

Universidad Nacional Autónoma de México

Facultad de Estudios Superiores Acatlán

Actuaría

Cóputas: modelado y medición de dependencia  
estocástica bivariada y su aplicación en el cálculo del  
Valor en Riesgo

T E S I S

Cid Omar Pérez Pérez

Asesor : Dr. Pablo Pérez Akaki

Enero, 2007



Universidad Nacional  
Autónoma de México



**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.

*En la orilla del tiempo dejé  
un eslabón grabado:*

*¿Qué buscas?*

*Dedicatoria*

*A mis padres Tere y Carlos por su amor,  
su ejemplo y su interminable compromiso con la excelencia.*

*A Sandy y a Luis por ser  
una fuente inagotable de motivación.*

*A Mamá Carmen  
por su invaluable ejemplo de vida.*

*A mis amigos por  
su entrega y permanencia*

*A mis profesores y amigos  
por sus conocimientos y ejemplo.*

*A mi querida Universidad.*

<b>Introducción</b>	<b>1</b>
<b>1 Cópulas</b>	<b>3</b>
1.1 Definiciones . . . . .	4
1.2 Propiedades básicas . . . . .	7
1.3 Teorema de Sklar . . . . .	11
1.4 Cópulas y variables aleatorias continuas . . . . .	14
1.5 Construcción de cópulas . . . . .	25
1.6 Cópulas Arquimedeanas . . . . .	29
<b>2 Dependencia</b>	<b>39</b>
2.1 Modos de dependencia . . . . .	40
2.2 Correlación . . . . .	44
2.3 Medidas de Concordancia . . . . .	48
2.4 Medidas de dependencia . . . . .	55
2.5 Cópulas empíricas . . . . .	57
<b>3 Ajuste de cópulas bivariadas</b>	<b>63</b>
3.1 Estimación paramétrica . . . . .	65
3.2 Estimación no paramétrica . . . . .	69
3.3 Revisión gráfica . . . . .	72
3.4 Revisión analítica . . . . .	74
3.5 Problemas en el ajuste . . . . .	78
<b>4 Aplicación de las cópulas: Valor en Riesgo</b>	<b>81</b>
4.1 Valor en Riesgo . . . . .	82
4.2 Empleo de una cópula para el cálculo del VaR . . . . .	85
4.3 Contraste de los resultados . . . . .	96
4.4 Otras aplicaciones de interés actuarial . . . . .	99
<b>Conclusiones</b>	<b>111</b>
<b>A Modos de dependencia y <math>\tau</math> de Kendall</b>	<b>113</b>
<b>B Algoritmos en Matlab</b>	<b>115</b>
<b>Bibliografía</b>	<b>121</b>

# Introducción

La ciencia actuarial frecuentemente está interesada en el modelado de fenómenos que incorporan tanto la interacción entre múltiples entidades, como el comportamiento propio de cada una de ellas. Ejemplos de ello son las funciones de supervivencia múltiple en las que los asegurados y la relación entre ellos son factores indispensables para una correcta especificación del modelo; el reaseguro, en el cual el conocimiento de las reclamaciones, los gastos de una aseguradora, así como de una relación entre ellas es indispensable para la determinación de las primas. Fuera del campo de los seguros, en otros aspectos financieros destacan la valuación de opciones en las que hay varios subyacentes involucrados con comportamientos individuales muy diversos; y como tema de gran interés y actualidad -debido principalmente a las iniciativas internacionales para la administración del riesgo de crédito y de mercado (p.ej. Basilea II)- el valor en riesgo (VaR), en el que el comportamiento conjunto de los activos o incumplimientos considerados es tan importante como el comportamiento individual de ellos.

Las herramientas existentes para abordar los tópicos anteriores y otros que requieren de modelado de dependencia son diversas y en muchos casos resuelven completamente el problema. Algunos ejemplos de ellas son las pruebas de independencia, de normalidad multivariada, la regresión lineal y no lineal por mencionar algunos. Sin embargo, en ocasiones el empleo de herramientas como las mencionadas anteriores requieren generalmente de supuestos que comúnmente no se cumplen en la práctica. Por ejemplo, la dependencia entre un par de activos financieros difícilmente es lineal y tampoco puede modelarse con una distribución normal, dado el comportamiento individual.

Cuando no es posible emplear herramientas como las anteriores, o bien, cuando éstas parecen ser insuficientes, existe una alternativa explorada en los últimos años cuya comprensión es bastante simple al dividir el problema del modelado de un comportamiento conjunto en dos aspectos básicos e intuitivamente claros: el comportamiento individual de los fenómenos de interés y una regla de interdependencia entre ellos.

Las *cóputas* (enlace entre dos cosas en castellano) permiten realizar precisamente la separación anterior, facilitando con ello la descripción y el análisis del comportamiento conjunto de dos o más fenómenos. En términos simples, estas herramientas son funciones matemáticas cuyos atributos garantizan la construcción de funciones de distribución probabilísticas.

Dado el potencial de las cópulas para el modelado de fenómenos multivariados de interés actuarial, el objetivo de esta presentación es mostrar algunas de las propiedades de ellas, así como su empleo en un caso el cálculo del VaR par-

tiendo de la hipótesis de que las cópulas permiten una mejor representación la alternativa de normalidad usada frecuentemente en la práctica.

Para lo anterior, dedicaremos el primer capítulo a la revisión de sus propiedades generales en términos matemáticos, su relación con las variables aleatorias, algunas estrategias para la construcción de ellas y las características de las llamadas cópulas arquimedeanas. Se presenta aquí el teorema de Sklar, resultado fundamental de la teoría de cópulas que formaliza la función antes descrita.

En el segundo capítulo, se presenta tema del modelado y medición de dependencia estocástica, mostrando la forma en que las cópulas permiten realizar ambas cosas con casi un mínimo de esfuerzo. Se hace una breve revisión del coeficiente de correlación lineal para la caracterización de la dependencia, a las medidas de concordancia y las medidas de dependencia. Como resultado principal de este capítulo, se estudia la cópula empírica cuya importancia para la teoría de cópulas es similar a la de la función empírica para la Estadística, al permitir la construcción de estadísticos de independencia y de bondad de ajuste para una muestra.

