(u)\phi'(v) \end{aligned} \quad (2.11)$$

Por otro lado, puesto que $\phi(\phi^{[-1]}(x)) = x$ sobre $Ran(\phi)$ se sigue que en estos puntos

$$\phi'(\phi^{[-1]}(x))\phi^{[-1]'}(x) = 1$$

con lo cual tenemos

$$\phi^{[-1]'}(x) = 1/\phi'(\phi^{[-1]}(x))$$

Así que derivando de nuevo obtenemos

$$\begin{aligned} \phi^{[-1]''}(x) &= -\frac{1}{(\phi'(\phi^{[-1]}(x)))^2} \phi''(\phi^{[-1]}(x))\phi^{[-1]'}(x) \\ &= -\frac{\phi''(\phi^{[-1]}(x))}{(\phi'(\phi^{[-1]}(x)))^3} \end{aligned}$$

Sustituyendo esto en (2.11) se obtiene el resultado deseado. ■

Hemos visto en el capítulo primero que la τ de Kendall se resume en la cópula mediante la expresión (1.42). El valor de $\mathbb{E}(C(U, V))$ puede ser complicado en función de C . Un modo de obtenerlo es através de la función de distribución de $C(U, V)$. En caso que C sea arquimediana hallaremos esta distribución utilizando la misma técnica que en el teorema 2.2

Teorema 2.3. *Sea C una cópula arquimediana con generador ϕ , y (U, V) con cópula C . Entonces, si $t \in I$*

$$K_C(t) = \mathbb{P}(C(U, V) \leq t) = t - \frac{\phi(t)}{\phi'(t^+)} \quad (2.12)$$

Demostración. Consideremos la partición (2.8) del teorema 2.2. Tomemos los rectángulos $T_k = [t_{k-1}, t_k] \times [0, t_{n-k+1}]$ que cubren la región

$$\{(u, v) \text{ t.q. } C(u, v) \leq t\}$$

mediante la unión $Z_n = \bigcup_{k=1}^n T_k$. En este caso, el área la probabilidad buscada está dada por la suma de la probabilidad del intervalo $[0, t] \times [0, 1]$ y $\lim_{n \rightarrow \infty} \mathbb{P}(Z_n)$ es decir

$$\mathbb{P}(C(U, V) \leq t) = t + \lim_{n \rightarrow \infty} \mathbb{P}(Z_n)$$

Para calcular $\mathbb{P}(Z_n)$ veamos que para cada k

$$\mathbb{P}(T_k) = C(t_k, t_n - k + 1) - t = \phi^{[-1]}(w - w/n) - \phi^{[-1]}(w)$$

Con lo cuál

$$\mathbb{P}(Z_n) = -w \left(\frac{\phi^{[-1]}(w) - \phi^{[-1]}(w - w/n)}{w/n} \right)$$

y, tomando límite, se verifica (2.12). ■

Con este resultado, calculemos la tau de Kendall para cópulas arquimedias.

Proposición 12. *Sea C una cópula arquimediana con generador ϕ , y sean X, Y tales que $C_{XY} = C$, entonces*

$$\tau_{X,Y} = 1 + \int_0^1 \frac{\phi(t)}{\phi'(t)} dt$$

Demostración. Tenemos que

$$4\mathbb{E}(C(U, V)) - 1 = 4 \int_0^1 t dK_C(t) - 1.$$

Integrando por partes, considerando $u = t$, $dv = dK_C(t)$, tenemos que

$$4\mathbb{E}(C(U, V)) - 1 = 4(tK_C(t))\Big|_0^1 - \int_0^1 K_C(t) dt - 1 = 3 - 4 \int_0^1 K_C(t) dt$$

Sabemos que $K_C(t) = t - \phi(t)/\phi'(t^+)$ según hemos probado en el teorema 2.3, con lo cual

$$\tau_{XY} = 3 - 4 \int_0^1 t - \phi(t)/\phi'(t^+) dt = 1 + 4 \int_0^1 \phi(t)/\phi'(t) dt$$

justificando el cambio de $\phi'(t^+)$ por $\phi'(t)$ dado que $\phi'(t^+) = \phi'(t)$ c.d.q. ■

Observación 6. La expresión del Teorema 2.3 para la τ de Kendall no es única. Si C es absolutamente continua podemos usar la forma

$$\tau_{XY} = 1 - 4 \iint_{I^2} \frac{\partial}{\partial u} C(u, v) \frac{\partial}{\partial v} C(u, v) du dv$$

para concluir que

$$\tau_{XY} = 1 - 4 \int_0^\infty u \left(\frac{d}{du} \phi^{[-1]}(u) \right)^2 du$$

Teorema 2.4. Si $(U, V) \sim C$ entonces la distribución de la pareja $(U, C(U, V))$ es

$$F(s, t) = \begin{cases} s, & \text{si } s \leq t \\ t - \frac{\phi(t) - \phi(s)}{\phi'(t^+)}, & \text{si } s > t. \end{cases} \quad (2.13)$$

para valores $s, t \in [0, 1]$.

Demostración. Si $s \leq t$ y $U \leq s$ entonces, como $C(U, V) \leq U$ se sigue que $C(U, V) \leq s \leq t$, es decir $\{C(U, V) \leq t, U \leq s\} = \{U \leq s\}$, y así $F(s, t) = s$. Si $s > t$, sean $z = \phi(s)$, $w = \phi(t)$. De modo similar a (2.8), considérese la partición del intervalo $[t, s]$ inducida por la partición regular del intervalo $[z, w]$. Tenemos que $t_{n-k} = \phi^{[-1]}(z + (k(w-z)/n))$ y que $C(t_j, t_k) = \phi^{[-1]}(w + (n-j-k)(w-z)/n)$. Haciendo el mismo procedimiento que antes, obtenemos (2.13). ■

Teorema 2.5. *Sea C una cópula arquimediana con generador ϕ , y sea (U, V) con distribución C . Entonces las variables*

$$S = \frac{\phi(U)}{\phi(U) + \phi(V)} \quad T = C(U, V)$$

son independientes y $S \sim U(0, 1)$

Demostración. Consideréense las variables transformadas $X = \phi(U)$, $Y = \phi(V)$ cuya distribución conjunta está dada por

$$\begin{aligned} \mathbb{P}(\phi(U) \leq x, \phi(V) \leq y) &= \mathbb{P}(U \geq \phi^{[-1]}(x), V \geq \phi^{[-1]}(y)) \\ &= 1 - \mathbb{P}(U \leq \phi^{[-1]}(x)) - \mathbb{P}(V \leq \phi^{[-1]}(y)) \\ &\quad + \mathbb{P}(U \leq \phi^{[-1]}(x), V \leq \phi^{[-1]}(y)) \\ &= 1 - \phi^{[-1]}(x) - \phi^{[-1]}(y) + \phi^{[-1]}(x + y) \end{aligned}$$

para cada x, y en el rango de ϕ . Siendo $\phi, \phi^{[-1]}$ funciones convexas, se sigue que dicha distribución admite una densidad casi en todo punto dada por

$$\begin{aligned} f_{XY}(x, y) &= \frac{\partial}{\partial y} \left(-\frac{1}{\phi'(\phi^{[-1]}(X))} + \frac{1}{\phi'(\phi^{[1]}(y))} \right) \\ &= -\frac{\phi^{[-1]'}(x + y)}{(\phi'(\phi^{[-1]}(x + y)))^2} \\ &= -\frac{1}{(\phi'(\phi^{[-1]}(x + y)))^3} \end{aligned}$$

Para cada $0 \leq s \leq 1$, y $t > 0$

$$\mathbb{P}(S \leq s, T \leq t) = \mathbb{P}\left(X \leq \frac{sY}{1-s}, X \geq \phi(t) - Y\right) \quad (2.14)$$

Siendo ϕ convexa, es derivable casi seguramente, de modo que Y admite una densidad, $-\frac{d}{dy}\phi^{[-1]}(y) = -1/\phi'(\phi^{[-1]}(y))$ sobre el rango de ϕ , con lo cual podemos calcular

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f_{XY}(x, y)}{f_Y(y)} \\ &= \frac{\phi'(\phi^{[-1]}(y))}{(\phi'(\phi^{[-1]}(x + y)))^3} \end{aligned}$$

Y con ella, para cada valor de Y tenemos

$$\begin{aligned}
 \mathbb{P}(X \leq x|Y = y) &= \int_0^x \frac{\phi'(\phi^{[-1]}(y))}{(\phi'(\phi^{[-1]}(\alpha + y)))^3} d\alpha \\
 &= \int_0^x \left(\frac{1}{\phi'(\phi^{[-1]}(\alpha + y))} \right)' d\alpha \\
 &= \phi'(\phi^{[-1]}(y)) \left(-\frac{1}{\phi'(\phi^{[-1]}(x + y))} + \frac{1}{\phi'(\phi^{[-1]}(y))} \right) \\
 &= 1 - \frac{\phi'(\phi^{[-1]}(y))}{\phi'(\phi^{[-1]}(x + y))}
 \end{aligned}$$

Siendo así podemos escribir la ecuación (2.14) condicionando sobre el valor de Y como

$$\begin{aligned}
 &\int_0^{\phi(0)} \mathbb{P}(\phi(v) - y \leq X \leq sy/(1-s)|Y = y) f_Y(y) dy \\
 &= \int_{(1-u)\phi(v)}^{\phi(v)} - \frac{\mathbb{P}(\phi(v) - y \leq X \leq sy/(1-s)|Y = y)}{\phi'(\phi^{[-1]}(y))} dy \\
 &\quad + \int_{\phi(v)}^{\phi(0)} - \frac{\mathbb{P}(X \leq sy/(1-s)|T = t)}{\phi'(\phi^{[-1]}(y))} dy
 \end{aligned}$$

Puesto que sobre el conjunto $[0, (1-u)\phi(v)]$

$$\mathbb{P}(\phi(v) - y \leq X \leq sy/(1-s)|Y = y) = 0$$

y sobre el conjunto $[\phi(v), \phi(0)]$

$$X \geq \phi(v) - y.$$

Evaluando estas integrales según los cálculos ya realizados se obtiene que

$$\mathbb{P}(S \leq s, T \leq t) = s \left(t - \frac{\phi(t)}{\phi'(t)} \right)$$

Por el Teorema 2.3 se concluye que

$$\mathbb{P}(S \leq s, T \leq t) = s\mathbb{P}(T \leq t)$$

con lo cual (S, T) son v.a. independientes y $S \sim U(0, 1)$. ■

Observación 7. Si C es absolutamente continua, esto es, derivable en todo punto, entonces podríamos probar este teorema con más facilidad utilizando el teorema de cambio de variable.

Demostración. Gracias al teorema de cambio de variable la densidad conjunta de S, T está dada por

$$\frac{\partial^2}{\partial u \partial v} C(u, v) \left| \frac{\partial(u, v)}{\partial(s, t)} \right|$$

Utilizando el resultado de (2.11) en la proposición 11 y que el jacobiano para la transformación es

$$\frac{\partial(u, v)}{\partial(s, t)} = \frac{\phi(t)\phi'(t)}{\phi'(u)\phi'(v)}$$

obtenemos que

$$h(s, t) = \left(-\frac{\phi''(t)\phi'(u)\phi'(v)}{(\phi'(t))^3} \right) \cdot \left(-\frac{\phi(t)\phi'(t)}{\phi'(u)\phi'(v)} \right) = \frac{\phi''(t)\phi(t)}{(\phi'(t))^2}$$

De donde, integrando,

$$H(s, t) = s \cdot \left(y - \frac{\phi(y)}{\phi'(y)} \right)_0^t = s \cdot (t - \phi(t)/\phi'(t))$$

como se buscaba probar. ■

Observación 8. Dado que la distribución de (U, V) se puede obtener a partir de aquella de (S, T) se sigue que si

1. $C(U, V) \sim K_C(t) = t - \frac{\phi(t)}{\phi'(t)}$

2. Las variables aleatorias

$$S = \frac{\phi(U)}{\phi(U) + \phi(V)} \quad T = C(U, V)$$

son independientes y

3. $S \sim U(0, 1)$

entonces C es arquimediana de generador ϕ .

Una consecuencia fundamental del teorema 2.3 es que a partir de la distribución de $C(U, V)$, $K_C(t)$, se puede recobrar la cópula C , es decir, se puede resolver la ecuación diferencial

$$\frac{\phi(t)}{\phi'(t)} = t - K_C(t) \quad (2.15)$$

para ϕ , puesto que es equivalente con

$$\frac{d}{dt} \ln(\phi(t)) = \frac{1}{t - K_C(t)}$$

y su solución es

$$\phi(t) = \exp\left(\int_{t_0}^t \frac{1}{x - K_C(x)} dx\right) = \exp\left(\int_{t_0}^t \frac{1}{\lambda(x)} dx\right) \quad (2.16)$$

donde $\lambda(x) = x - K_C(x) = \phi(x)/\phi'(x)$. Esta función está bien definida y genera una cópula arquimediana siempre que $t - K_C(t)$ sea negativo y permanezca acotado lejos del 0 en el intervalo unitario. Formalmente hablando

Proposición 13. Sean U, V variables aleatorias uniformemente distribuidas en $[0, 1]$ con cópula C . Defínase $K(t^-) = \lim_{s \nearrow t} K(s) = \mathbb{P}(C(U, V) < t)$. La función ϕ definida en (2.16) es convexa y decreciente y satisface que $\phi(1) = 0$ si y sólo si para cada $0 < t < 1$ $K(t^-) > t$.

Demostración. La demostración de este hecho puede encontrarse en [2]. ■

En vista de esta proposición se concluye que de entre todas las cópulas para las cuales la distribución de $X = C(U, V)$, F_X satisface que $F_X(x^-) > x$ en su dominio, la cópula arquimediana $C(u, v) = \phi^{[-1]}(\phi(u) + \phi(v))$ es la única para la cuál las partes 2 y 3 de la observación 8 son ciertas.

Con todo lo desarrollado somos capaces de generar cópulas arquimedianas siempre que brindemos un generador. No obstante, dada una cópula C , ¿Cómo saber si es o no es arquimediana? Se puede responder a esta pregunta mediante el teorema de Ling (1965) cuya prueba se halla en [4]

Teorema 2.6. Sea C una cópula. Entonces C es arquimediana si y sólo si

1. C es asociativa, i.e. $C(C(u, v), w) = C(u, C(v, w))$ para cada tríada de puntos u, v, w en el intervalo unitario.

2. C es simétrica, i.e. $C(u, v) = C(v, u)$ para cada pareja $(u, v) \in I^2$ ■

Aún si sabemos que C es arquimediana necesitamos algún modo de encontrarle un generador. Para ello se puede utilizar el siguiente teorema

Teorema 2.7. *Sea C una cópula arquimediana con generador ϕ entonces para casi toda u, v*

$$\frac{\phi'(u)}{\phi'(v)} = \frac{\frac{\partial}{\partial u} C(u, v)}{\frac{\partial}{\partial v} C(u, v)}$$

Demostración. Derivando la expresión $\phi(C(u, v)) = \phi(u) + \phi(v)$ con respecto a u tenemos que

$$\phi'(C(u, v)) \frac{\partial}{\partial u} C(u, v) = \phi'(u)$$

y derivando con respecto a v

$$\phi'(C(u, v)) \frac{\partial}{\partial v} C(u, v) = \phi'(v)$$

Como $\phi'(x) \neq 0$ cuando existe, dado que ϕ es estrictamente decreciente, se comprueba el resultado. ■

Utilicemos este resultado para una cópula específica

Ejemplo 2.6. *Sea $C_\theta(u, v) = \frac{uv}{1-\theta(1-u)(1-v)}$, con $\theta \in [-1, 1]$. Entonces C es arquimediana y es generada por la función $\ln \frac{1-\theta(1-t)}{t}$.*

Demostración. C_θ es simétrica para cada θ por conmutatividad simplemente. Probar la asociatividad es un proceso largo, pero sencillo, que se omite para enfocarnos en encontrar el generador. Evaluemos las derivadas parciales y usemos el teorema para obtener

$$\frac{\phi'(u)}{\phi'(v)} = \frac{v(1-\theta(1-v))}{u(1-\theta(1-u))}$$

Se concluye que $\phi'(u) = \alpha/t(1-\theta(1-t))$ -de hecho $\alpha > 0$ gracias a que $\phi'(t) < 0$ -, con lo cual

$$\phi(t) = \frac{\alpha}{1-\theta} \ln \left(\frac{1-\theta(1-t)}{t} \right)$$

Poniendo $\alpha = 1 - \theta$ para cada $\theta \in [-1, 1)$ se obtiene el generador buscado. Recuérdese que la constante se puede escoger en vista de lo demostrado en la proposición 8. ■

Para concluir el presente capítulo se presenta una breve tabla de cópulas arquimedianas.