Dentro del tercer capítulo se aborda el importante tema de estimación de parámetros y bondad de ajuste para cópulas, dejando de lado el tema de estimación para funciones marginales por asumirse completa familiaridad con él. Como resultados principales destacan los estadísticos de bondad de ajuste exclusivos para cópulas y la estimación para cópulas no necesariamente arquimedeanas.

En el cuarto capítulo se presenta la aplicación empírica, en donde se desarrolla con todo detalle el cálculo del VaR de dos portafolios compuestos por acciones que cotizan en la Bolsa Mexicana de Valores. Se muestra aquí un breve análisis comparativo que muestra las ventajas y desventajas del empleo de las cópulas en comparación con el cálculo a partir del supuesto de normalidad conjunta y se emite una conclusión al respecto. Se esbozan además algunas otras aplicaciones de interés actuarial donde las cópulas tienen cabida.

Por último, se presenta un apéndice donde se muestran las rutinas computacionales empleadas para la simulación, estimación de parámetros, pruebas de bondad de ajuste y cálculo del VaR empleados, en conjunto con algunas tablas que resumen las características de algunas cópulas.

Deseo hacer notar que se ha procurado que el contenido de esta tesis sea completo, en el sentido que contiene tanto los conocimientos formales necesarios -y dentro del alcance de un egresado de la licenciatura de Actuaría- para sustentar la teoría de cópulas, como algunas observaciones personales que en conjunto con el ejemplo práctico desarrollado, pretenden brindar un acercamiento intuitivo y digerible a los planteamientos teóricos.

Sólo resta mencionar que el trabajo se restringe al caso bivariado, ya que por una parte, la teoría multivariada eleva en gran medida el nivel de complejidad y por otra, esta aproximación es una base suficiente para la comprensión de esta extensión. Ahora bien, para un entendimiento completo de los planteamientos teóricos presentados, son indispensables conceptos básicos de análisis real y simulación de variables aleatorias, sólidos fundamentos de cálculo integral y diferencial, además de una base también sólida de probabilidad y estadística. Si bien algunos conceptos de teoría de la medida son también de gran utilidad, no serán indispensables.

# Capítulo 1

## Cóputas

En castellano, la palabra cóputa significa enlace, ligamiento o atadura de una cosa con otra. Fue empleada por primera vez en 1959 en un contexto matemático por Abe Sklar <sup>1</sup> donde se muestran las características de las funciones que permiten atar distribuciones unidimensionales para conformar distribuciones multivariadas además de un importante teorema que prueba la existencia de estas dentro de cualquier distribución n-dimensional dada, de ahí que las cóputas reciben precisamente este nombre.

Anterior a este trabajo, pueden encontrarse estudios sobre las cóputas en los trabajos de Wassily Hoeffding (1940) donde muestra distribuciones bivariadas estandarizadas cuya única diferencia con funciones formalizadas por Sklar corresponde al espacio en que estas tienen cabida (las cóputas son funciones en  $(0,1) \times (0,1)$  y las funciones de Hoeffding en  $(-\frac{1}{2}, \frac{1}{2}) \times (-\frac{1}{2}, \frac{1}{2})$ ).

Si bien la investigación en esta área se dió principalmente por el estudio de espacios métricos, recientemente se ha destacado su importancia para la estadística y la probabilidad y de hecho, de acuerdo con Nelsen <sup>2</sup> es a partir de la década de los noventas, con algunas conferencias en Roma, Seattle y Praga que esta iniciativa cobra mayor fuerza.

Las cóputas son de interés en la estadística y la probabilidad, debido a que, como lo afirma Fisher <sup>3</sup>, por una parte permiten el estudio de medidas de dependencia libres de escala y por otra son un punto de partida para la construcción de nuevas familias de distribuciones probabilísticas.

El objeto de este primer capítulo es la presentación de algunos resultados importantes acerca de las cóputas. Aquí las definiremos formalmente, haremos una revisión de sus características más importantes y las vincularemos con las variables aleatorias. Adicionalmente, mostramos algunas estrategias para la construcción de cóputas y las denominadas cóputas arquimedeanas cuyas características se explotarán en los siguientes capítulos.

Como hicimos notar, las cóputas existen en el sentido matemático más general y la interpretación y empleo probabilístico es tan solo una vertiente, por lo

---

<sup>1</sup>Sklar, A.(1959) 'Fonctions de répartition à n dimensions et leurs marges'. *Publ Publications de l'Institut de Statistique de L'Université de Paris*, 8, 229-231

<sup>2</sup>Nelsen, R. (1999). *An introduction to copulas*, Springer-Verlag.

<sup>3</sup>Fisher, N.I. (1997) 'Copulas', *Encyclopedia of Statistical Sciences*, Update Vol 1, John Wiley & Sons.



que se ha decidido mantener la generalidad en este primer capítulo sobre todo en las dos primeras secciones y vincular posteriormente estos conocimientos con conceptos probabilísticos de interés. De igual manera, se recuerda al lector que se tratará únicamente con cópulas bivariadas y cuando se hable de cópulas multivariadas nos referiremos a una dimensión mayor o igual a tres.

## 1.1. Definiciones

Previo a la revisión de los conceptos que son objeto de este capítulo es necesaria la introducción de la siguiente notación y definiciones de acuerdo con Erdely. <sup>4</sup>

De inicio, considérese a la recta real  $\mathbf{R} = (-\infty, +\infty)$ ; a los reales extendidos como  $\bar{\mathbf{R}} = \mathbf{R} \cup \{-\infty, \infty\}$  y al plano extendido  $\bar{\mathbf{R}}^2 = \bar{\mathbf{R}} \times \bar{\mathbf{R}}$

**Definición 1.1.1** Sean  $[x_1, y_1], [x_2, y_2] \in \bar{\mathbf{R}}^2$   $x_1 < x_2$ ,  $y_1 < y_2$  se le llama rectángulo en  $\bar{\mathbf{R}}^2$  a  $\mathbf{B} = [x_1, y_1] \times [x_2, y_2]$  y a los puntos  $(x_1, y_1), (x_1, y_2)$ ,  $(x_2, y_1), (x_2, y_2)$  se les denomina vértices.

Para nuestros fines, bastará considerar el rectángulo unitario

$$\mathbf{I}^2 = \mathbf{I} \times \mathbf{I} = [0, 1] \times [0, 1]$$

**Definición 1.1.2** Sean  $S_1, S_2$  subconjuntos no vacíos de  $\bar{\mathbf{R}}^2$ ,  $S$  subconjunto de  $\mathbf{R}$  y  $H : S_1 \times S_2 \rightarrow S$ .

Se dice que  $H$  es función real de doble entrada con  $\text{Dom}H = S_1 \times S_2$ ;  $\text{Ran}H = S$ .

**Definición 1.1.3** Sea  $H$  una función real de doble entrada y sea  $\mathbf{B} := [x_1, x_2] \times [y_1, y_2]$  un rectángulo cuyos vértices están contenidos en  $\text{Dom}H$ . El  $H$ -volumen de  $\mathbf{B}$  se denota y está dado por la siguiente expresión

$$V_H(\mathbf{B}) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1).$$

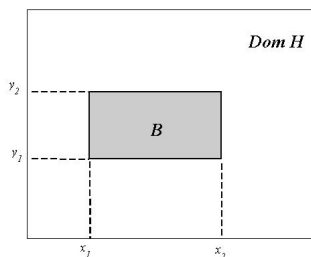


Figura 1.1: Rectángulo sobre el que se calcula el H-volumen

<sup>4</sup>Erdely, A.(2003) *Construcción de medidas de dependencia por medio de cópulas*. Tesis UNAM

**Definición 1.1.4** Sea  $H$  función real de doble entrada.

Se dice que  $H$  es 2-creciente si  $0 \leq V_H(B)$  para todo rectángulo  $B$  cuyos vértices están en  $\text{Dom}H$ .

**Ejemplo 1.1.1** Sea  $[u_1, u_2] \times [v_1, v_2]$  un rectángulo en  $\mathbf{I}^2$ ; las funciones siguientes son 2-crecientes considerando la función indicadora  $\mathbf{1}_A = 1$  para  $a \in A$  y 0 en otro caso.

1.  $\Pi(u, v) := uv$
2.  $M(u, v) := \min\{u, v\} = (u - v) \mathbf{1}_A + v$  con  $A := \{(u, v) \in \mathbf{I}^2 : u \leq v\}$
3.  $W(u, v) := \max\{u + v - 1, 0\} = (u + v - 1) \mathbf{1}_B$  con  $B := \{(u, v) \in \mathbf{I}^2 : u + v - 1 \geq 0\} \in \mathbf{I}^2$

$$1) V_{\Pi}([u_1, u_2] \times [v_1, v_2]) = u_2v_2 - u_2v_1 - u_1v_2 + u_1v_1 \\ = (u_2 - u_1)(v_2 - v_1) \geq 0$$

$$2) V_M([u_1, u_2] \times [v_1, v_2]) = \\ (u_2 - v_2) \mathbf{1}_A(u_2, v_2) - (u_2 - v_1) \mathbf{1}_A(u_2, v_1) \\ - (u_1 - v_2) \mathbf{1}_A(u_1, v_2) + (u_1 - v_1) \mathbf{1}_A(u_1, v_1);$$

Si todos los vértices están dentro o fuera del conjunto  $A$  el volumen es cero; si se encuentra solo alguno de los pares presentados en las funciones indicadoras, el volumen es positivo; si se encuentran los pares de la izquierda o los de la derecha, también es positivo; cualquier otro caso es cero.

$$3) V_W([u_1, u_2] \times [v_1, v_2]) = \\ (u_2 + v_2 - 1) \mathbf{1}_B(u_2, v_2) - (u_2 + v_1 - 1) \mathbf{1}_B(u_2, v_1) \\ - (u_1 + v_2 - 1) \mathbf{1}_B(u_1, v_2) + (u_1 + v_1 - 1) \mathbf{1}_B(u_1, v_1)$$

Siguiendo el mismo razonamiento hecho para la función anterior, se concluye que  $V_W(B) \geq 0$ .

La característica 2-creciente es mucho más general que lo que puede apreciarse a primera vista y que en la siguiente sección servirá para diferenciar entre cópulas y funciones de doble entrada. Algo importante respecto a esta característica, es que no debe confundirse con la monotonicidad creciente en los argumentos de la función. De hecho, como veremos en el siguiente ejemplo:

- i) 2-creciente  $\not\Rightarrow$  creciente monótona en sus argumentos
- ii) creciente monótona en ambos argumentos  $\not\Rightarrow$  2-creciente

**Ejemplo 1.1.2** i) Sea  $H_1(u, v) = v - u$  con  $H_1 : \mathbf{I} \times \mathbf{I} \rightarrow \mathbf{I}$

$$V_{H_1}([u_1, u_2] \times [v_1, v_2]) = (v_2 - u_2) - (v_1 - u_2) - (v_2 - u_1) + v_1 - u_1 \geq 0$$

pero claramente para una  $v_0$  fija, la función es decreciente en  $u$  a partir de  $u \geq v_0$

ii) Sea  $H_2(u, v) = 1 - uv$  con  $H_2 : [-1, 0] \times [-1, 0] \rightarrow \mathbf{I}$   $H_2$  es creciente para  $u$  y  $v$ ; sin embargo,

$$\begin{aligned} V_{H_2}([u_1, u_2] \times [v_1, v_2]) &= 1 - (u_2v_2) - (u_2v_1) - (u_1v_2) + u_1v_1 \\ &= (u_1 - u_2)(v_2 - v_1) \end{aligned}$$

Tómese por ejemplo  $a \in \text{Dom}H_2$  Entonces  $V_{H_2}([a, 0] \times [a, 0]) = -a^2 \leq 0$

**Definición 1.1.5** Sean  $S_1, S_2$  subconjuntos no vacíos de  $\mathbf{R}$  con  $a_1 := \inf(S_1)$  y  $a_2 := \inf(S_2)$ . La función real de doble entrada  $H$  está fijada si:  $H(x, a_2) = 0 = H(a_1, y)$ , para todo  $(x, y)$  en  $\text{Dom}H = S_1 \times S_2$ .