Familia de.	$C_\theta(u, v)$	$\phi_\theta(t)$	$\theta \in$
Ali-Mikhail-Haq	$\frac{uv}{1-\theta(1-u)(1-v)}$	$\ln \frac{1-\theta(1-t)}{t}$	$[-1, 1)$
Gumbel-Barnett	$uv \exp(-\theta \ln u \ln v)$	$\ln(1 - \theta \ln t)$	$(0, 1]$
Gumbel	$\exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right)$	$-\ln \frac{e^{-\theta t}-1}{e^{-\theta}-1}$	$[1, \infty)$
Frank	$-\frac{1}{\theta} \ln\left(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right)$	$-\ln \frac{e^{-\theta t}-1}{e^{-\theta}-1}$	$\mathfrak{R} \setminus \{0\}$
Cook-Johnson	$\max\left([u^{-\theta} + v^{-\theta} - 1]^{-1/\theta}, 0\right)$	$\frac{1}{\theta}(t^{-\theta} - 1)$	$[-1, \infty) \setminus \{0\}$

Tabla 2.1: Familias de Cópulas arquimedianas con un parámetro

Estas familias aparecen con frecuencia en la modelación. Encontramos que, dado el umbral de sus parámetros, presentan ciertos casos especiales, como se resume en la tabla siguiente

2.4 Conclusiones

La familia de las cópulas arquimedianas es ciertamente rica al contar con muchos y muy variados miembros según lo afirma el teorema 2.1. Un aspecto importante de esta riqueza es que tanto la cópula producto como la cota inferior de Fréchet son cópulas arquimedianas a pesar de que la cota superior no lo sea.

Una gran virtud de las cópulas arquimedianas es que la distribución de la variable $C(U, V)$ hallada en el teorema 2.3 es suficiente para conocer el generador de C y así a C misma. Con esto, las cópulas arquimedianas son

Familia de Cópulas	Casos límite y especiales
Ali-Mikhail-Haq	$C_0 = \Pi, \quad C_1 = \frac{uv}{u+v-uv}$
Gumbel-Barnett	$C_0 = \Pi$
Gumbel	$C_1 = \Pi \quad C_\infty = M$
Frank	$C_{-\infty} = W, \quad C_0 = \Pi, \quad C_\infty = M$
Cook-Johnson	$C_{-1} = W \quad C_0 = \Pi \quad C_\infty = M \quad C_1 = \frac{uv}{u+v-uv}$

Tabla 2.2: Casos especiales

caracterizadas por una función univariada. Como esta función es una distribución, podremos aproximarla dada una colección de datos y con ello podremos construir una cópula arquimediana a partir de muestras gracias a este teorema como se hará en el siguiente capítulo.

Capítulo 3

Procesos de Markov

3.1 Introducción

Un tema importante en probabilidad es el de los procesos de Markov. Estos tienen la propiedad de que las variables aleatorias que le conforman dependen de las anteriores sólo a través de la más reciente, en palabras llanas, son procesos para los cuales el futuro y el pasado son independientes dado el presente. En este capítulo veremos cómo las propiedades de un proceso de Markov se relacionan con las cópulas y estudiaremos un modo de definir procesos Markovianos.

En la sección 3.2 definimos un producto en el espacio de las cópulas bidimensionales con el cuál mostramos una relación entre las cópulas de parejas de variables y la independencia condicional entre ellas. En la sección 3.3 se estudia la propiedad de Chapman–Kolmogorov que poseen las probabilidades de transición de cualquier proceso de Markov y se muestra que es equivalente con que la familia de cópulas del proceso satisfaga una condición relativa al producto antes definido. Con esto se propone un nuevo modo de construir procesos estocásticos no a través de las probabilidades de transición sino a través de las distribuciones marginales y una familia de cópulas.

En la sección 3.4 abordamos de modo sencillo el caso del movimiento Browniano para construir con sus probabilidades de transición una familia de cópulas que satisfaga la condición de Chapman–Kolmogorov. Construimos otra familia de cópulas que satisfacen esta condición usando una familia paramétrica mediante una transformación en el parámetro y calculamos la cópula asociada a la transición de m pasos de una cadena de Markov.

En la sección 3.5 definimos un nuevo producto para caracterizar de modo completo la propiedad de Markov con cópulas. Aprovechando esta equivalencia construimos usando cópulas un proceso que satisface la condición de Chapman–Kolmogorov pero no es de Markov. Concluye el capítulo con un estudio sobre algunas propiedades analíticas y algebraicas que resultan del producto definido sobre el espacio de las cópulas bidimensionales. Con los resultados obtenidos se construye un proceso determinístico salvo en un intervalo acotado de tiempo.

El material del presente capítulo se basó principalmente en [3]. Como apoyo se usó [1].

3.2 Un producto para cópulas bidimensionales

En el estudio de los diversos modos de dependencia entre variables aleatorias a través de su cópula hemos recurrido hasta ahora a características fundamentalmente geométricas. Por ejemplo, en la sección 1.5 estudiamos el soporte y su relación con nexos deterministas entre las variables. Estudiamos también que la comparación con la cópula producto define la dependencia por cuadrante y que la monotonía de la función $C(u, v)/u$ determina la monotonía de colas. Demostramos que las medidas usuales de concordancia son tan sólo distancias promedio en algún sentido.

Definiremos y analizaremos un producto para cópulas bidimensionales que nos ayudará a extender este estudio. Para ello, es necesario hacer algunas aclaraciones.

Recordemos del capítulo primero que dada una cópula C , las derivadas parciales $\partial C/\partial u$ y $\partial C/\partial v$ existen y son crecientes, con lo cual son derivables casi dondequiera (c.d.q.), es decir, tiene sentido hablar de

$$\begin{aligned} C^{,1}(u, v) &= \frac{\partial}{\partial u} C(u, v) & C^{,2}(u, v) &= \frac{\partial}{\partial v} C(u, v) \\ C^{,1,2} &= \frac{\partial}{\partial u \partial v} C(u, v) & C^{,2,1} &= \frac{\partial}{\partial v \partial u} C(u, v) \end{aligned}$$

Podemos ahora definir el siguiente producto para cópulas: Dadas A, B cópulas y $u, v \in I$ definimos

$$A * B(u, v) = \int_0^1 A^{,2}(u, t) B^{,1}(t, v) dt$$

Esta integral existe por tratarse de funciones derivables (y por tanto continuas) c.d.q.

Observación 9. La definición de este producto implica que es distributivo por la derecha y por la izquierda sobre combinaciones convexas.

Teorema 3.1. *La operación $*$ es cerrada en \mathcal{C} , el conjunto de las cópulas bidimensionales.*

Demostración. Sean $A, B \in \mathcal{C}$ y sea $C = A * B$, probemos que $C \in \mathcal{C}$. Para las evaluaciones en cero

$$C(u, 0) = \int_0^1 A,^2(u, t)B,^1(t, 0)dt = \int_0^1 0dt = 0$$

y análogamente para $C(0, v)$. Ahora bien

$$C(1, v) = \int_0^1 A,^2(1, t)B,^1(t, v)dt = \int_0^1 B,^1(t, v)dt = v$$

La última igualdad se obtiene por ser $B_v(x) = B(x, v)$ absolutamente continua como se muestra en el Lema 2. El trabajo para la otra variable es análogo.

Queda por demostrar que C es bicreciente. Para ello, sean $u_1 \leq u_2$, $v_1 \leq v_2$ puntos en el intervalo unitario, entonces

$$\begin{aligned} & C(u_1, v_1) + C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) \\ &= \int_0^1 A,^2(u_1, t)B,^1(t, v_1)dt + \int_0^1 A,^2(u_2, t)B,^1(t, v_2)dt \\ &\quad - \int_0^1 A,^2(u_1, t)B,^1(t, v_2)dt - \int_0^1 A,^2(u_2, t)B,^1(t, v_1)dt \\ &= \int_0^1 A,^2(u_2, t)(B,^1(t, v_2) - B,^1(t, v_1))dt \\ &\quad - \int_0^1 A,^2(u_1, t)(B,^1(t, v_2) - B,^1(t, v_1))dt \\ &= \int_0^1 (A,^2(u_2, t) - A,^2(u_1, t))(B,^1(t, v_2) - B,^1(t, v_1))dt \geq 0 \end{aligned}$$

Ha quedado probado que C es una cópula. ■

Calculemos el producto de una cópula C dada con las tres cópulas M, W, Π

$$\Pi * C(u, v) = \int_0^1 u C,^1(t, v) dt = u(C(1, v) - C(0, v)) = uv$$

$$C * \Pi(u, v) = \int_0^1 C,^2(u, t) v dt = v(C(u, 1) - C(u, 0)) = vu$$

$$M * C(u, v) = \int_0^1 \mathbb{1}(t < u) C,^1(t, v) dt = C(u, v) - C(0, v) = C(u, v)$$

$$C * M(u, v) = \int_0^1 C,^2(u, t) \mathbb{1}(t < v) dt = C(u, v) - C(v, 0) = C(u, v)$$

$$W * C(u, v) = \int_0^1 \mathbb{1}(t > 1 - u) C,^1(t, v) dt = C(1, v) - C(1 - u, v) = v - C(1 - u, v)$$

$$C * W(u, v) = \int_0^1 C,^2(u, t) \mathbb{1}(t > 1 - v) dt = C(u, 1) - C(u, 1 - v) = u - C(u, 1 - v)$$

Observación 10. Para el producto $*$ el elemento π es nulo y el elemento M neutro.

Para interpretar en términos probabilistas este producto, probemos que está relacionado con la distribución condicional de parejas de variables aleatorias.

Teorema 3.2. Sean X, Y variables aleatorias continuas con cópula C , entonces

$$\mathbb{P}(X \leq x | Y) \stackrel{c.s.}{=} C,^2(F_X(x), F_Y(Y)) \quad (3.1)$$

$$\mathbb{P}(Y \leq y | X) \stackrel{c.s.}{=} C,^1(F_X(X), F_Y(y)) \quad (3.2)$$

Demostración. Probaremos la relación (3.1), siendo análoga la prueba para (3.2). Si denotamos por $\sigma(Y)$ la sigma-álgebra generada por Y , debemos probar que para cada $E \in \sigma(Y)$

$$\int_E \mathbb{1}(X \leq x) d\mathbb{P}(w) = \int_E C,^2(F_X(x), F_Y(Y(w))) d\mathbb{P}(w)$$

En particular, basta probarlo para conjuntos de la forma $E = Y^{-1}(-\infty, a]$.

Calculemos

$$\begin{aligned} \int_E C,^2(F_X(x), F_Y(Y(w)))d\mathbb{P}(w) &= \int_{-\infty}^a C,^2(F_X(x), F_Y(\nu))dF_Y(\nu) \\ &= \int_0^{F_Y(a)} C,^2(F_X(x), \nu)d\nu = C(F_X(x), F_Y(a)) - C(F_X(x), 0) \\ &= C(F_X(x), F_Y(a)) \end{aligned}$$

Ahora bien,

$$\int_E \mathbb{1}(X \leq x) d\mathbb{P}(w) = \int_{\Omega} \mathbb{1}(X \leq x, Y \leq a) d\mathbb{P}(w) = C(F_X(x), F_Y(a))$$

Queda así probado (3.1). ■

Corolario 7. Sean X, Y, Z variables aleatorias continuas tales que X y Y sean condicionalmente independientes dado Z . Entonces

$$C_{XY} = C_{XZ} * C_{ZY}.$$

Demostración. Por la independencia condicional, sabemos que

$$\mathbb{P}(X \leq x, Y \leq y|Z) = \mathbb{P}(X \leq x|Z)\mathbb{P}(Y \leq y|Z)$$

Aplicando en esta expresión el resultado del Teorema 3.2 tenemos que

$$\mathbb{P}(X \leq x, Y \leq y|Z) = C,^2_{XZ}(F_X(x), F_Z(Z))C,^1_{ZY}(F_Z(Z), F_Y(y))$$

Integrando con respecto a los valores de Z obtenemos

$$\begin{aligned} &\int_{\text{Ran}(Z)} \mathbb{P}(X \leq x, Y \leq y|\theta)dF_Z(\theta) \\ &= \int_{\text{Ran}(Z)} C,^2_{XZ}(F_X(x), F_Z(\theta))C,^1_{ZY}(F_Z(\theta), F_Y(y))dF(\theta) \\ &= \int_0^1 C,^2_{XZ}(F_X(x), \psi)C,^1_{ZY}(\psi, F_Y(y))d\psi = C_{XZ} * C_{ZY} \end{aligned}$$

De modo que $C_{XY} = C_{XZ} * C_{ZY}$. ■

3.3 El producto * y las ecuaciones de Chapman-Kolmogorov

La independencia condicional es un concepto importante en probabilidad. A partir de ella se define un proceso de Markov como sigue.

Definición 9. Sea $\{X_t\}_{t \in I}$ una colección de variables aleatorias. Diremos que $\{X_t\}_{t \in I}$ es un proceso de Markov si para todo conjunto de índices $t_1 < t_2 < \dots < t_n < s < t$ tenemos

$$\mathbb{P}(X_t \leq x | X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_n} = x_n, X_s = y) = \mathbb{P}(X_t \leq x | X_s = y)$$

es decir si X_t es condicionalmente independiente de la colección $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ dado X_s .

Como vimos en la sección anterior el producto * que hemos definido se relaciona con la independencia condicional. En lo siguiente relacionaremos una propiedad de los procesos de Markov con este producto.

Hagamos la convención de que dada una función de dos (o más) variables $f(x, y)$, el símbolo $\int_A g(t) f(dt, y)$ representa la integral sobre el conjunto A de la función g con respecto a la función $h(t) = f(t, y)$, y fijo.

Teorema 3.3. Sea $\{X_t\}_{t \in T}$ un proceso de Markov, y para cada $s, t \in T$ denotemos C_{st} la cópula del vector (X_s, X_t) . Entonces, son equivalentes

1. Las probabilidades de transición $P(s, x, t, A) = P(X_t \in A | X_s = x)$ del proceso satisfacen las ecuaciones de Chapman-Kolmogorov

$$P(s, x, t, A) = \int_{\mathfrak{R}} P(u, \psi, t, A) P(s, x, u, d\psi) \quad (3.3)$$

para todo conjunto Boreliano A , para cada $s < u < t$ y para casi todo $x \in \mathfrak{R}$.

2. Para cada $s < u < t$

$$C_{st} = C_{su} * C_{ut} \quad (3.4)$$

Antes de demostrar este Teorema, necesitaremos de dos lemas.