Complementario a lo mencionado con respecto la propiedad 2-creciente, si la función en cuestión se fija, puede concluirse que:

1.  $H(x, y) \geq 0$ , haciendo  $x_1 = y_1 = 0$  en el H-volumen
2.  $H$  es monótona creciente por argumentos

A saber, para  $x_1 < x_2 \in S_1$  la función fijada  $G(t) := H(x_2, t) - H(x_1, t)$  es no decreciente en  $S_2$  como consecuencia directa de la definición:

$$H(x_2, t+h) - H(x_1, t+h) - H(x_2, t) + H(x_1, t) > 0 \Rightarrow G(t+h) > G(t)$$

Para  $F(t) := H(t, y_2) - H(t, y_1)$  se tiene algo similar en  $S_1$

A continuación se presentan las definiciones generales de función de distribución, función de distribución conjunta y funciones de distribución marginales. Como se observará las definiciones en el sentido probabilístico guardan gran similitud con ellas.

**Definición 1.1.6** Sea  $F$  una función con  $\text{Dom}F = \bar{\mathbf{R}}$ . Si  $F$  es no decreciente y  $F(-\infty) = 0$  y  $F(+\infty) = 1$ ,  $F$  es una función de distribución.

**Definición 1.1.7** Sea  $H$  una función real de doble entrada. Si  $H$  es 2-creciente, está fijada para  $a = -\infty$  y  $H(\infty, \infty) = 1$  entonces  $H$  es función de distribución conjunta.

**Definición 1.1.8** Sean  $S_1, S_2$  subconjuntos no vacíos de  $\mathbf{R}$  tales que:  $b_1 := \sup(S_1)$  y  $b_2 := \sup(S_2)$ . Decimos entonces que la función de distribución  $H : S_1 \times S_2$  tiene marginales :

$$F : S_1 \rightarrow \mathbf{R} \text{ con } F(x) := H(x, b_2),$$

$$G : S_2 \rightarrow \mathbf{R} \text{ con } G(y) := H(b_1, y)$$

**Definición 1.1.9** Sea  $F$  una función de distribución. La cuasi-inversa de  $F$  denotada por  $F^{-1}$  con dominio  $\mathbf{I}$  es tal que:

1. Si  $t \in \text{Ran}F$ ,  $F^{(-1)}(t) = x$  con  $x \in \bar{\mathbf{R}}$  y  $F(x) = t$  i. e.  $F(F^{-1}(t)) = t$
2. Si  $t$  no está en  $\text{Ran}F$ ,  $F^{(-1)}(t) = \inf \{x : F(x) \geq t\} = \sup \{x : F(x) \leq t\}$ .

*Observación:* Si  $F$  es estrictamente creciente, la cuasi-inversa es  $F^{-1}$  que desde luego es única.

Para finalizar esta primer sección se presenta una proposición que adelante resultará de utilidad.

**Proposición 1.1** *Sea  $H$  fijada y 2-creciente,  $F$  y  $G$  sus marginales. Sean  $(x_1, y_1)$  y  $(x_2, y_2) \in \text{Dom } H$ . Entonces:*

$$|H(x_2, y_2) - H(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)|$$

*Demostración:*

Usando la desigualdad del triángulo

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

por ser 2-creciente y fijada, de acuerdo con las definiciones anteriores

$$|H(x_2, y_2) - H(x_1, y_2)| \leq |F(x_2) - F(x_1)|$$

análogo para  $G$ ;

$$|H(x_1, y_2) - H(x_1, y_1)| \leq |G(y_2) - G(y_1)|$$

□

*Observación:* Por lo anterior, puede decirse que  $H$  cumple la condición de Lipschitz en  $\mathbf{I}^2$ .

Como se habrá notado, la relación de las definiciones anteriores con las definiciones en Probabilidad es grande. Con el fin de evidenciar esta relación se presenta el siguiente ejemplo.

**Ejemplo 1.1.3** *Sean  $X, Y$  variables aleatorias, con funciones de distribución  $F(x) = P(X \leq x)$  y  $G(y) = P(Y \leq y)$  respectivamente y sea  $H = P(X \leq x, Y \leq y)$  la función de distribución conjunta asociada a estas variables aleatorias.*

*Es claro que estas funciones de distribución cumplen con con las definiciones generales presentadas con anterioridad. De hecho, tienen otras particularidades; a saber, ser continuas por la derecha y estar relacionadas directamente con las funciones de densidad  $f, g$  y  $h$ . Sobre la función  $H$  puede decirse cumple con ser 2-creciente y estar fijada. La cuasi-inversa de una función de probabilidad acumulada se calcula de manera similar al incorporar los límites por la izquierda. Sobre la última proposición es necesario recordar que  $P(A \cap B) \leq P(A) + P(B)$  para dos eventos cualesquiera; en otros términos,  $H(x_2, y_2) \leq F(x_2) + G(y_2)$  y algo parecido para  $H(x_1, y_1)$ , con lo cual nos aproximamos bastante al resultado de la proposición.*

## 1.2. Propiedades básicas

Una vez presentada la notación a emplear, iniciamos con nuestro estudio sobre cópulas. En esta sección rescatamos la definición formal de cópula y sus características básicas.

**Definición 1.2.1** Sea  $C$  una función real de doble entrada. Se dice que  $C$  es una 2-cópula o cópula bivariada si cumple con las siguientes características

1.  $DomC = \mathbf{I} \times \mathbf{I}$
2.  $C$  esta fijada en cero y es 2-creciente
3. Para todo  $u, v$   $C(u,1) = u$  y  $C(1,v)=v$

Como se observa a partir de la definición anterior, las cópulas son funciones bastante simples.

Estas son un caso particular de las 2-subcópulas, definidas de manera idéntica a las 2-cópulas con la diferencia de que para  $C'$  subcópula  $DomC' = S_1 \times S_2$  con  $S_i \subseteq \mathbf{I}$  y  $\{0,1\} \subseteq S_i$

En adelante trabajaremos únicamente con cópulas bivariadas y las nombraremos únicamente cópulas, enunciando la dimensión de estas solo si es necesario diferenciarlas.

**Ejemplo 1.2.1** Retomando las funciones  $\Pi$ ,  $M$  y  $W$  presentadas en la sección anterior, tenemos que el dominio de ellas es precisamente  $\mathbf{I}^2$  y para todo  $u, v$  en  $\mathbf{I}$

1.  $\Pi(u,0) = u * 0 = 0 = 0 * v = \Pi(0,v)$ ;  
 $\Pi(u,1) = u * 1 = u, \Pi(1,v) = 1 * v = v$ .  
 Por lo tanto  $\Pi$  cumple con la Definición 1.2.1 y por su estructura se le denomina *cópula producto*.
2.  $M(u,0) = \min\{u,0\} = 0 = \min\{0,v\} = M(0,v)$ ;  
 $M(u,1) = \min\{u,1\} = u, M(1,v) = \min\{1,v\} = v$ . Por tanto es cópula
3.  $W(u,0) = \max\{u+v-1,0\} = 0 = \max\{u+v-1,0\} = W(0,v)$ ,  
 $W(u,1) = \max\{u,0\} = u, W(1,v) = \max\{v,0\} = v$ . Y también  $W$  es cópula.

Un ejemplo claro de funciones que no son cópulas, son las definidas en el ejemplo 1.1.2, que de hecho, no cubren ninguna de las características de la necesarias.

Como consecuencia directa de la definición anterior, al ser  $C$  una función fijada y 2-creciente, esta es monótona creciente en sus argumentos y de hecho:

$$C(u,v) \leq C(1,1) \leq 1.$$

Por otra parte, al estar fijada en cero:  $0 \leq C(u,v)$ ; con lo que se llega a

$$0 \leq C(u,v) \leq 1 \tag{1.1}$$

También como consecuencia de la Definición 1.2.1, por la Proposición 1.1 se tiene:

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1| \tag{1.2}$$

lo que implica que las cópulas (y también las sub-cópulas si se observa la definición) son uniformemente continuas en  $\mathbf{I}$  y por tanto sus derivadas existen en casi todas partes. De acuerdo a esto, para  $v$  fija

$$|C(u+\Delta u, v) - C(u, v)| \leq \Delta u \Rightarrow \frac{|C(u+\Delta u, v) - C(u, v)|}{\Delta u} \leq 1$$

por lo anterior

$$0 \leq \frac{\partial}{\partial u} C(u, v) \leq 1 \quad (1.3)$$

y de manera similar

$$0 \leq \frac{\partial}{\partial v} C(u, v) \leq 1 \quad (1.4)$$

Lo que indica que las variaciones de la cópula con respecto a cada una de sus variables están acotadas. Ahora bien, si consideramos las derivadas parciales como funciones de  $u$  y  $v$  respectivamente, empleando el hecho de que  $C$  es monótona creciente en cada uno de sus argumentos, vemos que las derivadas parciales son funciones monótonas crecientes, lo que indica que las variaciones tienden a permanecer iguales o a incrementarse de forma suave.

Por la desigualdad 1.1 sabemos que las cópulas están acotadas por abajo al ser siempre no negativas y por arriba al ser siempre menores o iguales que la unidad. Esta restricción en conjunto con la restricción sobre sus derivadas parciales nos habla de un comportamiento muy específico.

En el lema siguiente mostramos que las cópulas  $W$  y  $M$  revisadas con anterioridad, son cotas de las cópulas, haciendo así aún más específico el espacio en que tienen cabida.

**Lema 1.2.1** *Sea  $C$  una cópula. Entonces:*

$$W(u, v) \leq C(u, v) \leq M(u, v) \quad (1.5)$$

*Demostración:*

i)  $C(u, 1) = u$  y  $C(1, v) = v$  por ser no decrecientes

$$C(u, v) \leq \min\{u, v\} = M(u, v)$$

ii)  $V_C([u, 1] \times [v, 1]) \geq 0 \Rightarrow C(u, v) \geq u + v - 1$  como también  $C(u, v) \geq 0$  entonces

$$C(u, v) \geq W(u, v)$$

□

A  $W$ ,  $M$  se les denomina cotas inferior y superior respectivamente y de manera trivial verificamos que  $W \leq C \leq M$ . Si se observa con cuidado, el

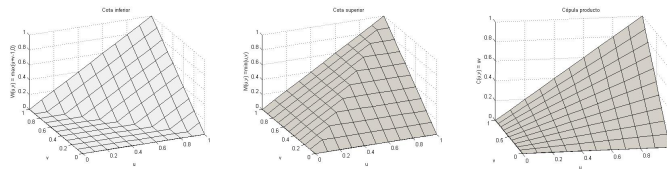


Figura 1.2: Cópulas  $W$ ,  $M$  y  $\Pi$

espacio comprendido entre el gráfico 1 y el 2 es precisamente el lugar geométrico donde las cópulas tienen cabida. Para una cópula bivariada es fácil visualizar su

comportamiento aunque puede optarse por los *conjuntos de nivel*. Para  $t \in \mathbf{I}$  fija, estos corresponden a

$$A_t(C) := \{(u, v) \in \mathbf{I}^2 | C(u, v) = t\}$$

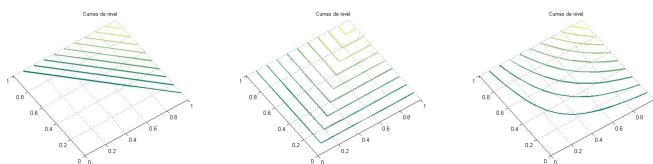


Figura 1.3: Curvas de nivel de las cópulas  $W$ ,  $M$  y  $\Pi$

A partir de la figura 1.3 podemos diferenciar perfectamente entre las cópulas del ejemplo. Además, dado que  $M$  y  $W$  son cotas, la gráfica de las curvas de nivel de cualquier cópula estará encerrada entre los triángulos generados por las curvas de nivel de estas; ejemplo de ello son las curvas de nivel de la cópula producto.