3.3. EL PRODUCTO * Y LAS ECUACIONES DE CHAPMANN-KOLMOGOROV79

Lema 11. Sea B una cópula y $f \in \mathcal{L}_1([0, 1])$, entonces, si $x \leq 1$

$$\int_0^x \int_0^1 f(t)B,^1(u, dt)du = \int_0^1 f(t)B(x, dt)$$

Demostración. Observemos que los operadores

$$T(f) = \int_0^x \left(\int_0^1 f(t)B,^1(u, dt) \right) du, \quad S(f) = \int_0^1 f(t)B(x, dt)$$

son lineales y acotados (De hecho $\|T\| \leq \|B\|$ y $\|S\| \leq \|B\|$), y por tanto, continuos. Además, si $f(t) = \mathbb{1}((a, b))$ entonces

$$T(f) = \int_0^x B,^1(u, b) - B,^1(u, a)du = B(x, b) - B(x, a)$$

y por otro lado

$$S(f) = \int_a^b B(x, dt) = B(x, b) - B(x, a)$$

Es decir $T(f) = S(f)$ para funciones indicadoras. De este modo, si f es una función escalonada continua por la derecha entonces los operadores S y T coinciden. Siendo esta clase de funciones densa en $\mathcal{L}_1([0, 1])$ se sigue de la continuidad de S y T que $S \equiv T$. ■

Lema 12. Sean A, B dos cópulas. Entonces, para casi toda u

$$\frac{\partial}{\partial u} \int_0^1 A,^1(t, v)B,^2(u, t)dt = \int_0^1 A,^1(t, v)B,^1(x, dt) \quad (3.5)$$

Demostración. Consideremos $\varphi \in C^\infty$ tal que $\varphi(0) = \varphi(1) = 0$, entonces

$$\int_0^1 \left(\varphi(u) \int_0^1 A,^1(t, v)B,^1(u, dt) \right) du = - \int_0^1 \varphi'(u) \left(\int_0^u \int_0^1 A,^1(t, v)B,^1(\psi, dt)d\psi \right) du.$$

Esto se verifica integrando por partes. Usando ahora el lema anterior, Lema Lema 11, podemos escribir lo anterior como

$$- \int_0^1 \varphi'(u) \left(\int_0^1 A,^1(t, v)B(u, dt) \right) du \quad (3.6)$$

Como $B(u, t)$ es derivable con respecto a t y su derivada $B,^2(u, t)$ es continua, podemos escribir (3.6) como

$$- \int_0^1 \varphi'(u) \left(\int_0^1 A,^1(t, v) B(u, t) dt \right) du$$

Al integrar por partes en la expresión

$$\int_0^1 \varphi(u) \frac{\partial}{\partial u} \left(\int_0^1 A,^1(t, v) B,^2(u, t) dt \right) du$$

observamos que

$$\begin{aligned} - \int_0^1 \varphi'(u) \left(\int_0^1 A,^1(t, v) B(u, t) dt \right) du = \\ \int_0^1 \varphi(u) \frac{\partial}{\partial u} \left(\int_0^1 A,^1(t, v) B,^2(u, t) dt \right) du \end{aligned}$$

Hemos probado que para toda $\varphi \in C^\infty$ con $\varphi(0) = \varphi(1) = 0$

$$\begin{aligned} \int_0^1 \left(\varphi(u) \int_0^1 A,^1(t, v) B,^1(u, dt) \right) du = \\ \int_0^1 \varphi(u) \frac{\partial}{\partial u} \left(\int_0^1 A,^1(t, v) B,^2(u, t) dt \right) du \end{aligned}$$

de modo que la igualdad (3.5) es válida. ■

Demostración al Teorema 3.3. Probemos primero que si $C_{st} = C_{su} * C_{st}$ entonces se satisfacen las ecuaciones de Chapman–Kolmogorov. Es suficiente verificarlo para conjuntos A de la forma $(-\infty, a)$ (pues $A \mapsto P(s, x, t, A)$ es una medida de probabilidad para cada $s < t$, y casi todo x). En estos conjuntos tenemos que las probabilidades de transición son

$$P(s, x, t, A) = \mathbb{P}(X_t \leq a | X_s = x) = C_{st},^1(F_s(x), F_t(a)) \quad (3.7)$$

según el teorema 3.2. Siendo de este modo, para conjuntos $A = (-\infty, a)$

tenemos

$$\begin{aligned}
 & \int_{\mathfrak{R}} P(u, \psi, t, A) P(s, x, u, d\psi) \\
 &= \int_{\mathfrak{R}} C_{ut},^1(F_u(\psi), F_t(a)) C_{su},^1(F_s(x), F_u(d\psi)) \\
 &= \int_0^1 C_{ut},^1(\theta, F_t(a)) C_{su},^1(F_s(x), d\theta) \\
 &= \left. \frac{\partial}{\partial \beta} \int_0^1 C_{ut},^1(\theta, F_t(a)) C_{su}^2(\beta, \theta) d\theta \right]_{\beta=F_s(x)} \\
 &= (C_{su} * C_{ut}),^1(F_s(x), F_t(a))
 \end{aligned} \tag{3.8}$$

Y gracias a haber supuesto que $C_{st} = C_{su} * C_{ut}$ se sigue la igualdad con (3.7). Ahora bien, si las ecuaciones de Chapman-Kolmogorov se satisfacen, la primera expresión de (3.8) es igual a $C_{st},^1(F_s(x), F_t(a))$. Esto quiere decir que

$$(C_{su} * C_{ut}),^1(F_s(x), F_t(a)) = C_{st},^1(F_s(x), F_t(a))$$

para $x \in \mathfrak{R}$, $a \in \mathfrak{R}$ se obtiene que $C_{st},^1 = (C_{su} * C_{st}),^1$. Finalmente, de la expresión

$$C_{st}(x, y) = \int_0^x C_{st},^1(\alpha, y) d\alpha$$

se obtiene que $C_{st} = C_{su} * C_{ut}$. ■

Este hecho permite aproximarse al estudio de los procesos de Markov de un modo distinto al tradicional en que se da la distribución inicial y una familia $\{P_t\}$ que satisfaga Chapman-Kolmogorov. Ahora, daremos todas las distribuciones marginales y una familia de cópulas bivariadas que satisfagan (3.4). Una diferencia importante entre estos métodos es que dando la distribución inicial y las probabilidades de transición (enfoque tradicional) se varían las distribuciones marginales necesariamente, mientras que al fijar las cópulas y variar la distribución inicial no se afecta ninguna otra marginal.

3.4 Ejemplos

Ejemplo 3.1. *Podemos cacular las cópulas de algún proceso conocido, como el Browniano, y después especificar un nuevo proceso cambiando las dis-*

tribuciones marginales, obteniendo un movimiento Browniano que tenga marginales no-Gaussianas.

Las probabilidades de transición en el movimiento Browniano son

$$P(s, x, t, (-\infty, y]) = \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^y \exp(-(u-x)^2/2(t-s)) du$$

para cada $0 < s < t$. Sabiendo que

$$P(s, x, t, (-\infty, y]) = C_{s,t}^{-1}(F_s(x), F_t(y))$$

podemos integrar para obtener

$$C_{st}(F_s(x), F_t(y)) = \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^x \left(\int_{-\infty}^y \exp(-(u-v)^2/2(t-s)) du \right) dF_s(v)$$

para $0 < s < t$. Hagamos el cambio de variable $F_s(v) = z$, entonces $dF_s(v) = dz$ y si $v = -\infty$ entonces $z = 0$, y si $v = x$ entonces $z = F_s(x)$, de modo que

$$C_{st}(F_s(x), F_t(y)) = \frac{1}{\sqrt{2\pi(t-s)}} \int_0^{F_s(x)} \left(\int_{-\infty}^y \exp(-(u - F_s^{-1}(z))^2/2(t-s)) du \right) dz$$

Para encontrar $C_{st}(x, y)$ evaluemos la anterior expresión en $F_s^{-1}(x), F_t^{-1}(y)$ para obtener

$$C_{st}(x, y) = \frac{1}{\sqrt{2\pi(t-s)}} \int_0^x \left(\int_{-\infty}^{F_t^{-1}y} \exp(-(u - F_s^{-1}(z))^2/2(t-s)) du \right) dz$$

Haciendo el cambio de variable

$$\theta = \frac{u - F_s^{-1}(z)}{\sqrt{t-s}} \quad d\theta = \frac{du}{\sqrt{t-s}}$$

y notando que si $u = -\infty$ entonces $\theta = -\infty$ y que si $u = F_t^{-1}(y)$ entonces $\theta = (F_t^{-1}(y) - F_s^{-1}(z))/\sqrt{t-s}$ se obtiene que

$$C(x, y) = \int_0^x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta} \exp(-\theta^2/2) d\theta dz$$

usando $\beta = (F_t^{-1}(y) - F_s^{-1}(z))/\sqrt{t-s}$. Si suponemos que $X_0 \stackrel{c.s.}{=} 0$ y nombramos por $\phi(x)$ la distribución normal estándar evaluada en x tenemos que $F_t(x) = \phi(x/\sqrt{t})$ para todo t , de modo que

$$C(x, y) = \int_0^x \phi\left(\frac{\sqrt{t}\phi^{-1}(y) - \sqrt{s}\phi^{-1}(u)}{\sqrt{t-s}}\right) du$$

Por construcción estas cópulas satisfacen (3.4) por lo cual basta especificar cualesquiera marginales continuas distintas a la Gaussiana para obtener un proceso con cópulas iguales a las del Movimiento Browniano, pero con marginales no-Gaussianas. ■

Ejemplo 3.2. *Podemos construir procesos de Markov utilizando familias paramétricas de cópulas como*

$$C_\alpha(u, v) = uv + 3\alpha uv(1-u)(1-v)$$

para cada $\alpha \in [-1/3, 1/3]$.

Notemos que en este caso

$$C_\alpha * C_\beta = C_{\alpha\beta}$$

calculando directamente

$$\int_0^1 [u + 3\alpha u(1-u)(1-t) - 3\alpha ut(1-u)][v + (1-t)3\beta v(1-v) - 3\beta vt(1-v)] dt$$

según la definición del producto $*$, tenemos

$$\begin{aligned} & xy + 3\alpha u(1-u)v \int_0^1 (1-t) dt - 3\alpha uv(1-u) \int_0^1 t dt + 3\alpha v(1-v)u \int_0^1 (1-t) dt \\ & + v(1-v)3\alpha u(1-u) \int_0^1 (1-t)^2 dt - 3\alpha v(1-v)3\beta u(1-u) \int_0^1 (1-t)t dt \\ & - 3\alpha v(1-v)u \int_0^1 t dt - 3\alpha v(1-v)3\beta u(1-u) \int_0^1 t(1-t) dt \\ & + 3\alpha v(1-v)3\beta u(1-u) \int_0^1 t dt. \end{aligned}$$

Simplificando la ecuación obtenemos

$$C_\alpha * C_\beta = C_{\alpha \cdot \beta}. \quad (3.9)$$

Si consideramos, para cada $s \leq t$ que $\alpha(s, t) = \exp(s - t)$ y nombramos $C_{st} = C_{\alpha(s, t)}$ entonces, las propiedades de la exponencial aunadas a (3.9) implican que la familia C_{st} satisface las ecuaciones de Chapman–Kolmogorov (3.4), de modo que al escogerse una familia de marginales continuas $\{F_s\}$ se tiene un proceso Markoviano.

Obsérvese el comportamiento asintótico

$$\lim_{t \rightarrow \infty} C_{st} = \lim_{\alpha \rightarrow 0} C_\alpha = C_0 = \pi$$

de modo que conforme t se "aleja" más de s , las variables X_s, X_t tendrán una cópula cada vez más parecida a la de variables aleatorias independientes. ■

Ejemplo 3.3. *Considérese una cadena de Markov $\{X_n\}_{n \in \mathbb{N}}$ con valores en un conjunto finito de enteros positivos. En este ejemplo examinaremos cómo asociar una matriz de transición de m pasos con una cópula para dicha transición en cadenas estacionarias.*

Consideremos, para simplificar los cálculos, que la cadena de Markov $\{X_n\}$ toma valores en el conjunto $\{1, 2, 3\}$. Tomemos la matriz de transición en m pasos por $Q = P^m$, donde P es la matriz de transición en un paso. De este modo $Q_{jk} = \mathbb{P}(X_{n+m} = k | X_n = j)$, con $m > 0$. Denotemos por p_k la probabilidad de que X_n tome el valor $k \in \{1, 2, 3\}$ que podemos calcular utilizando la distribución inicial y las probabilidades de transición de los estados anteriores del proceso mediante la recursión.

$$\mathbb{P}(X_n = k) = \sum_{j=1}^3 \mathbb{P}(X_n = k | X_{n-1} = j) \mathbb{P}(X_{n-1} = j)$$

Escribamos $p = (p_1, p_2, p_3)$, la densidad de X_n y notemos que para X_{n+m} tenemos

$$\mathbb{P}(X_{n+m} = 1) = \sum_{k=1}^3 \mathbb{P}(X_{n+m} = 1 | X_n = k) p_k$$

por lo cual la densidad de X_{n+m} es pQ . Ahora bien, las distribuciones son

$$\begin{aligned} F_n(1) &= p_1 & F_{n+m}(1) &= (pQ)_1 \\ F_n(2) &= p_1 + p_2 & F_{n+m}(2) &= (pQ)_1 + (pQ)_2 \\ F_n(3) &= 1 & F_{n+m}(3) &= 1 \end{aligned} \quad (3.10)$$

Si utilizamos el Teorema 3.2 tendremos que

$$C(F_n(x), F_{n+m}(y)) = \int_0^x \mathbb{P}(X_{n+m} \leq y | X_n = \theta) dF_n(\theta)$$

Evaluando esta expresión obtenemos los valores de C en los nueve puntos que las distribuciones determinan. Recordando la prueba de Sklar podemos extender esta función a una cópula de modo que en los cuadrados de la forma

$$[F_n(i), F_n(j)] \times [F_{n+m}(k), F_{n+m}(l)]$$

se acumula probabilidad $p_j Q_{j\ell}$ uniformemente según muestra la figura.

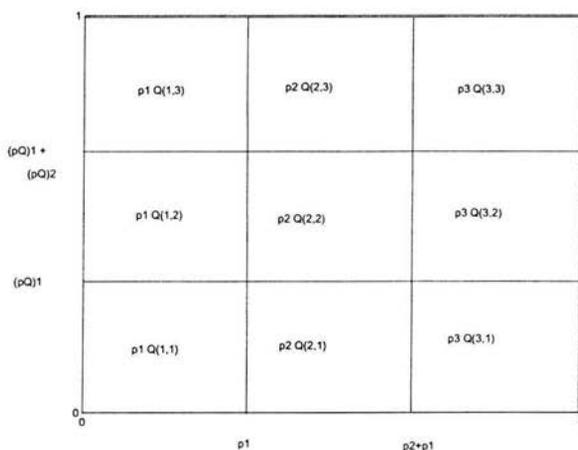


Figura 3.1: Probabilidad asignada por la cópula C a los nueve rectángulos del cuadrado unitario

Es importante notar que esta cópula no depende sólo del tamaño de la transición, es decir, de la matriz de transición Q sino de la distribución de X_n . De este modo, si se varían las distribuciones de las variables anteriores a X_n , aunque se preserve la distribución inicial, las cópulas del proceso han de variar. ■

3.5 El producto $*$ y la propiedad de Markov

Habiendo caracterizado las ecuaciones de Chapman–Kolmogorov con el producto $*$, definiremos ahora un nuevo producto para caracterizar la propiedad de Markov (que, más adelante se verá, no son equivalentes). Esta vez, tomemos cópulas A de m variables, B de n variables y definamos la operación $A \bullet B : [0, 1]^{m+n-1} \rightarrow [0, 1]$ como

$$\begin{aligned} A \bullet B(x_1, x_2, \dots, x_{m+n-1}) \\ = \int_0^{x_m} A,^m(x_1, x_2, \dots, x_{m-1}, \psi) B,^1(\psi, x_{m+1}, \dots, x_{m+n-1}) d\psi \end{aligned}$$

Recordemos que una colección $\{X_t\}_{t \in T}$ es un proceso de Markov si y sólo si para todo conjunto de índices $t_1 < t_2 < \dots < t_n < s < t$ tenemos

$$\mathbb{P}(X_t \leq x | X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_n} = x_n, X_s = y) = \mathbb{P}(X_t \leq x | X_s = y) \quad (3.11)$$

Tenemos el siguiente resultado.

Teorema 3.4. *Un proceso estocástico real $\{X_t\}_{t \in T}$ es de Markov si y sólo si para cada entero positivo n y para cada colección t_1, t_2, \dots, t_n estrictamente creciente de T*

$$C_{t_1, t_2, \dots, t_n} = C_{t_1, t_2} \bullet C_{t_2, t_3} \cdots \bullet C_{t_{n-1}, t_n} \quad (3.12)$$

donde C_{t_1, t_2, \dots, t_n} es la cópula de $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ y C_{t_i, t_j} la cópula de X_{t_i}, X_{t_j} .

Demostración. Debemos mostrar que (3.12) del Teorema implica (3.11) y viceversa. En lo sucesivo, por facilidad en la notación, pondremos F_1 en vez de F_{t_1} , $C_{1,2}$ por C_{t_1, t_2} , etc. Examinemos la demostración cuando $n = 3$. Para ello, tomemos $t_1 < t_2 < t_3$ elementos de T , y probemos que (3.11) si y sólo si

$$\mathbb{P}(X_1 \leq x_1, X_3 \leq x_3 | X_2) \stackrel{c.s.}{=} \mathbb{P}(X_1 \leq x_1 | X_2) \mathbb{P}(X_3 \leq x_3 | X_2) \quad (3.13)$$

es decir, a la independencia condicional de X_1 y X_3 dado X_2 .