Una herramienta más para la visualización de cópulas bivariadas se encuentra en la siguiente

**Definición 1.2.2** Sea  $C$  una cópula y sea  $a \in \mathbf{I}$ . Definimos:

- i) sección horizontal en  $a$  como la función  $\tilde{h} : \mathbf{I} \rightarrow \mathbf{I}$  dada por  $\tilde{h}_a(t) := C(t, a)$
- ii) sección vertical en  $a$  como la función  $v : \mathbf{I} \rightarrow \mathbf{I}$  dada por  $v_a(t) := C(a, t)$
- iii) sección diagonal como la función  $\delta : \mathbf{I} \rightarrow \mathbf{I}$  dada por  $\delta_C(t) := C(t, t)$

Nótese que por los resultados anteriores, las secciones verticales y horizontales definidas cumplen con ser uniformemente continuas, no decrecientes, y sus derivadas son no negativas y menores que la unidad

Una de las propiedades más interesantes de las cópulas es que bajo la sumas convexas generan un espacio cerrado. Para formalizar esto se presenta el siguiente resultado.

**Proposición 1.2** Sean  $a_i \in \mathbf{I}$  con  $\Sigma a_i = 1$  y para  $C_i$  cópulas;  $i = 1, 2, \dots, n$ . Entonces la combinación convexa  $\Sigma a_i C_i$  es una cópula

*Demostración:*

1. De inicio  $C(u, 0) = \Sigma a_i C_i(u, 0) = \Sigma(a_i \cdot 0) = 0 = C(0, v)$
2.  $C(u, 1) = \Sigma a_i C_i(u, 1) = \Sigma(a_i u) = u$ ; análogamente  $C(1, v) = v$
3.  $V_C([u_1, v_1] \times [u_1, v_1]) = \Sigma a_i C_i(u_2, v_2) - \Sigma a_i C_i(u_2, v_1) + \Sigma a_i C_i(u_1, v_2) - \Sigma a_i C_i(u_1, v_1) = a_1 V_{C_1}([u_1, v_1] \times [u_1, v_1]) + a_2 V_{C_2}([u_1, v_1] \times [u_1, v_1]) + \dots + a_n V_{C_n}([u_1, v_1] \times [u_1, v_1]) = \Sigma a_i V_{C_i}([u_1, v_1] \times [u_1, v_1]) \geq 0$

□

**Ejemplo 1.2.2** Sean  $a_1 = \alpha$ ,  $a_2 = 1 - \alpha - \beta$ ,  $a_3 = \beta$ ;  $C_1 = M$ ;  $C_2 = \Pi$  y  $C_3 = W$ ;  $C$  definida como  $C_{(\alpha, \beta)}(u, v) = \sum a_i C_i(u, v)$  es una cópula.

Notamos que  $C$  será una cópula mientras los parámetros satisfagan  $\alpha + \beta \leq 1$ . y de hecho todos los elementos del conjunto  $\{C_{\alpha, \beta} : \alpha + \beta \leq 1\}$  son cópulas. En general, a cualquier conjunto de cópulas

$$\{C_{\alpha_1, \alpha_2, \dots, \alpha_n} : \alpha_1 \in A_1, \alpha_2 \in A_2, \dots, \alpha_n \in A_n\}$$

se le denomina *familia de cópulas parametrizada*.

Observamos también que incorpora a las cópulas  $\Pi, M, W$  'barriendo' todo el espacio en que las cópulas tienen sentido de acuerdo al valor de sus parámetros. A este tipo de familias se les denomina *familia comprensiva*.

### 1.3. Teorema de Sklar

En la sección anterior se mostró que las cópulas son funciones bien comportadas y con cualidades interesantes.<sup>5</sup> En esta sección mostramos el resultado más importante en la teoría de cópulas y tal vez también el más importante para su aplicación en estadística.

El Teorema de Sklar formaliza el papel que juegan las cópulas en la relación entre funciones multivariadas y sus marginales comentada con anterioridad. La demostración de este resultado se efectúa siguiendo los argumentos en Nelsen<sup>6</sup> y requiere la presentación de dos lemas cuya demostración puede encontrarse en este mismo texto o bien en Erdelyi<sup>7</sup> y aquí la hemos omitido por no ser nuestro principal interés además de requerir de otras herramientas matemáticas. Su presentación aquí tendrá el objetivo de desarrollar intuición sobre la estructura de la prueba de este importante resultado.

Por una parte, el primer lema garantiza la existencia y unicidad de una sub-cópula que une las distribuciones marginales para conformar la distribución conjunta.

**Lema 1.3.1** Sea  $H$  una distribución conjunta con marginales  $F$  y  $G$ . Existe una única 2-subcópula  $C'$  tal que:

- i)  $Dom C' = Ran F \times Ran G$
- ii)  $H(x, y) = C'(F(x), G(y))$  para  $x, y \in \bar{\mathbf{R}}$

□

La demostración del resultado requiere definir una relación tal que se cumplan los dos puntos anteriores y después probar que esta es una sub-cópula y que es única. El segundo lema garantiza que una sub-cópula cualquiera puede 'extenderse' a una cópula. Como se recordará, la diferencia entre estas dos funciones es únicamente el dominio de ellas; mientras que las cópulas tienen por dominio el

<sup>5</sup>Nos referimos a que son funciones perfectamente definidas, están acotadas, su derivada existe casi en todo su dominio y cumplen la característica 2-creciente

<sup>6</sup>Nelsen, Op.Cit. p18

<sup>7</sup>Erdelyi, Op.Cit pp 15-18



cuadrado unitario  $I \times I$ , las sub-cópulas tienen su dominio en dos subconjuntos arbitrarios cuya única restricción es que incluyan al 0 y al 1. Intuitivamente, la extensión 'estira' la sub-cópula de forma que tenga por dominio el cuadrado unitario y mantenga sus propiedades.