Supongamos pues (3.11) y tomemos B un Boreliano, entonces

$$\begin{aligned}
& \int_{X_2^{-1}(B)} \mathbb{1}(X_1 \leq x_1) \mathbb{1}(X_3 \leq x_3) d\mathbb{P} \\
&= \int_{X_2^{-1}(B) \cap X_1^{-1}(-\infty, x_1]} \mathbb{1}(X_3 \leq x_3) d\mathbb{P} \\
&= \int_{X_2^{-1}(B) \cap X_1^{-1}(-\infty, x_1]} \mathbb{P}(X_3 \leq x_3 | X_1, X_2) d\mathbb{P} \\
&= \int_{X_2^{-1}(B)} \mathbb{1}(X_1 \leq x_1) \mathbb{P}(X_3 \leq x_3 | X_2) d\mathbb{P} \\
&= \int_{X_2^{-1}(B)} \mathbb{E} \left(\mathbb{1}(X_1 \leq x_1) \mathbb{P}(X_3 \leq x_3 | X_2) | X_2 \right) d\mathbb{P} \\
&= \int_{X_2^{-1}(B)} \mathbb{P}(X_1 \leq x_1 | X_2) \mathbb{P}(X_3 \leq x_3 | X_2) d\mathbb{P}
\end{aligned}$$

Utilizando el hecho de que

$$\int_{X_2^{-1}(B)} \mathbb{1}(X_1 \leq x_1) = \int_{X_2^{-1}(B)} \mathbb{P}(X_1 \leq x_1 | X_2) d\mathbb{P}$$

para cualquier conjunto Boreliano B y que $\mathbb{P}(X_3 \leq x_3 | X_2)$ es medible en $\{X_2^{-1}(B) : B \text{ Boreliano}\}$, la sigma-álgebra generada por X_2 . Nótese que (3.11) se utilizó para que $\mathbb{P}(X_3 \leq x_3 | X_1, X_2) = \mathbb{P}(X_3 \leq x_3 | X_2)$. De este modo, (3.11) implica (3.13). Para probar el recíproco basta utilizar las mismas igualdades en distinto orden.

Si integramos ambos lados de (3.13) sobre el conjunto $X_2^{-1}(-\infty, x_2)$ obtenemos que

$$\begin{aligned}
& \int_{X_2 \leq x_2} \mathbb{P}(X_1 \leq x_1, X_3 \leq x_3 | X_2) d\mathbb{P} = \int_{X_2 \leq x_2} \mathbb{1}(X_1 \leq x_1, X_3 \leq x_3) d\mathbb{P} \\
&= \int_{X_2 \leq x_2} C_{12}^2(F_1(x_1), F_2(X_2)) C_{23}^1(F_2(X_2), F_3(x_3)) d\mathbb{P} \\
&= \int_{-\infty}^{x_2} C_{1,2}^2(F_1(x_1), F_2(\theta)) C_{2,3}^1(F_2(\theta), F_3(x_3)) dF_2(\theta) \\
&= \int_0^{F_2(x_2)} C_{12}^2(F_1(x_1), \alpha) C_{23}^1(\alpha, F_3(x_3)) d\alpha \\
&= C_{12} \bullet C_{23}(F_1(x_1), F_2(x_2), F_3(x_3))
\end{aligned}$$

Queda así probado que, en el caso en que $n = 3$, (3.12) es resultado de la propiedad de Markov (3.13).

Para probar que (3.12) implica (3.13) en caso que $n = 3$, hemos de probar que para cada conjunto $B \in \sigma(X_2)$

$$\int_B \mathbb{1}(X_1 \leq x_1) \mathbb{1}(X_3 \leq x_3) d\mathbb{P} = \int_B \mathbb{P}(X_1 \leq x_1 | X_2) \mathbb{P}(X_3 \leq x_3 | X_2) d\mathbb{P}$$

Es suficiente demostrarlo para los generadores de la forma $X_2^{-1}(-\infty, x_2]$, y para estos conjuntos tendremos que

$$\begin{aligned} \int_{X_2^{-1}(-\infty, x_2]} \mathbb{1}(X_1 \leq x_1) \mathbb{1}(X_3 \leq x_3) d\mathbb{P} &= C_{123}(F_1(x_1), F_2(x_2), F_3(x_3)) \\ &= \int_0^{F_2(x_2)} C_{12, 2}(F_1(x_1), \theta) C_{23}^1(\theta, F_3(x_3)) d\theta \\ &= \int_{-\infty}^{x_2} C_{12, 2}(F_1(x_1), F_2(\alpha)) C_{23}^1(F_2(\alpha), F_3(x_3)) dF_2(\alpha) \\ &= \int_{X_2^{-1}(-\infty, x_2]} C_{12, 2}(F_1(x_1), F_2(X_2)) C_{23}^1(F_2(X_2), F_3(x_3)) d\mathbb{P} \\ &= \int_{X_2^{-1}(-\infty, x_2]} \mathbb{P}(X_1 \leq x_1 | X_2) \mathbb{P}(X_3 \leq x_3 | X_2) d\mathbb{P} \end{aligned}$$

En caso que $n > 3$ la prueba es similar. En este caso, tendremos que la condición de Markov es equivalente con que

$$\begin{aligned} \mathbb{P}(X_1 \leq x_1, \dots, X_{n-1} \leq x_{n-1}, X_t \leq x | X_n) \\ \stackrel{c.s.}{=} \mathbb{P}(X_1 \leq x_1, \dots, X_{n-1} \leq x_{n-1} | X_n) \mathbb{P}(X_t \leq x | X_n) \end{aligned} \quad (3.14)$$

En breve, para probar que la propiedad de Markov (3.14) implica (3.11)

tendremos

$$\begin{aligned}
 C(F_1(x_1), \dots, F_n(x_n), F_t(x)) &= \\
 & \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} \mathbb{P}(X_t \leq x | X_1 = t_1, \dots, X_n = t_n) dF_{12\dots n}(t_1, t_2, \dots, t_n) \\
 &= \int_0^{F_1(x_1)} \dots \int_0^{F_n(x_n)} C_{nt, 1}(\theta_n, F_t(x)) C_{12\dots n}(d\theta_1, \dots, d\theta_n) \\
 &= \int_0^{F_n(x_n)} C_{nt, 1}(F_1(x_1), F_2(x_2), \dots, F_{n-1}(x_{n-1}), d\theta_n) \\
 &= C_{12\dots n} \bullet C_{nt}(F_1(x_1), \dots, F_n(x_n), F_t(x))
 \end{aligned}$$

Y el reverso es análogo al que se ha hecho para el caso $n = 3$. ■

Corolario 8. *Las ecuaciones de Chapman-Kolmogorov no determinan un proceso de Markov.*

Demostración. Construyamos un proceso del siguiente modo. Para las variables X_1, X_2, X_3 (o cualesquiera otras tres), tomemos la cópula

$$C_\alpha(u, v, w) = uvw + \alpha u(1-u)v(1-v)w(1-w)$$

para $|\alpha| < 1$. Digamos que cualquier conjunto finito de variables que no contenga a las tres X_1, X_2, X_3 es independiente, mientras que la cópula de un conjunto que los contenga será

$$C_{1,\dots,m}(u_1, \dots, u_m) = C_\alpha(u_1, u_2, u_3) \bullet P_{4\dots m}(u_4, \dots, u_m)$$

Como se argumenta en [3], por el Teorema de Kolmogorov, esto conforma un proceso estocástico al establecer distribuciones marginales. Gracias a que las variables de este proceso son independientes por parejas y a que Π es \ast -neutro se satisfacen las ecuaciones de Chapman-Kolmogorov, no obstante

$$\begin{aligned}
 C_{123}(u_1, u_2, u_3) &\neq C_{12} \bullet C_{23}(u_1, u_2, u_3) \\
 &= \int_0^{u_2} C_{12, 2}(u_1, \theta) C_{23, 1}(\theta, u_3) d\theta = u_1 u_2 u_3
 \end{aligned}$$

De modo que, no cumpliendo con (3.12) el proceso especificado no es Markoviano. ■

3.6 Estructura de $(\mathcal{C}, *)$.

Un aspecto importante de toda la construcción hecha hasta el momento es que al dotar al espacio de las cópulas bidimensionales \mathcal{C} con el producto $*$ hemos generado una estructura algebraica. ¿Que propiedades tiene esta estructura y en qué se traducen? Para examinar esta pregunta, enriquezcamos el panorama definiendo otra operación, $(\cdot)^T: \mathcal{C} \rightarrow \mathcal{C}$, pidiendo que

$$A^T(u, v) = A(v, u) \quad \forall u, v.$$

Con estos dos elementos veamos que significa en términos de procesos de Markov la existencia de elementos inversos o extremos.

Definición 10. Una cópula $A \in \mathcal{C}$ es un elemento extremo cuando el hecho de que

$$A = \lambda B + (1 - \lambda)C \quad \lambda \in (0, 1)$$

implica que $B = C = A$.

Definición 11. Una cópula A es invertible por la izquierda cuando existe $B \in \mathcal{C}$ tal que $B * A = M$. Diremos que es invertible por la derecha si existe $C \in \mathcal{C}$ tal que $A * C = M$. Llamaremos a B y C inverso izquierdo e inverso derecho respectivamente. En caso que ambos existan diremos que C es invertible.

Observación 11. Si C es invertible entonces sus inversos derecho e izquierdo coinciden.

Demostración. Sean A y B dichos inversos, entonces

$$C * A = M = A * B \quad \text{y así} \quad C = C * M = C * A * B = M * B = B$$

■

Teorema 3.5. Una cópula $C \in \mathcal{C}$ es invertible por la izquierda si y sólo si para cada $v \in [0, 1]$, $C^1(u, v) = 0$ ó 1 para casi todo $u \in [0, 1]$; y en este caso C^T es un inverso izquierdo. C tiene inverso derecho si y sólo si para todo $u \in [0, 1]$ $C^2(u, v) = 0$ ó 1 para casi todo $v \in [0, 1]$; y en este caso C^T es un inverso derecho.

Demostración. Demostraremos sólo para invertibilidad por la izquierda siendo análogo el otro caso. Supongamos primero que para cada $v \in [0, 1]$, $C,^1(u, v) = 0$ ó 1 para casi todo u . Como el mapeo $v \rightarrow C,^1(u, v)$ es no-decreciente entonces para cada $v_1 \leq v_2$ tenemos $C,^1(u, v_1)C,^1(u, v_2) = C,^1(u, v_1)$ para casi todo $u \in [0, 1]$ puesto que si $C,^1(u, v_1) = 1$ entonces $C,^1(u, v_2) = 1$. Por ello

$$\begin{aligned} C^T * C(x, y) &= \int_0^1 (C^T),^2(u, t)C,^1(t, v)dt \\ &= \int_0^1 C_1(t, u)C,^1(t, v)dt \\ &= \int_0^1 C_1(t, \min(u, v)) \\ &= \min(u, v) = M(u, v) \end{aligned}$$

Con lo cual C tiene inverso izquierdo y este es C^T .

Supongamos ahora que C es invertible por la izquierda y que A es un inverso izquierdo, entonces, utilizando la desigualdad de Hölder y el hecho de que las derivadas de primer orden de cualquier cópula están entre 0 y 1 casi donde quiera, y como $v = \min(v, v)$ tenemos

$$\begin{aligned} v &= \int_0^1 A,^2(v, t)C,^1(t, v)dt \\ &\leq \left(\int_0^1 (A,^2(v, t))^2 dt \right)^{1/2} \left(\int_0^1 (C,^1(t, v))^2 dt \right)^{1/2} \\ &\leq \left(\int_0^1 A,^2(v, t) dt \right)^{1/2} \left(\int_0^1 (C,^1(t, v))^2 dt \right)^{1/2} \\ &\leq v^{1/2} \left(\int_0^1 (C,^1(t, v))^2 dt \right)^{1/2} \\ &\leq v^{1/2} \left(\int_0^1 C,^1(t, v) dt \right)^{1/2} \\ &= v^{1/2} v^{1/2} = v \end{aligned}$$

Se concluye la igualdad de todas las cantidades y por ello que

$$v^{1/2} \left(\int_0^1 (C,^1(t, v))^2 dt \right)^{1/2} = v^{1/2} \left(\int_0^1 C,^1(t, v) dt \right)^{1/2}$$

Es decir

$$\int_0^1 [C^1(t, v) - (C^1(t, v))^2] dt = 0$$

Sabiendo que el integrando es positivo c.d.q., tenemos que para cada $v > 0$, $C^1(u, v) = 0$ ó 1 para casi todo u . Para terminar, nótese que si $v = 0$ entonces $C^1(u, v) = 0$, y así $C^1(u, v) = 0$ ó 1 para casi todo u , con $v \in [0, 1]$. ■

Teorema 3.6. *Todo elemento invertible por la derecha o la izquierda de C es extremo.*

Demostración. Por el Teorema anterior, si C es invertible por la izquierda entonces $C(u, v) = 0$ ó $C(u, v) = 1$ para casi todo u con $v \in [0, 1]$. Como el mapeo $v \rightarrow C^1(u, v)$ es creciente, podemos garantizar que existen dos conjuntos disjuntos $U, V \subseteq [0, 1]$ tales que $U \cup V = [0, 1]$, $C^1(u, v) = 0$ c.d.q. en U y $C^1(u, v) = 1$ c.d.q. en V . Si escribimos $C = \lambda A + (1 - \lambda)B$ con $\lambda \in (0, 1)$ entonces, derivando, $C^1(u, v) = \lambda A^1(u, v) + (1 - \lambda)B^1(u, v)$ de donde $0 = A^1(u, v) = B^1(u, v)$ c.d.q. en U y $1 = A^1(u, v) = B^1(u, v)$ c.d.q. en V . Se concluye que en casi todo punto $A^1 = B^1 = C^1$ de donde, integrando, se obtiene que C es un elemento extremo. ■

Teorema 3.7. *Los inversos derechos e izquierdos son únicos.*

Demostración. Si C es una cópula invertible por la izquierda (derecha) de inversos izquierdos (derechos) A y B entonces la cópula $\frac{1}{2}A + \frac{1}{2}B$ es inverso izquierdo por que $*$ se distribuye sobre combinaciones convexas. Siendo así $\frac{1}{2}A + \frac{1}{2}B$ es un elemento extremo, al tener inverso derecho (izquierdo) C . Se sigue inmediatamente que $A = B$ (pues un elemento extremo ha sido escrito como combinación lineal no-trivial de A y B). ■

En términos de procesos de Markov, las cópulas invertibles corresponden a nexos deterministas entre las variables del proceso de modo que si C_{st} es invertible entonces existe una función Borel-medible f tal que $X_t \stackrel{c.s.}{=} f(X_s)$ (i.e. X_t es $\sigma(X_s)$ medible y viceversa). Formalmente tenemos el siguiente Teorema.

Teorema 3.8. *Sean X_1, X_2 variables aleatorias continuas. Son equivalentes*

1. *Existe una función Borel-medible, f , tal que $X_2 = f(X_1)$ c.s.*

2. Para todo u , $\mathbb{P}(X_2 \leq x | X_1) = 0$ o 1 c.s.

3. La cópula C_{12} es invertible por la izquierda.

Demostración. La equivalencia entre (2) y (3) se sigue de los Teoremas 3.2 y 3.5 por ser X_1, X_2 continuas. Mostremos la equivalencia entre (1) y (2). Supóngase (2). Entonces, para cada conjunto Boreliano, B

$$\begin{aligned} \int_{X_1^{-1}(B)} \mathbb{P}(X_2 \leq x | X_1) \mathbb{1}(X_2 \leq x) d\mathbb{P} &= \int_{X_1^{-1}(B)} \mathbb{E}(\mathbb{P}(X_2 \leq x | X_1) \mathbb{1}(X_2 \leq x) | X_1) d\mathbb{P} \\ &= \int_{X_1^{-1}(B)} X_1^{-1}(B) \mathbb{E}(\mathbb{1}(X_2 \leq x) | X_1)^2 d\mathbb{P} \\ &= \int_{X_1^{-1}(B)} \mathbb{E}(\mathbb{1}(X_2 \leq x) | X_1) d\mathbb{P} \\ &= \int_{X_1^{-1}(B)} \mathbb{1}(X_2 \leq x) d\mathbb{P} \\ &= \int_{X_1^{-1}(B)} \mathbb{1}(X_2 \leq x)^2 d\mathbb{P} \end{aligned}$$

por lo cual podemos afirmar que

$$\begin{aligned} 0 &= \int_{X_1^{-1}(B)} (\mathbb{1}(X_2 \leq x)^2 - 2\mathbb{E}(\mathbb{1}(X_2 \leq x) | X_1) \mathbb{1}(X_2 \leq x) \\ &\quad + \mathbb{E}(\mathbb{1}(X_2 \leq x) | X_1)^2) d\mathbb{P} \\ &= \int_{X_1^{-1}(B)} (\mathbb{1}(X_2 \leq x) - \mathbb{E}(\mathbb{1}(X_2 \leq x) | X_1))^2 d\mathbb{P} \end{aligned}$$

Por tanto $\mathbb{1}(X_2 \leq x) = \mathbb{E}(\mathbb{1}(X_2 \leq x) | X_1)$ c.s. con lo que X_2 es $\sigma(X_1)$ -medible y (1) se concluye.

Ahora bien, si (1) es cierto, entonces $\sigma(X_2) \subseteq \sigma(X_1)$ de donde $\mathbb{1}(X_2 \leq x) = \mathbb{E}(\mathbb{1}(X_2 \leq x) | X_1)$ c.s. con lo que (2) es cierto. ■

La invertibilidad derecha no es ni necesaria ni suficiente para afirmar la existencia de una relación determinista entre las variables. De hecho, si consideramos un proceso Markoviano en que son invertibles todas las cópulas para $u < v \leq s$ y para $t \leq u < v$ donde $s < t$, y C es un elemento invertible por

la derecha pero no por la izquierda y consideramos $E = C^T * C$ y

$$\begin{aligned} C_{uv} &= C_{ut} * C, & u \leq s < v < t \\ C_{uv} &= E, & s < u < v < t \\ C_{uv} &= C^T * C_{tv}, & s < u < t \leq v. \end{aligned}$$

entonces el proceso satisface (3.4). Nótese que $C_{st} = C * C^T = M$ de modo que hay una relación determinista entre X_s y X_t , pero si $s < u < v < t$, $C_{uv} = E$ que no es invertible por la izquierda, de modo que en el intervalo (s, t) el proceso no es determinístico. De este modo la cópula C invertible por la derecha permite que haya un rasgo de aleatoriedad en el proceso que de otro modo sería determinístico.

3.7 Conclusiones

Hemos extendido nuestro estudio sobre la posibilidad de explicarnos una relación entre variables aleatorias mediante su cópula. Por ejemplo, ahora sabemos que en las cópulas de (X, Y) y de (Y, Z) se resume la independencia condicional de dos de ellas dada la tercera. Mostramos que las cópulas asociadas a un proceso determinan la propiedad de Markov. Este método nos llevó a la interesante conclusión de que podemos determinar un proceso estocástico de Markov usando cópulas lo cual nos da la oportunidad de generar procesos en que la relación entre parejas de variables sea muy similar pero las distribuciones marginales sean distintas.

Capítulo 4

Cóputas Empíricas

4.1 Introducción

Este capítulo está dedicado al uso de las cópulas en la práctica, a traer a uso la riqueza de las cópulas. En la sección 4.2 proponemos modos de simular observaciones de distribuciones conjuntas a partir de simular observaciones de cópulas. Presentamos resultados generales y abordamos el caso particular de las cópulas arquimedianas. Mostramos las gráficas de observaciones de distintas distribuciones bivariadas obtenidas a partir de observaciones de cópulas.

En la sección 4.3 damos una versión empírica C_n de la cópula C que relaciona a dos variables a partir de una muestra aleatoria. Se demuestra que esta función converge de modo uniforme hacia la cópula C . A partir de este hecho estudiamos una prueba estadística para probar la hipótesis de independencia y examinamos un tanto su efectividad. Después aplicamos esta muestra a una serie de tiempo real obtenida con los precios del tipo de cambio peso-dólar, fix. Presentamos las rutinas de Matlab utilizadas para los análisis y un resumen de los resultados.

En la sección 4.4, fundamentados en el capítulo anterior, analizamos un modo de estimar cópulas arquimedianas a partir de una muestra y examinamos este estimador. Después aplicamos este método a la muestra de datos del tipo de cambio peso-dólar, fix. Presentamos tanto las rutinas de Matlab utilizadas como un resumen de los resultados.

Utilizamos para esta sección el capítulo 2 de [1], y los artículos [2] y [4].

4.2 Simulación Estocástica

Dada una pareja aleatoria (X, Y) con función de distribución conjunta F , la simulación estocástica consiste en generar observaciones $\{(x_i, y_i)\}_{i=1}^n$ de dicha pareja. Podemos encontrar en la literatura diversos métodos para generar observaciones de una variable aleatoria U uniformemente distribuída en $(0, 1)$. A partir de ello podemos generar observaciones de variables aleatorias continuas con distribución G mediante la composición $X = G^{-1}(U)$.

En el caso bivariado tenemos al menos dos alternativas.

1. Diseñar métodos para generar observaciones de una distribución F directamente, o
2. Utilizar las cópulas y el teorema de Sklar en combinación con el caso univariado

Según el teorema 1.2, de Sklar, la función de distribución bivariada F es la composición de alguna cópula C y las funciones de distribución marginales F_1, F_2 , es decir

$$F(x, y) = C(F_1(x), F_2(y)) \quad \forall (x, y) \in \mathfrak{R}^2$$

De este modo si tenemos un vector (U, V) con distribución C , es decir

$$\mathbb{P}(U \leq u, V \leq v) = C(u, v) \quad \forall (u, v) \in (0, 1)^2$$

entonces el vector aleatorio $(F_1^{-1}(U), F_2^{-1}(V))$ sigue la distribución F pues

$$\begin{aligned} \mathbb{P}(F_1^{-1}(U) \leq x, F_2^{-1}(V) \leq y) &= \mathbb{P}(U \leq F_1(x), V \leq F_2(y)) \\ &= C(F_1(x), F_2(y)) = F(x, y) \end{aligned}$$

Tomando esto en cuenta, podemos simular observaciones de la distribución F del modo siguiente

1. Generamos observaciones (u, v) de (U, V) , el vector con distribución C .
2. Hacemos las composiciones $x = F_1^{-1}(u)$, $y = F_2^{-1}(v)$.
3. Tomamos la pareja (x, y) .

Entonces (x, y) es una observación de la distribución F . Este método simplifica la simulación de distribuciones bivariadas a la simulación de aquellas distribuciones con marginales uniformes. Veamos algunos métodos para la simulación de observaciones de una cópula.

Podemos usar la distribución condicional de V dado que $U = u$, que es

$$\mathbb{P}(V \leq v | U = u) = \lim_{h \rightarrow 0} \frac{C(u+h, v) - C(u, v)}{h} = \frac{\partial}{\partial u} C(u, v) = c_u(v).$$

que existe c.d.q. Usando esta función tenemos el siguiente algoritmo

1. Generamos (u, t) del vector (U, T) de variables aleatorias independientes uniformes en $(0, 1)$,
2. Definimos $v = c_u^{-1}(t)$ como la observación de V ,
3. Tomamos la pareja (u, v) como observación de la cópula C .

La pareja (u, v) se toma como observación de C porque siendo T uniforme la variable $V = c_u^{-1}(T)$ sigue la distribución c_u dado que $U = u$, de modo que para cada $u, v \in (0, 1)$,

$$\mathbb{P}(U \leq u, V \leq v) = \int_0^1 \mathbb{P}(V \leq v | U = u) f_U(u) du = \int_0^1 c_u(v) du = C(u, v)$$

con lo cuál el vector (U, V) sigue la distribución C .

Ejemplo 4.1. Consideremos la cópula dada por

$$C(u, v) = \frac{uv}{u+v-1}$$

Para simularla observemos que $c_u(v) = \left(\frac{v}{u+v-uv}\right)^2$ cuya inversa es $c_u^{-1}(t) = \frac{u\sqrt{t}}{1-(1-u)\sqrt{t}}$. De este modo el algoritmo se expresa como sigue

1. Generar dos uniformes independientes $U = u, T = t$,
2. Definir $v = \frac{u\sqrt{t}}{1-(1-u)\sqrt{t}}$,
3. Tomar la pareja (u, v) .

Ejemplo 4.2. Podemos ahora generar observaciones de dos distribuciones bivariadas con la misma cópula C y distintas marginales.

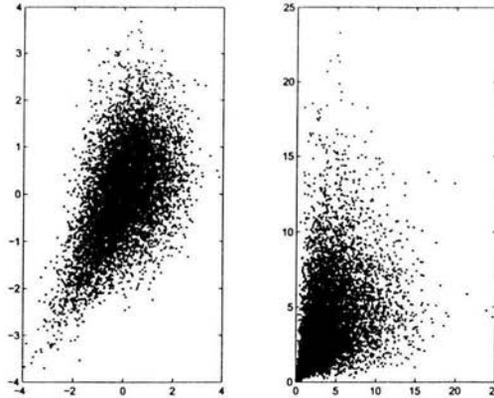


Figura 4.1: 10,000 simulaciones de C compuesta con distintas marginales. A la izquierda marginales normales estándar, a la derecha $\Gamma(2, 2)$.

Podemos también simular utilizando ahora las probabilidades de supervivencia expresadas en la cópula como

$$\mathbb{P}(U \geq u, V \geq v) = 1 - u - v + C(u, v)$$

Expresando esto en términos de las funciones marginales de supervivencia $1 - u$ y $1 - v$ tenemos

$$\mathbb{P}(U \geq u, V \geq v) = (1 - u) + (1 - v) - 1 + C(1 - (1 - u), 1 - (1 - v))$$

y así podemos decir que $\mathbb{P}(U \geq u, V \geq v)$ es la composición de las marginales $1 - u$, y $1 - v$ de supervivencia con la función

$$\tilde{C}(u, v) = u + v - 1 + C(1 - u, 1 - v).$$

Llamaremos a \tilde{C} la cópula de supervivencia.

Proposición 14. Si (U, V) sigue la distribución C , entonces $(1 - U, 1 - V)$ tiene distribución \tilde{C} .

Demostración. Por cálculo directo

$$\begin{aligned} \mathbb{P}(1 - U \leq a, 1 - V \leq v) &= \mathbb{P}(U \geq 1 - a, V \geq 1 - b) \\ &= 1 - (1 - a) - (1 - b) + C(1 - a, 1 - b) = a + b - 1 + C(1 - a, 1 - b) \end{aligned}$$

■

El algoritmo que expresamos antes puede utilizarse para obtener variables con una cópula de supervivencia dada, pues si $U \sim U(0, 1)$ entonces $(1-U) \sim U(0, 1)$ también. Así, podemos pensar en el primer paso del algoritmo como la generación de observaciones de $(1-U, 1-T)$ y a la cópula obtenida como su distribución \bar{C} . De este modo la pareja $(1-u, 1-v)$ representa una observación correspondiente a la cópula C .

En el desarrollo de este método C puede ser cualquier cópula. No obstante para las cópulas arquimedianas tenemos otros métodos que no involucran a la distribución condicional de V dado $U = u$ y que dependen sólo del generador ϕ y la función de distribución $K_C(t)$ correspondiente a la variable $C(U, V)$.

Proposición 15. *Si C es una cópula arquimediana, entonces el siguiente algoritmo genera observaciones de la pareja (U, V) con distribución C .*

1. Tórnense $S = s \perp T = t$ variables generadas uniformes en $[0, 1]$,
2. Defínase $w = K_C^{-1}(t)$
3. Si $u = \phi^{[-1]}(s\phi(w))$ y $v = \phi^{[-1]}((1-s)\phi(w))$, entonces u, v son la pareja deseada.

Demostración. Siendo S, T variables aleatorias uniformes, la distribución de $W = K_C^{-1}(T)$ es K_C , y además S es independiente de W por serlo de T . Esto nos dice que

$$\mathbb{P}(S \leq s, W \leq w) = F_S(s)F_W(w) = sK_C^{-1}(w)$$

Por el teorema 2.5, si (U, V) se distribuye C y nombramos

$$S' = \frac{\phi(U)}{\phi(U) + \phi(V)}, \quad W' = \phi^{[-1]}(\phi(U) + \phi(V)) = C(U, V)$$

entonces los vectores (S, W) y (S', W') siguen la misma distribución de modo que las observaciones generadas (s, w) pueden considerarse como del vector (S', W') .

De este modo la función $s\phi(w)$ representa una observación de $S'(\phi(U) + \phi(V)) = \phi(U)$ y con ello $u = \phi^{[-1]}(s\phi(w))$ es una observación de U . Del mismo modo $(1-s)\phi(w)$ representa una observación de $\phi(V)$. Esto termina la prueba. ■

Ejemplo 4.3. Consideremos la familia paramétrica de cópulas arquimedianas

$$C(u, v) = \left(1 + [(u^{-1} - 1)^\alpha + (v^{-1} - 1)^\alpha]^{1/\alpha}\right)^{-1}$$

con $\alpha \in [1, \infty)$, generada por la función $\phi(t) = (1/t - 1)^\alpha$ y generemos observaciones de ella.

Calculando, se obtiene lo siguiente

$$\phi^{-1}(u) = (u^{1/\alpha} + 1)^{-1}.$$

$$\phi'(t) = -\frac{\alpha}{t^2}(1/t - 1)^{\alpha-1}, \text{ con lo cual, simplificando}$$

$$K(t) = 2t - t^2 \text{ Al invertir}$$

$$K^{-1}(u) = 1 \pm \sqrt{1 - u}.$$

La inversa $K^{-1}(u)$ resulta tomando el signo negativo, pues de otro modo su evaluación no sería menor a 1. El algoritmo manda pues

1. Generar dos variables uniformes en $[0, 1]$ $S = s, T = t$, independientes,
2. definir $w = 1 - \sqrt{t}$,
3. considerar $u = (s^{1/\alpha}(1/w - 1) + 1)^{-1}$, $v = ((1 - s)^{1/\alpha}(1/w - 1) + 1)^{-1}$.

Mostramos la gráfica de la composición de la cópula de parámetro 1 correspondiente a la familia con marginales normales estándar y gamma de parámetros (2, 2).

Otro método posible para cópulas arquimedianas se expresa en la siguiente proposición.

Proposición 16. Si C es una cópula arquimediana, entonces el siguiente algoritmo genera observaciones de la pareja (U, V) con distribución C .

1. Generemos (u, t) del vector (U, T) de variables aleatorias independientes con distribución uniforme en $(0, 1)$;
2. Definamos $w = \phi^{(-1)}(\phi(u)/t)$;
3. Consideremos $v = \phi^{[-1]}(\phi(w) - \phi(u))$;

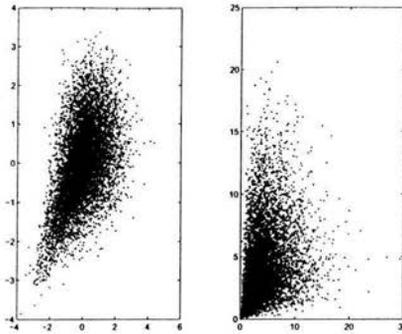


Figura 4.2: 10,000 simulaciones de C_1 compuesta con marginales normales estándar (izquierda) y $\text{gamma}(2,2)$

4. La pareja deseada es (u, v) .

Demostración. Sean U, T variables aleatorias independientes uniformemente distribuidas en $(0, 1)$, y sea $W = \phi'^{(-1)}(\phi'(U)/T)$. Demostremos que W se distribuye K_C . Como ϕ es convexa, la función ϕ' es decreciente, de modo que

$$\mathbb{P}\left(\phi'^{(-1)}\left(\frac{\phi'(U)}{T}\right) \leq x\right) = \mathbb{P}\left(\frac{\phi(U)}{T} \geq \phi'(x)\right)$$

Como $T \geq 0$

$$\mathbb{P}\left(\frac{\phi(U)}{T} \geq \phi'(x)\right) = \mathbb{P}(\phi'(U) \geq T\phi'(x)) = \mathbb{P}\left(T \leq \frac{\phi'(U)}{\phi'(x)}\right)$$

Sabiendo que (U, T) tiene por densidad

$$f_{U,T}(x, y) = \mathbb{1}((x, y) \in [0, 1] \times [0, 1])$$

calculamos esta probabilidad como

$$\begin{aligned} \mathbb{P}\left(T \leq \frac{\phi'(U)}{\phi'(x)}\right) &= \int_0^x \int_{\phi'(u)/\phi'(x)}^1 1 dt du \\ &= \int_0^x \left(1 - \frac{\phi'(u)}{\phi'(x)}\right) du = x - \frac{\phi(x)}{\phi'(x)} \end{aligned} \tag{4.1}$$

Obsérvese que el límite de la primera integral de (4.1) obedece a que $T \leq 1$ y $\phi'(U)/\phi'(x) \geq 1$ si y sólo si $U \geq x$ por ser ϕ' decreciente.