**Lema 1.3.2** *Cualquier subcópula  $C'$  puede extenderse a una cópula  $C$ . Esta extensión en general no es única.*

□

Veamos ahora el Teorema de Sklar y comentemos un poco sobre las dos implicaciones que tiene. La primera es que, dada una distribución conjunta, esta puede descomponerse como una cópula y sus distribuciones marginales. La segunda es que dadas dos distribuciones marginales y una cópula, se puede componer una distribución conjunta a partir de ellas.

**Teorema 1.3.1 (Sklar, (1959))** *Sea  $H$  una función de distribución conjunta con marginales  $F$  y  $G$ . Entonces existe una cópula  $C$  tal que*

$$H(x, y) = C(F(x), G(y)) \quad (1.6)$$

*Si  $F$  y  $G$  son continuas entonces  $C$  es única, de otro modo  $C$  es única en  $\text{Ran}F \times \text{Ran}G$ .*

*Por otra parte, si  $C$  es una cópula y  $F$  y  $G$  son funciones de distribución entonces  $H$  como se define en (1.6) es una función de distribución conjunta con marginales  $F$  y  $G$ .*

*Demostración:*

1. De acuerdo al Lema 1.3.1, existe una única subcópula  $C'$  tal que  $H(x, y) = C'(F(x), G(y))$  con  $\text{Dom}C' = \text{Ran}F \times \text{Ran}G$ . Por el Lema 1.3.2  $C'$  puede extenderse a una cópula  $C$ . Si  $F$  y  $G$  son continuas, entonces  $\text{Ran}F = \text{Ran}G = I$  y la única sub-cópula existente es la cópula  $C$ . Si las funciones no son continuas, entonces el rango de ellas no necesariamente es  $I$ , por lo cuál su extensión a partir del Lema 1.3.2 existe pero no es única.
2. Sea  $C$  una cópula y  $F$  y  $G$  funciones de distribución y defínase  $H(x, y) := C(F(x), G(y))$  y como se observa,  $H$  hereda las características de  $C$  puesto que por una parte para cualquier rectángulo  $B$   $V_H \geq 0$ . Por otra parte, tenemos  $H(-\infty, y) = G(y)$  también  $H(x, -\infty) = F(x)$  y  $H(\infty, \infty) = 1$  y por tanto  $H$  es función de distribución conjunta con marginales  $F$  y  $G$ .

□

Veamos ahora un resultado que provee de una poderosa herramienta para la obtención de cópulas y que en términos generales extrae la cópula de la función de distribución conjunta cuya existencia se garantiza en el primer punto del Teorema de Sklar. Para ello, únicamente es necesario el conocimiento de  $H$ , ya que a partir de ella se pueden obtener  $F$  y  $G$  y en un segundo nivel las cuasi-inversas de ellas.

**Corolario 1.3.1** *Sean  $H$ ,  $F$ ,  $G$ ,  $C$  como en el Teorema 1.3.1 y sean  $F^{(-1)}$  y  $G^{(-1)}$  las cuasi-inversas de  $F$  y  $G$ , respectivamente. Entonces para todo  $(u, v) \in \text{Dom}C$ .*

$$C(u, v) = H(F^{(-1)}(u), G^{(-1)}(v)) \quad (1.7)$$

*Demostración:*

$$H(F^{(-1)}(u), G^{(-1)}(v)) = C[F(F^{(-1)}(u)), G(G^{(-1)}(v))];$$

si  $(u, v) \in \text{Dom}C = \text{Ran}F \times \text{Ran}G$   
entonces tenemos que

$$H[F^{(-1)}(u), G^{(-1)}(v)] = C(u, v)$$

□

Ejemplifiquemos ahora el procedimiento descrito por el corolario y el teorema anterior.

### Ejemplo 1.3.1

Sea  $H$  función de distribución conjunta definida como

$$H(x, y) = [1 + e^{-x} + e^{-y} + (1 - \theta)e^{-x-y}]^{-1} \quad \theta \in [-1, 1]$$

Sabemos que:

$$F(x) = H(x, \infty) = \frac{1}{1 + e^{-x}} \text{ y también } G(y) = H(\infty, y) = \frac{1}{1 + e^{-y}}$$

es decir, distribuciones logísticas estandar; las cuasi-inversas están dadas por

$$x = -\log(u^{-1} - 1) \text{ y } y = -\log(v^{-1} - 1)$$

respectivamente.

Ahora, de acuerdo al Corolario 1.3.1:

$$\begin{aligned} C(u, v) &= H(-\log(u^{-1} - 1), -\log(v^{-1} - 1)) \\ &= \left[ \frac{1}{u} + \frac{1}{v} - 1 + (1 - \theta) \left( \frac{1}{u} - 1 \right) \left( \frac{1}{v} - 1 \right) \right]^{-1} \\ &= \left[ \frac{u + v - uv + (1 - \theta)(1 - u - v + uv)}{uv} \right]^{-1} \\ C(u, v) &= \frac{uv}{1 - \theta(1 - u)(1 - v)} \quad \theta \in [-1, 1] \end{aligned} \quad (1.8)$$

es la cópula que ata las marginales anteriores.

Ahora bien, para ejemplificar el segundo punto del Teorema de Sklar, incorporemos las siguientes marginales a la estructura anterior.

$$F_1(x) = \begin{cases} 0 & x < -1 \\ (x + 1)/2 & x \in [-1, 1] \\ 1 & x > 1 \end{cases}$$

$$G_1(y) = \begin{cases} 0 & y < 0 \\ 1 - e^{-y} & y \geq 0 \end{cases}$$

Hacemos:  $H_1(x, y) = C(F_1(x), G_1(y))$  obtenemos

$$H_1(x, y) = \begin{cases} \frac{(x+1)(e^y-1)}{x+2e^y-1} & (x, y) \in [-1, 1] \times [0, \infty]; \\ 1 - e^{-y} & (x, y) \in [1, \infty] \times [0, \infty]; \\ 0 & \text{otro caso;} \end{cases}$$

que es una función de distribución con marginales  $F_1$  y  $G_1$  que aunque difiere notablemente de la primera distribución conjunta, comparte con esta la misma estructura  $C$ . A la cópula presentada en (1.8) se le conoce como *familia Ali-Makhail-Haq* y más adelante nos será de utilidad.