De este modo W sigue la distribución K_C . Como sabemos ésta es la distribución correspondiente a la composición $W' = C(U, V)$ donde (U, V) sigue la distribución C . Con ello, la observación w puede ser tomada como de la variable $C(U, V)$. De este modo la variable $v = \phi^{[-1]}(\phi(w) - \phi(u))$ representa una observación de

$$\phi^{[-1]}(\phi(C(U, V) - \phi(U)) = \phi^{[-1]}(\phi(U) + \phi(V) - \phi(U)) = V$$

con lo cuál (u, v) representa una observación de C . ■

Ejemplo 4.4. Consideremos la familia paramétrica de cópulas arquimedianas

$$C(u, v) = \max([u^{-\theta} + v^{-\theta} - 1]^{-1/\theta}, 0)$$

con $\theta \in (0, \infty)$, generada por la función $\phi(t) = (1/\theta)(t^{-\theta} - 1)$ y generemos observaciones de ella.

Calculando, obtenemos que

$$\phi'(t) = -t^{-(\theta+1)},$$

$$\phi'^{-1}(z) = (-z)^{-1/(\theta+1)},$$

$$\phi^{[-1]}(z) = (1 + \theta z)^{-1/\theta},$$

Con lo cual, el algoritmo manda

1. Generar observaciones de variables aleatorias independientes (u, t) con distribución uniforme en $(0, 1)$,
2. Definir $w = u(t^{1/(\theta+1)})$,
3. Considerar $v = [1 + u^{-\theta}(t^{-\theta/(\theta+1)} - 1)]^{-1/\theta}$.
4. La pareja buscada es (u, v) .

Presentamos la gráfica de C compuesta con distintas marginales.

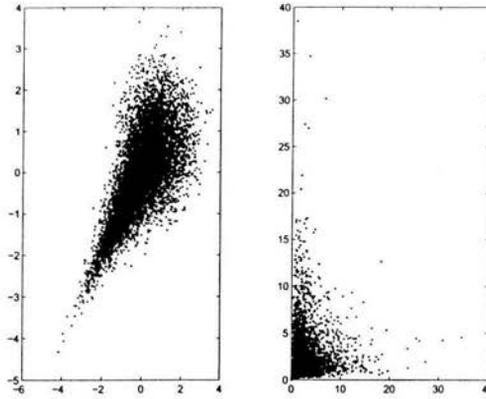


Figura 4.3: 10,000 simulaciones de C compuesta con distintas marginales. A la izquierda marginales normales estándar, a la derecha lognormales estándar

4.3 Cópulas Empíricas

Dada una función de distribución de p -variables, sabemos por el teorema de Sklar, expresado mediante (1.15), que existe C una cópula tal que

$$\forall \bar{x} \in \mathfrak{R}^p \quad F(x_1, \dots, x_p) = C(F_1(x_1), \dots, F_p(x_p))$$

La idea de las cópulas empíricas es aproximar C de un modo no-paramétrico. Para ello, consideremos una muestra aleatoria $(\hat{X}_1, \hat{X}_2, \dots, \hat{X}_n)$ de un vector $\hat{X} = (X(1), X(2), \dots, X(p))$ con función de distribución F y funciones marginales $\{F_j, 1 \leq j \leq p\}$ continuas. Definamos lo siguiente:

1. Los estadísticos de orden de la muestra son

$$X_{(1)}(j) < X_{(2)}(j) < \dots < X_{(n)}(j) \quad \text{para cada } 1 \leq j \leq p \quad (4.2)$$

2. La permutación $(X_1(j), \dots, X_n(j)) \xrightarrow{\pi_j} (X_{(1)}(j), \dots, X_{(n)}(j))$ de asignaciones $X_i(j) \mapsto X_{(\pi_j(i))}(j)$ será llamada j -ésima función de rangos de la muestra. El rango de $X_i(j)$ será $\pi_j(i)$ y se denotará por $r_i(j)$. En palabras sencillas el rango de la variable $X_i(j)$ es el lugar que ocupa en la muestra ordenada (4.2).

3. La función de distribución empírica de la muestra es

$$F^{(n)}(\bar{x}) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i(1) \leq x_1, \dots, X_i(p) \leq x_p) \quad (4.3)$$

Para cada $\omega \in \Omega$. Sus marginales serán denotadas como $F_j^{(n)}$, para cada $1 \leq j \leq p$.

Para cada $\omega \in \Omega$, gracias al teorema de Sklar aplicado a la función de distribución empírica $F^{(n)}(\omega)$, podemos afirmar que existe al menos una cópula C_n tal que

$$\forall \bar{x} \in \mathfrak{R}^p \quad F^{(n)}(\bar{x}) = C_n(F_1^{(n)}(x_1), \dots, F_p^{(n)}(x_p)) \quad (4.4)$$

Definición 12. A cualquier función C_n que satisfaga (4.4) se le denomina *cópula empírica asociada a $\hat{X}_1, \dots, \hat{X}_n$* .

Nótese que esta función está determinada de modo único sobre el conjunto

$$I_n := \left\{ \frac{k_1}{n}, \frac{k_2}{n}, \dots, \frac{k_p}{n} \right\} \quad 1 \leq k_j \leq n, 1 \leq j \leq p$$

Por el corolario 4 al teorema de Sklar, sabemos que

$$C_n(u_1, \dots, u_p) = C([F_1^{(n)}]^{-1}(u_1), \dots, [F_p^{(n)}]^{-1}(u_p))$$

donde $[F_j^{(n)}]^{-1}$ denota la pseudo-inversa de la j -ésima función de distribución marginal empírica. Recordemos que, en términos de los estadísticos de orden, si $x \in \mathfrak{R}$, entonces

$$F_j^{(n)}(x) = \begin{cases} 0, & \text{si } x < X_{(1)}(j), \\ k/n, & \text{si } X_{(k)}(j) \leq x < X_{(k+1)}(j), \\ 1, & \text{si } X_{(n)}(j) \leq x. \end{cases}$$

Por lo cuál su pseudo-inversa es

$$[F_j^{(n)}]^{-1}(k/n) = X_{(k)}(j)$$

Componiendo con (4.3) obtenemos que

$$C_n(k_1/n, \dots, k_p/n) = \begin{cases} \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_1(j) \leq X_{(k_1)}(j), \dots, X_p(j) \leq X_{(k_p)}(j)) & k_j \neq 0 \\ 0 & \text{e.o.c.} \end{cases}$$

lo cual es equivalente a la expresión en términos de los rangos

$$C_n \left(\frac{k_1}{n}, \dots, \frac{k_p}{n} \right) = \begin{cases} \frac{1}{n} \sum_{i=1}^n \mathbb{I}(r_i(1) \leq k_1, \dots, r_i(p) \leq k_p), & \text{si } \forall j, k_j \neq 0, \\ 0, & \text{si } k_j = 0 \text{ p.a. } j. \end{cases} \quad (4.5)$$

Teorema 4.1. *La cópula C_n definida por (4.5) es un estimador suficiente para C .*

Demostración. Obsérvese que

1. La distribución de C_n no depende de F_1, F_2, \dots, F_p , por (4.5),
2. La distribución de $F_1^{(n)}, \dots, F_p^{(n)}$ no depende de C , pues por definición depende sólo de las marginales F_j de F

Así dada C_n sobre I_n y las marginales $\{F_j^{(n)}, 1 \leq j \leq p\}$, se determina $F^{(n)}$ de modo biunívoco. ■

Para hablar de consistencia del estimador, tomemos C_n una cópula empírica cualquiera para C en $[0, 1]^p$, definida en I_n según (4.5). Ya que la distribución de C_n depende sólo de C , podemos en adelante suponer que las funciones de distribución marginales son uniformes y, con ello, que $F = C$.

Teorema 4.2. *Cualquier estimador C_n definido en I_n mediante (4.5) converge uniformemente a C , es decir,*

$$\lim_{n \rightarrow \infty} \sup_{\bar{u} \in [0,1]^p} |C_n(\bar{u}) - C(\bar{u})| = 0 \quad \text{c.s.}$$

Demostración. Utilicemos primero la versión p -dimensional del lema 2 para afirmar que para cualquier cópula A

$$\forall 0 \leq \bar{u}_1, \bar{u}_2 \leq 1, \quad |A(\bar{u}_1) - A(\bar{u}_2)| \leq \sum_{j=1}^p |u_1(j) - u_2(j)| \quad (4.6)$$

Tomemos ahora $\bar{u} \in [0, 1]^p$. Por definición de I_n existe un punto $\bar{\theta} \in I_n$ tal que $|u_j - \theta_j| \leq 1/n \quad \forall 1 \leq j \leq p$. Es así que

$$\begin{aligned} |C_n(\bar{u}) - C(\bar{u})| - |C_n(\bar{\theta}) - C(\bar{\theta})| &\leq |C_n(\bar{u}) - C_n(\bar{\theta})| + |C(\bar{u}) - C(\bar{\theta})| \\ &\leq 2 \sum_{j=1}^p |u_j - \theta_j| \leq \frac{2p}{n} \end{aligned}$$

Esto significa que

$$\sup_{\bar{u} \in [0,1]^p} |C_n(\bar{u}) - C(\bar{u})| - \sup_{\bar{u} \in I_n} |C_n(\bar{u}) - C(\bar{u})| \leq \frac{2p}{n} \quad (4.7)$$

de modo que basta examinar la convergencia uniforme sobre I_n .

Por las definiciones de I_n y de las funciones de distribución empíricas marginales, tenemos que

$$\begin{aligned} \sup_{\bar{u} \in I_n} |C_n(\bar{u}) - C(\bar{u})| = \\ \sup_{\bar{u} \in [0,1]^p} |C_n(C_1^{(n)}(u_1), \dots, C_p^{(n)}(u_p)) - C(C_1^{(n)}(u_1), \dots, C_p^{(n)}(u_p))| \end{aligned}$$

Obsérvese que por la desigualdad del triángulo y el lema 2 tenemos

$$\begin{aligned} & \sup_{\bar{u} \in [0,1]^p} |C_n(C_1^{(n)}(u_1), \dots, C_p^{(n)}(u_p)) - C(C_1^{(n)}(u_1), \dots, C_p^{(n)}(u_p))| \\ \leq & \sup_{\bar{u} \in [0,1]^p} |C_n(C_1^{(n)}(u_1), \dots, C_p^{(n)}(u_p)) - C(\bar{u})| + \sup_{\bar{u} \in [0,1]^p} |C(\bar{u}) - C(C_1^{(n)}(u_1), \dots, C_p^{(n)}(u_p))| \\ \leq & \sup_{\bar{u} \in [0,1]^p} |C_n(C_1^{(n)}(u_1), \dots, C_p^{(n)}(u_p)) - C(\bar{u})| + \sum_{j=1}^p \sup_{0 \leq u \leq 1} |u - C_j^{(n)}(u)| \end{aligned} \quad (4.8)$$

Sabemos, por el Teorema de Glivenko–Cantelli, cuya prueba se puede encontrar en [5], que

$$\sup_{\bar{u} \in [0,1]^p} |C^{(n)}(\bar{u}) - C(\bar{u})| \xrightarrow{n \rightarrow \infty} 0, \quad c.s.$$

donde $C^{(n)}$ es la distribución empírica asociada con C según (4.3) y también que, $C_j^{(n)}$ converge de modo uniforme hacia C_j , es decir,

$$\forall j, \quad \sup_{0 \leq u \leq 1} |C_j^{(n)}(u_j) - (u_j)| \xrightarrow{n \rightarrow \infty} 0, \quad c.s.$$

De modo que tomando límites en (4.8)

$$\lim_{n \rightarrow \infty} \sup_{\bar{u} \in [0,1]^p} |C_n(C_1^{(n)}(u_1), \dots, C_p^{(n)}(u_p)) - C(C_1^{(n)}(u_1), \dots, C_p^{(n)}(u_p))| = 0 \quad c.s.$$

con lo cual

$$\lim_{n \rightarrow \infty} \sup_{\bar{u} \in I_n} |C_n(\bar{u}) - C(\bar{u})| = 0 \quad c.s.$$

completando la prueba. ■

Como un caso particular de esta convergencia de C_n hacia C , la cópula del vector \bar{X} , tenemos que si las coordenadas $X(1), \dots, X(p)$ son independientes entonces C_n converge hacia Π . Pensando en esto podemos proponer la prueba de independencia basada en el estadístico

$$T_n = \sup_{[0,1]^p} |C_n(\bar{u}) - \Pi(\bar{u})|$$

en la que, teniendo una muestra del vector \bar{X}

1. Calcularemos la distribución de T_n ,
2. evaluaremos el estadístico T_n para la muestra.
3. Rechazaremos independencia de las coordenadas $X(1), \dots, X(p)$ siempre que el valor obtenido antes supere el cuantil $1 - \alpha$ de la distribución de T_n (α es el nivel de significancia).

El cálculo de la distribución de T_n puede complicarse bastante en función tanto del número de coordenadas de \bar{X} como del tamaño de la muestra. Una opción es basarnos en Glivenko–Cantelli para en vez de calcular de modo exacto aproximar la distribución de T_n mediante simulaciones. Gracias a que C_n no depende de las marginales (ni depende de ella la convergencia que da sentido a la prueba) la simulación puede hacerse con variables aleatorias uniformes. La idea es sencilla, se genera “una gran cantidad” de evaluaciones de T_n y con ellas se obtiene un aproximador de su función de distribución cuyos cuantiles usaremos como si fuesen los reales.

Para el caso bivariado podemos utilizar el siguiente programa de Matlab para generar las evaluaciones de T_n y con ellas encontrar el cuantil de orden “ $0 < c < 1$ ”.

```
function c=cuantilTn(c,n,m)
a=zeros(n,1);
```

```

ind=(1:m)^(1:m);
for k=1:n
.   y=rand(m,1);
.   yo=sort(y);
.   dep=zeros(m,m);
.   temp=0.^(abs(y-yo(1)*ones(m,1)));
.   temp=tril(ones(m,m))*temp;
.   dep(1,:)=temp';
.   for i=2:m
.       temp=0.^(abs(y-yo(i)*ones(m,1)));
.       temp=tril(ones(m,m))*temp;
.       dep(i,:)=dep(i-1,:)+temp';
.   end
.   a(k)=max(max(abs(m*dep-ind)))/(m^2);
end
A=(1:m)/m;
c=min(find(A>=c));
ao=sort(a);
c=ao(c);

```

Después, para evaluar el estadístico T_n según la muestra podemos utilizar el siguiente programa (también para Matlab)

```

function d=distaCnPi(x,y)
% x y y son vectores columna de las parejas de datos
m=sortrows([x y]);
n=length(x);
y=sort(y);
x=sort(x);
dep=zeros(n,n);
temp=0.^(abs(y-m(1,2)*ones(n,1)));
t=min(find(temp==1));
temp=[zeros(t-1,1);ones(n-t+1,1)];
dep(1,:)=temp';
for i=2:n

```

```

.   temp=0.^(abs(y-m(i,2)*ones(n,1)));
.   t=min(find(temp==1));
.   temp=[zeros(t-1,1);ones(n-t+1,1)];
.   dep(i,:)=dep(i-1,:)+temp';
end
h=find(x(2:n)-x(1:n-1)==0);
h=flipud(h);
for i=h'
.   dep(i,:)=dep(i+1,:);
end
ind=(1:n)'*(1:n);
d=max(max(abs(n*dep-ind)))/(n^2);

```

Para explorar un poco esta prueba, tomemos como muestra la simulación de un proceso $\{Z_n\}$ formado a partir de un proceso independiente $\{X_n\}$ como sigue

$$\begin{aligned}
 Z_1 &= X_1 \\
 Z_2 &= \frac{X_1 + X_2}{2} \\
 Z_n &= \frac{X_{n-2} + X_{n-1} + X_n}{3}
 \end{aligned}$$

Esta construcción implica que $|n - m| \geq 3$ si y sólo si $Z_n \perp Z_m$. Podemos probar la independencia para Z_k, Z_{k+r} con $r < 3$ y examinar los resultados. Para ello, tomaremos muestras de la forma $(Z_1, Z_{1+r}), (Z_4, Z_{4+r}), \dots, (Z_{1+3n}, Z_{1+3n+r})$ para asegurar que son independientes las observaciones. Para simular el proceso y obtener la muestra indicada se utilizarán las siguientes rutinas

```

function z=proceso(n)
z=zeros(n,1);
x=rand(n,1);
x=-log(x);% x son n v.a.i. distribuidas exp(1)
z(1)=x(1);

```

```

z(2)=(x(2)+x(1))/2;
for k=3:n
.   z(k)=(x(k-2)+x(k-1)+x(k))/3;
end

```

```

function A=muestraZn(z,c,r,t)
%obtiene las muestras (Z_k, Z_{k+r}), (Z_{k+r+t}, Z_{k+t+2r})
%comenzando en (Z_c, Z_{c+r})
n=length(z);
x(1)=z(c);
y(1)=z(c+r);
p=floor((n-1-c+t)/(r+t));
for k=2:p
.   x(k)=z(c+(t+r)*(k-1));
.   y(k)=z(c+(t+r)*(k-1)+r);
end
A=[x' y'];

```

Obtenemos los cuantiles para T_n simulando 1000 observaciones de T_n para 100 variables independientes cada una y utilizamos una muestra del proceso Z_n de 5000 variables.

Orden del cuantil	Valor del cuantil
0.90	.0744
0.95	.0815

Tabla 4.1: Cuantiles para T_n

Por lo tanto, se rechaza la independencia en el primer caso con cualquier cuantil, en el segundo con el de .90.

Valor r	Evaluación de T_n
1	.1293
2	.0790

Tabla 4.2: Evaluaciones de T_n

Aplicando la prueba para $r = 3$, se obtiene que la evaluación del estadístico es $T_n = .0265$ dejando en claro la independencia (que se conocía por construcción).

Efectuemos ahora una prueba con datos reales. Contamos con los precios del tipo de cambio peso-dólar, fix desde 1995 hasta 2002. Estos precios tienen una clara dependencia de los días anteriores como lo muestra la gráfica. Trabajaremos entonces con la variación porcentual del fix definida



Figura 4.4: Precios del fix. 1995 a 2002

como sigue: Si X_n denota el valor del fix al día n , entonces la variación Y_n

para este día será

$$Y_n = \frac{X_n - X_{n-1}}{X_{n-1}}$$

En este caso no se tiene esta notoria dependencia como observamos en la gráfica. Busquemos rechazar independencia en parejas del estilo (Y_n, Y_{n+k}) .

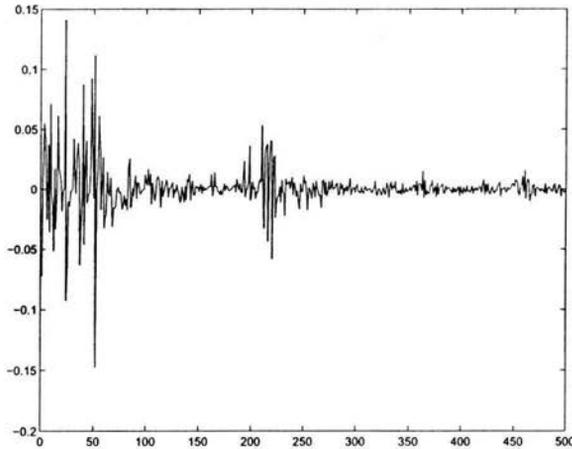


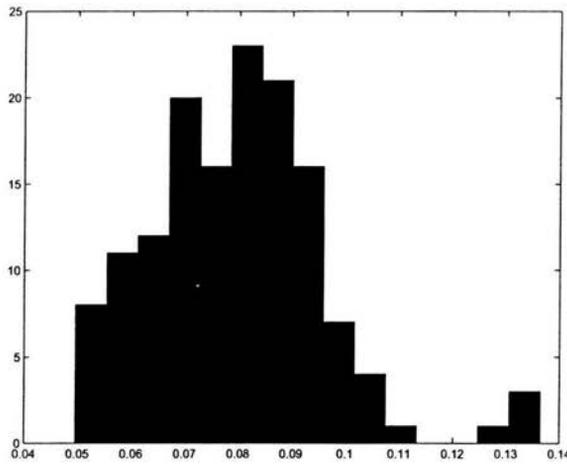
Figura 4.5: Variación porcentual del fix. Gráfica de 500 datos

Para ello, supongamos que la variación porcentual pierde memoria tras 30 días o menos; es decir que para cada natural n , Y_n es independiente de Y_{n+30} . Tomaremos, pues, muestras de la forma (Y_n, Y_{n+k}) , $(Y_{n+k+30}, Y_{n+2k+30})$. Usaremos para ello la rutina de Matlab que usamos en la sección anterior.

Elijamos $k = 7$. Tenemos ya los cuantiles para T_n .

Debido al método de selección de las parejas, se pierde mucha información muestral. Esto nos da la opción de comenzar la selección no en el primero sino en el segundo día o más allá. En un intento por utilizar toda la información disponible, se ha evaluado el estadístico T_n en todas las muestras de la forma antes mencionada siempre que éstas tengan tamaño no menor a 50 elementos. La conclusión a que se llega es que el 66% de las veces se excede el cuantil de .90 y el 48% el de 95. Presentamos el histograma de evaluaciones para el estadístico T_n .

Orden del cuantil	Valor del cuantil
0.90	.0744
0.95	.0815

Tabla 4.3: Cuantiles para T_n Figura 4.6: Histograma de las evaluaciones de T_n

Es así que rechazamos la independencia de las parejas $(Y_n, Y_{(n+7)})$ suponiendo pérdida de memoria en 30 días o menos. En la siguiente sección ajustaremos una cópula arquimediana a estas parejas.

4.4 Ajuste de cópulas arquimedianas paramétricas

En un contexto general puede ser difícil el encontrar una cópula que ajuste a un conjunto de datos, no obstante, podemos facilitar la labor si suponemos

que la cópula que buscamos es arquimediana. En este caso, la proposición 13 nos dice que aproximando la distribución de la variable aleatoria $C(U, V)$ podemos aproximar el generador de C . Un método es el siguiente.

Supongamos que contamos con una muestra aleatoria $(X_1, Y_1), \dots, (X_n, Y_n)$ correspondiente al vector (X, Y) de cópula C y marginales continuas F_X, F_Y . Supongamos también que C es arquimediana. En particular, y sin perder generalidad dado que el método de ajuste se centrará en la cópula de modo independiente de las marginales, pensemos que F_X y F_Y son las distribuciones uniformes en $[0, 1]$, de modo que la distribución de (X, Y) será C . Como muestra la proposición 13, el comportamiento estocástico de $V = C(X, Y)$ caracteriza al generador de C y, con ello, a C . Buscamos pues un estimador para la distribución $K(v) = \mathbb{P}(C(X, Y) \leq v)$. Esto puede llevarse a cabo en dos pasos

1. Primero, construimos la distribución empírica bivariada $C^{(n)}(x, y)$ asociada con C del modo usual, i.e.

$$C^{(n)}(x, y) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x, Y_i \leq y)$$

2. Calculamos $C^{(n)}(X_i, Y_i)$ y con estos valores un estimador de K como

$$K_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(C^{(n)}(X_i, Y_i) \leq t)$$

Uno puede prescindir del estimador $C^{(n)}$ notando que las cantidades $C^{(n)}(X_i, Y_i)$ representan la proporción de las observaciones muestrales que no exceden (X_i, Y_i) componente a componente y utilizar las variables

$$A_j = \frac{1}{n} \sum_{j=1}^n \mathbb{1}(X_j \leq X_i, Y_j \leq Y_i)$$

para generar el estimador de $K(t)$. Sin embargo, como $C^{(n)}(X_i, Y_i)$ siempre es mayor que $1/n$ y es deseable tener un estimador que tome valores en $(0, 1)$, será más conveniente el utilizar las variables

$$B_j = \frac{1}{n-1} \sum_{j \neq i} \mathbb{1}(X_j \leq X_i, Y_j \leq Y_i)$$

Si denotamos por $\delta(x) = \mathbb{1}(x > 0)$ entonces, un estimador para $K(t)$ está dado por

$$K_n(t) = \frac{1}{n} \sum_{i=1}^n \delta(t - B_i) \quad (4.9)$$

Nótese que con ello somos capaces ya de producir un estimador de la cópula C gracias a la relación (2.16) definiendo $\lambda_n(t) = t - K_n(t)$ para cada $n \in \mathbb{N}$ y $t \in (0, 1)$. Siempre que $K_n(t^-) > t$ en su dominio, el estimador así generado será una cópula arquimediana. Obsérvese que esto es cierto aún cuando no supongamos que estamos estimando una cópula arquimediana. Es así que con este método construimos una cópula arquimediana a partir de la muestra. Siendo así, podemos pensar que la cópula construida C_n es una aproximación arquimediana a la cópula real C . Examinemos cómo llevar este método a la práctica.

Primero necesitamos estimar el generador

$$\phi_n(t) = \int_{t_0}^t \frac{1}{\lambda_n(x)} dx$$

Después necesitamos invertirlo para tener

$$C_n(u, v) = \phi_n^{-1}(\phi_n(u) + \phi_n(v))$$

Este paso, la inversión del generador ϕ_n , es computacionalmente muy costosa pues el único procedimiento para hallar ϕ_n^{-1} es usando el teorema de la función implícita que nos da una expresión para ella en términos locales y además como una serie. Esta es una seria limitante del método.

Un procedimiento alternativo surge al notar que el estimador propuesto para $K(t)$ guarda estrecha relación con la τ de Kendall.

Recordemos que la τ muestral es la diferencia entre la probabilidad de concordancia y la de discordancia, es decir

$$\tau_n = \frac{\#\text{pares concordantes} - \#\text{pares discordantes}}{\binom{n}{2}}$$

Rigurosamente hablando, podemos considerar las variables

$$I_{ij} = \begin{cases} 1 & \text{si } X_j < X_i \text{ y } Y_j < Y_i, \\ 0 & \text{en caso contrario.} \end{cases}$$

para cada $1 \leq i, j \leq n$. Con estas variables, dos parejas $(X_i, Y_i), (X_j, Y_j)$ son concordantes si y sólo si $I_{ij} + I_{ji} = 1$, de modo que el número total de parejas concordantes es $\sum_{ij} I_{ij}$. Tomando en cuenta que $\#$ discordantes = $\binom{n}{2}$ - $\#$ concordantes podemos deducir que

$$\tau_n = \frac{2 \sum_{ij} I_{ij}}{\binom{n}{2}} - 1$$

Gracias a que $B_i = (n-1) \sum_j I_{ij}$, se sigue que $\sum_{ij} I_{ij} = (n-1) \sum_i B_i$, y con ello

$$\tau_n = \frac{2(n-1) \sum_i B_i}{\binom{n}{2}} - 1$$

lo cual puede simplificarse como

$$\tau_n = 4\bar{B} - 1$$

donde $\bar{B} = \frac{1}{n} \sum_i B_i$. Podemos explicarnos esta relación como la versión muestral de la proposición 12. De este modo, si suponemos que la cópula que deseamos estimar es arquimediana y paramétrica podemos escoger el parámetro pidiendo que τ_n coincida con (o esté un una banda de confianza alrededor de) el verdadero valor de τ . Es importante notar que este método sí necesita el supuesto de que busquemos aproximar una cópula arquimediana.

Examinemos ahora las propiedades del estimador $K_n(t)$. Primeramente, probaremos que converge a $K(t)$ para cada valor de t , después daremos una aproximación $o(1/n)$ para su variancia asintótica.

Proposición 17. *Si B_i son idénticamente distribuidas, entonces la distribución $K_n(t)$ converge de modo puntual hacia $K(t)$ para todo valor de t .*

Demostración. Si condicionamos a que $(X_1, Y_1) = (x_1, y_1)$ entonces, la variable $(n-1)V_1$ sigue una distribución binomial de parámetros $(n-1, C(x_1, y_1))$ por la independencia de la muestra. Esto significa que

$$\mathbb{E}(e^{tB_1} | X_1 = x_1, Y_1 = y_1) = (1 - C(x_1, y_1) + C(x_1, y_1)e^{t/(n-1)})^{(n-1)}$$

de modo que,

$$\mathbb{E}(e^{tB_1}) = \mathbb{E} \left[(1 - C(X_1, Y_1) + C(X_1, Y_1)e^{t/(n-1)})^{(n-1)} \right]$$

El lado derecho de esta igualdad puede ser escrito como

$$\mathbb{E} \left(1 + \frac{(n-1)C(X_1, Y_1) [e^{t/(n-1)} - 1]}{n-1} \right)^{(n-1)}$$

Argumentemos ahora que

$$\lim_{n \rightarrow \infty} \left(1 + \frac{(n-1)C(X_1, Y_1) [e^{t/(n-1)} - 1]}{n-1} \right)^{(n-1)} = e^{tC(X_1, Y_1)} \quad (4.10)$$

Utilizando la regla de L'Hopital sobre la función

$$f(x) = \frac{e^{t/x} - 1}{1/x}$$

observamos que conforme $x \rightarrow \infty$, $f(x) \rightarrow t$ por lo cual

$$\lim_{n \rightarrow \infty} (n-1)C(X_1, Y_1) [e^{t/(n-1)} - 1] = tC(X_1, Y_1)$$

Esto prueba (4.10), puesto que $(1 - x/n)^n \rightarrow e^{-x}$ de modo uniforme sobre intervalos acotados.

Utilizando ahora el teorema de la convergencia acotada de Lebesgue se sigue que

$$\lim_{n \rightarrow \infty} \mathbb{E}(e^{tB_1}) = \mathbb{E}(e^{tC(X_1, Y_1)})$$

lo cual demuestra que

$$\lim_{n \rightarrow \infty} \mathbb{P}(B_1 \leq v) = \mathbb{P}(C(X_1, Y_1) \leq v).$$

Resta notar que $K_n(t)$ es la distribución de B_1 . ■

Proposición 18. *Si las variables B_i son idénticamente distribuidas, entonces la función de distribución empírica de las B_i , $K_n(t)$, es un estimador \sqrt{n} -consistente de $K(t)$, y una aproximación $o(1/n)$ para la varianza asintótica de $K_n(t)$ está dada por*

$$[K(t)(1 - K(t)) + k(t)\{k(t)R(t) - 2t(1 - K(t))\}]/n$$

donde $k(t) = K'(t)$ es la densidad asociada con $K(t)$, misma que existe casi seguramente y

$$R(t) = \mathbb{E} \left(C(\min(X_1, X_2), \min(Y_1, Y_2)) - t^2 \middle| C(X_1, Y_1) = C(X_2, Y_2) = t \right) \quad (4.11)$$

Demostración. La prueba de este hecho se puede encontrar en ??.

Gracias a que la cópula C es arquimediana, podemos conocer la forma de $R(t)$ definido en (4.11) como sigue.

Proposición 19. Si C es una cópula arquimediana de generador ϕ , entonces

$$R(t) = 2 \int_0^1 (1-v)\phi^{[-1]} \{(1+v)\phi(t)\} dv - t^2 \quad (4.12)$$

Demostración. Para las cópulas arquimedianas la condición $C(X_1, Y_1) = t$ puede escribirse como $\phi(X_1) + \phi(Y_1) = \phi(t)$.

Por ello, dado que $C(X_1, Y_1) = C(X_2, Y_2) = t$, si $X_1 = \min(X_1, X_2)$ entonces $Y_2 = \min(Y_1, Y_2)$. Resultado de esta simetría

$$\begin{aligned} & \mathbb{E} \left(C(\min(X_1, X_2), \min(Y_1, Y_2)) \middle| C(X_1, Y_1) = C(X_2, Y_2) = t \right) \\ &= 2\mathbb{E} \left(C(X_1, Y_2) \mathbb{1}((X_2 - X_1) > 0) \middle| C(X_1, Y_1) = C(X_2, Y_2) = t \right) \end{aligned} \quad (4.13)$$

Considerando que $\phi(t) = \phi(X_2) + \phi(Y_2)$, si tomamos las variables $S_1 = \phi(X_1)/\phi(t)$ y $S_2 = \phi(X_2)/\phi(t)$, podemos escribir

$$C(X_1, Y_2) = \phi^{[-1]}(\phi(t)S_1 - \phi(t) + \phi(t)S_2) \quad (4.14)$$

Sabemos por el teorema 2.5 que si (U, V) tienen cópula C , entonces la variable

$$S = \frac{\phi(U)}{\phi(U) + \phi(V)}$$

sigue una distribución uniforme en $(0, 1)$. De este modo, condicionado a que $C(X_1, Y_1) = C(X_2, Y_2) = \phi(t)$ las variables S_1 y S_2 tienen por distribución la uniforme en $(0, 1)$. Es así que la ecuación (4.13) se reduce a

$$2 \int_0^1 \int_0^{u_1} \phi^{[-1]}(\phi(t)u_1 + \phi(t) - \phi(t)u_2) du_2 du_1.$$

Si definimos $v = u_1 - u_2$ en la segunda integral obtenemos

$$2 \int_0^1 \int_0^{u_1} \phi^{[-1]}(u_1\phi(t) + \phi(t) - u_2\phi(t)) dv du_1$$

Cambiando el orden de la integral tenemos

$$2 \int_0^1 \int_{u_2}^1 \phi^{[-1]}(v\phi(t) + \phi(t)) du_1 dv$$

Resolviendo esta integral se termina la prueba. ■

Por ejemplo, para la cópula de variables aleatorias independientes, Π , generada por $\phi(t) = \log(1/t)$ tenemos que

$$R(t) = 2 \int_0^1 (1 - v) \exp((1 + v) \log(t)) dv - t^2$$

Para resolver la integral podemos separarla en las integrales

$$\int_0^1 \exp((1 + v) \log(t)) dv \quad y \quad \int_0^1 v \exp((1 + v) \log(t)) dv$$

La primera se resuelve directamente mediante la substitución $u = (1 + v) \log(t)$, y la segunda por partes pidiendo $f = t, g' = \exp((1 + v) \log(t))$ y usando lo anterior. De este modo, tras simplificar, obtenemos

$$R(t) = \frac{2t(t - \log(t) - 1)}{(\log(t))^2} - t^2$$

Con esto podemos encontrar la forma del aproximador $o(1/n)$ de la varianza para $K_n(t)$ siendo este igual a $t(t - \log(t) - 1)/n$.

Efectuemos un ajuste con este método. En la sección anterior probamos la hipótesis de independencia para parejas (Y_n, Y_{n+7}) de observaciones de la variación porcentual del precio del fix y la rechazamos. Esto da sentido a la búsqueda de una cópula para estas parejas. Supondremos que es arquimediana y además paramétrica. Consideraremos como posibles las familias Clayton, Genest y Gumbel-Hougaard.

Familia	$\phi(t)$	Rango de α	$\lambda(t) = \phi(t)/\phi'(t)$	τ
Clayton	$(t^{-\alpha} - 1)/\alpha$	$(0, \infty)$	$t(1 - t^\alpha)/\alpha$	$\alpha/(\alpha + 2)$
Genest	$(1 - t^{1/\alpha})^\alpha$	$[1, \infty)$	$-t^{(\alpha-1)/\alpha}(1 - t^{1/\alpha})$	$(4\alpha - 6)/(4\alpha - 2)$
Gumbel-Hougaard	$[-\log(t)]^{\alpha+1}$	$[0, \infty)$	$-t \log(t)/(\alpha + 1)$	$\alpha/(\alpha + 1)$

Tabla 4.4: Familias Clayton, Genest y Gumbel-Hougaard de cópulas arquimedianas.

Usaremos para el análisis las siguientes rutinas.

```
function t=taum(x,y)
n=length(x);
b=zeros(n,1);
for i=1:n
.   b(i)=(length(find(x<x(i)& y<y(i))))/(n-1);
end
t=4*(sum(b)/n)-1;
```

```
function k=distCuv(x,y,t);
n=length(x);
b=zeros(n,1);
for i=1:n
.   b(i)=(length(find(x<=x(i)& y<=y(i))))/(n-1);
end
for i=1:length(t)
.   k(i)=length(find(b<=t(i)));
end
k=k/n;
```

La primera para calcular el valor de la τ muestral y la segunda para evaluar el estimador $K_n(t)$ de la distribución de $C(U, V)$. Obtenemos que $\tau = -.0847$. A través de calcular $\|\lambda_n - \lambda\| = \sup_{t \in [0,1]} |\lambda_n(t) - \lambda(t)|$ en cada caso llegamos a la decisión de que la mejor elección será la cópula Gumbel-Hougaard de parámetro $\alpha = -.0781$. Presentamos un resumen de los resultados.

Familia	Valor del parámetro
Clayton	-.1563
Genest	1.6731
Gumbel-Hougaard	-0.0781

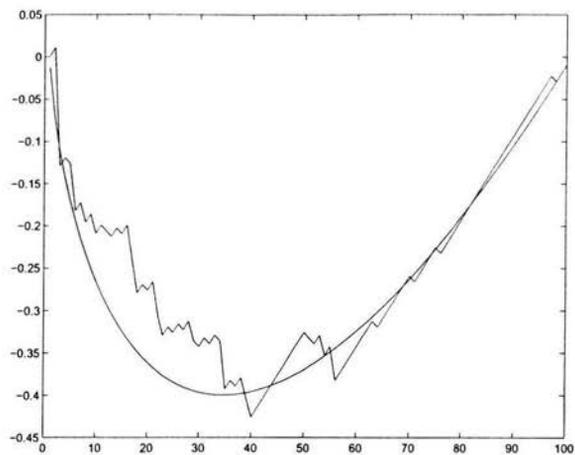
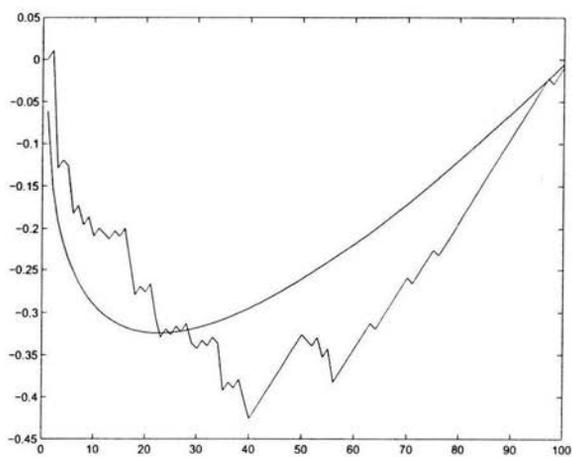
Tabla 4.5: Parámetros de las distintas familias.

Hemos de notar que la familia Gumbel-Hougaard contiene a la

Familia	$\ \lambda_n - \lambda\ $
Clayton	.1331
Genest	.1629
Gumbel-Hougaard	.1107

Tabla 4.6: Comparación del ajuste de λ_n .

cópula Π , que corresponde al parámetro $\alpha = 0$. El estimador que obtuvimos, $-.0781$ es verdaderamente pequeño comparando con el rango posible, $[0, \infty)$. La cópula que hemos obtenido, pues, debe ser muy similar a la cópula de variables aleatorias independientes. Recordemos que, si bien es cierto que en la sección anterior rechazamos la independencia de las parejas (Y_n, Y_{n+7}) en la mayoría de los casos, también es cierto que en una gran proporción de las muestras (el 44%) no fue posible hacerlo. Este ajuste nos explica que lo que sucede es que la cópula de estas variables es muy similar a la cópula Π .

Figura 4.7: λ de la familia Clayton contra λ_n Figura 4.8: λ de la familia Genest contra λ_n

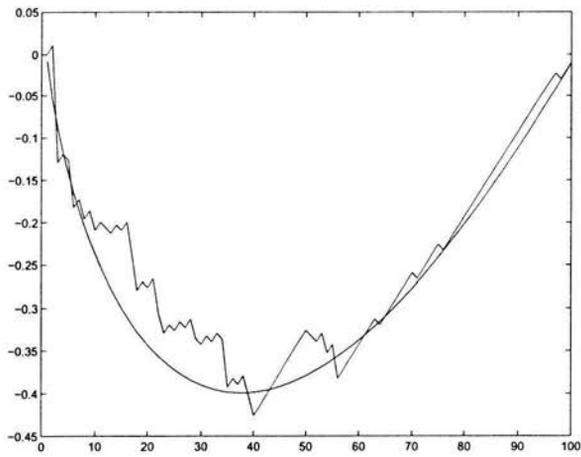


Figura 4.9: λ de la familia Gumbel-Hougaard contra λ_n

4.5 Conclusiones

En el terreno de lo empírico las cópulas nos revelan información interesante. Nos dejan ver que los aspectos de la relación que dependen puramente de la cópula pueden pasar de largo ciertos “detalles”. Por ejemplo, tenemos los reordenamientos, en que siendo determinísticamente dependientes las variables, se puede reportar independencia. No obstante hemos de reconocer que si un reordenamiento se semeja mucho a la cópula producto, entonces sus observaciones pueden ser utilizadas como observaciones independientes. Lo que se concluye es que no hay un claro límite conocido entre la independencia y la dependencia ... en cada prueba que se concluya la independencia, tal vez se estaría concluyendo con la misma razón la existencia de una función invertible que relacione las variables en cuestión, y así, un nexo determinista entre ellas.

Claro está que la prueba de la cópula empírica fallaría en un caso como este, sin embargo, no es el único caso en que puede fallar. Se pueden construir ejemplos de variables aleatorias dependientes cuya cópula semeja tanto a la de la independencia que en la mayoría de los casos la prueba basada en T_n sea incapaz de rechazar. Tal es el caso siguiente: Tómnese X, Y, Z v.a. independientes con distribución $N(0, 1)$, y tómnese $W = X/Y, U = Y/Z$. Calculando su densidad conjunta se ve que no son independientes (pues difiere del producto de dos Cauchy estándar como corresponderá). Generando mediante simulación muestras aleatorias de 1000 observaciones de estas variables $-W, U-$ y evaluando T_n se tiene que el 95% de las veces no podemos rechazar la independencia de (W, U) . En este caso, la independencia condicional de (W, U) dado Y parece jugar un papel importante en la incapacidad de la prueba. Curiosamente, cuando se realiza la misma prueba con distribuciones normales no simétricas –por ejemplo con media $\mu = 1$ – la prueba es muy eficiente.

Ciertamente, en caso de rechazo de la hipótesis de independencia, la prueba de la cópula empírica es muy eficiente, sin embargo, su naturaleza, por decirlo de algún modo, universal, hace que en muchos casos no sea suficiente para detectar dependencia.

Por otro lado, el método de ajuste de cópulas arquimedianas paramétricas nos da la posibilidad de utilizar las cópulas como nexo entre dos distribuciones univariadas obtenidas a partir de una muestra para producir posibles distribuciones bivariadas de esta misma muestra. De este modo, este método trae a la práctica la riqueza de las cópulas.

Conclusiones Generales

Hemos visto cómo la función de distribución de un vector aleatorio está compuesta por información relativa exclusivamente a las variables, información marginal, y está compuesta también por información relativa al comportamiento conjunto de las variables aleatorias. Esta situación puede ser confusa, pues puede dar la idea de que la definición de independencia o de dependencia es una suerte de esencia que no guarda relación con las distribuciones marginales. Hemos notado que no es así. Por ejemplo, si un vector está formado por una variable aleatoria X y la suma $X + Y$ con Y independiente de X , entonces la cópula asociada a este vector cambia radicalmente al elegir las marginales para las variables X, Y . De este modo, aunque la cópula esté íntimamente relacionada con la dependencia de un vector aleatorio, no siempre es posible extraer información útil o relevante de este hecho.

Por otra parte, en los problemas de datos reales, para ajustar una cópula podríamos utilizar las cópulas empíricas. El problema de esta aproximación es que calcular la cópula empírica tiene complicaciones numéricas importantes. Sería costoso, en términos computacionales, generar las evaluaciones de la cópula empírica en, digamos, 100 puntos (¡que sería lo correspondiente a partir el intervalo unitario en sólo 10 segmentos!), y mucho más costoso aún el interpolar con el método presentado en la prueba de Sklar para extender esta sub-cópula al cuadrado unitario. No sería práctico. Por ello el método de las cópulas Arquimedianas me parece más adecuado.

Al respecto de este método podemos plantear, con mucha razón, la interrogante de ¿por qué utilizar cópulas Arquimedianas en un problema específico? pensando en que la cópula real puede no pertenecer a esta familia. No tengo una respuesta definitiva. Lo único que puedo decir es que hay una gran variedad de relaciones entre variables en que podemos tener una cópula Arquimediana, es decir, la familia Arquimediana no está limitada a no poder describir tal o cual modalidad de dependencia, de modo que si

bien no hay razón para decidir que la cópula es Arquimediana, tampoco hay restricciones teóricas en cuanto a los ámbitos en que debemos o no usarlas. El método expuesto es aún más específico en la selección ya que se basa en encontrar el parámetro suponiendo una familia paramétrica de cópulas Arquimedianas. En favor de esto sólo puedo decir que no se elige una única familia paramétrica sino varias. En realidad se eligen familias en que el proceso sea viable, es decir, en que podamos invertir la función que define la tau de Kendall ya sea directamente o numericamente. Como apoyo al resultado (la selección de una cópula), se pueden ajustar distribuciones marginales y generar con ello tablas de frecuencia de la distribución conjunta resultante y compararlas con los datos. Si hemos encontrado una distribución que consistentemente recupera la estructura general de nuestros datos, podemos usarla como la distribución de la muestra siempre que en la práctica refleja similares resultados. Existen otros métodos que se tratan en la literatura para ajustar cópulas Arquimedianas. Al respecto es interesante preguntarse si son similares los resultados y si se ha probado la mayor efectividad de alguno de ellos y en qué casos.

La prueba de independencia desarrollada en el último Capítulo parte de un contexto muy general: la convergencia de la cópula empírica. Este hecho es a la vez favorable y desfavorable. Podemos asegurar que en caso de rechazar independencia hemos de confiar firmemente en esta conclusión. Al mismo tiempo cuando no somos capaces de rechazar independencia, no hay cosa alguna que decir. Se pueden construir muchos ejemplos en que la dependencia es sutil, en que la cópula es muy similar a la cópula independiente y esta prueba no detecta aún a niveles muy altos de confianza. No podemos pedirle más a un procedimiento tan universal. Queda mucho por estudiar de esta prueba. Para empezar, encontrar la distribución del estadístico propuesto y calcular sus cuantiles para dar un carácter objetivo a la selección de la región de rechazo. Un cálculo similar a este se puede hallar en [4] donde se construye una prueba bastante más compleja para independencia.

Esta misma prueba y los reordenamientos de M son una pareja algo escandalosa. Es algo impresionante, de primer momento, que podamos encontrar una función invertible para relacionar dos variables que siempre van a reportar ser independientes. Esto revela que en el ámbito de lo práctico la independencia y la dependencia más estrecha no son tan distintas, sino que pueden ser casi iguales.

Bibliografía

- [1] Nelsen, Roger B. *An introduction to Copulas*. Springer-Verlag, 1999
- [2] Genest, Ch. & Rivest, L. *Statistical Inference Procedures for Bivariate Achimedean Copulas* Journal of the American Statistical Association, **88**, 1034-1043. 1993
- [3] Darsow, W.F., Nguyen, B. & Olsen, E.T. *Copulas and Markov Processes* Illinois Journal of Mathematics, **36**, 600-642. 1992
- [4] Deheuvels, P. *La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance* Bulletin Acad. Royal Belgique Cl.Sci. **65**, 274-292. 1979
- [5] Mood, A., Graybill, F., Boes, D. *Introduction to the theory of statistics* McGraw-Hill Book Company 1986
- [6] Klugman, S., Parsa, R. *Fitting bivariate loss distributions with copulas* Insurance: Mathematics & Economics, **24**, 139-148. 1999
- [7] Frees, E., Valdez, E. *Understanding relationships using Copulas* North American Actuarial Journal, Vol 2, No. 1 1997
- [8] Contreras-Cristán, A., González-Barrios, J.M. *A statistical Method for the determination of the appropriate order in a general class of time series models*
- [9] Embrechts, P., McNeil, A., Straumann, D. *Correlation and Dependence in Risk Management*
- [10] Schweizer, B., Sklar, A. *Probabilistic Metric Spaces* North-Holland, New York 1983

- [11] Bartle, R.G. *The Elements of Integration and Lebesgue Measure* Wiley 1996
- [12] Royden *Real Analysis* Macmillan 1971