Cerramos esta sección con otro ejemplo importante. En él extendemos una cópula cualquiera en su dominio para conformar una distribución conjunta.

### Ejemplo 1.3.2

Sea  $H$  definida como

$$H(x, y) = \begin{cases} C(x, y) & (x, y) \in \mathbf{I}^2; \\ x & x \in \mathbf{I}, y > 1; \\ y & y \in \mathbf{I}, x > 1; \\ 1 & x > 1, y > 1; \\ 0 & \text{otro caso;} \end{cases}$$

Como se observa,  $H$  es una distribución conjunta con marginales uniformes. En general, esta representación nos facilitará las cosas en la siguiente sección pues nos permite ver a las cópulas como restricciones a  $\mathbf{I}^2$  de funciones de distribución con marginales uniformes.

## 1.4. Cópulas y variables aleatorias continuas

Hasta el momento no se ha mencionado mucho sobre la relación de las cópulas con variables aleatorias. Sin embargo, si se consideran a  $F$  y  $G$  distribuciones de las variables aleatorias continuas  $X$  y  $Y$  y a  $H$  como su función de distribución conjunta, los resultados anteriores se mantienen por los razonamientos del Ejemplo 1.1.3.

Es prudente decir que en nuestra revisión sobre las relaciones de las cópulas con variables aleatorias, además de restringirnos al caso bivariado, nos dedicaremos a variables continuas debido a que de acuerdo con el Teorema de Sklar las cópulas obtenidas serán únicas y esto representa una ventaja cuando se trata de caracterizar variables observadas. Si bien esto puede representar un impedimento cuando se desea trabajar con variables aleatorias discretas, en nuestro caso no es así debido a las aplicaciones por presentar.

Cabe mencionar que en esta sección no únicamente se reexpresan los resultados anteriores, sino que también se busca profundizar en las implicaciones de ellos para clarificar el sentido de su empleo en la modelación de procesos aleatorios bivariados. Iniciemos con el resultado principal de la sección anterior.

**Teorema 1.4.1 (Sklar)** Sean  $X$  y  $Y$  variables aleatorias con funciones de distribución  $F$  y  $G$  respectivamente y función de distribución conjunta  $H$ . Entonces existe una cópula  $C$  tal que la igualdad 1.6 se mantiene y si  $F$  y  $G$  son continuas  $C$  es única, de otra manera  $C$  sólo es única en  $\text{Ran}F \times \text{Ran}G$

En adelante a  $C$  en el Teorema 1.4.1 la denominaremos *cópula de  $X$  y  $Y$*  y la denotaremos por  $C_{XY}$  cuando nos refiramos a variables aleatorias. Como la reexpresión del Corolario 1.3.1 se hace de manera muy similar, se omitirá y en su lugar, bosquejaremos las implicaciones de estos en términos de variables aleatorias, para lo que nos bastarán en un inicio las siguientes observaciones:

1. Si se conoce a  $H$  distribución conjunta de  $X$  y  $Y$  variables aleatorias continuas, es posible extraer a  $C_{X,Y}$  bajo el procedimiento descrito en el ejemplo 1.3.1.
2. Si por el contrario se conoce  $C$  y se incorporan  $F$  y  $G$  conocidas se obtiene una distribución conjunta  $H$ .

Ejemplifiquemos las observaciones anteriores considerando a un par de variables aleatorias con distribución normal bivariada.

**Ejemplo 1.4.1** Considérese a  $(X, Y) \sim N(x, y; 0, 1, r_{XY})$ ;  $X \sim N(0, 1)$  y  $Y \sim N(0, 1)$

$$H(x, y) = \int_{-\infty}^x \int_{-\infty}^y \frac{1}{2\pi\sqrt{1-r_{XY}^2}} \exp \frac{r_{X,Y}st - s^2 - t^2}{2(1-r_{X,Y}^2)} dsdt$$

La cuasi-inversa de  $X$  es  $\Phi^{-1}$  al igual que la de  $Y$ . Por el corolario anterior

$$\begin{aligned} C(u, v) &= H(\Phi^{-1}(u), \Phi^{-1}(v)) \\ \Rightarrow C(u, v) &= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} c(r) e^{c(s,t)} dsdt \end{aligned} \quad (1.9)$$

$$\text{con } c(r) = \frac{1}{2\pi\sqrt{1-r_{XY}^2}} \text{ y } c(s, t) = \frac{r_{XY}st - s^2 - t^2}{2(1-r_{XY}^2)}$$

La expresión en (1.9) es la *cópula gaussiana* y representa la estructura de dependencia existente en una normal bivariada. Para hacer esto un poco más claro, supóngase a priori que la distribución conjunta de dos variables es gaussiana, sin embargo, una de ellas tiene comportamiento exponencial y la otra de una distribución F de Fisher. Lo único que restaría por hacer para modelar este fenómeno sería evaluar en la función anterior  $u = F_1(x)$  exponencial y  $v = G_1(y)$  distribución F.

Como se habrá notado, el fenómeno descrito en el ejemplo anterior es bastante raro. Su presentación tiene por objeto mostrar la variedad de fenómenos que pueden modelarse a partir del empleo de cópulas y distribuciones continuas, pero sobre todo mostrar que son precisamente las cópulas las que describen como se relacionan las variables aleatorias, es decir, son estructuras de dependencia para cualesquiera distribuciones marginales.

El hecho de que las características de dependencia entre variables aleatorias en una distribución conjunta puede 'leerse' en la cópula de  $X$  y  $Y$  tiene fundamento directo en el Teorema de Sklar pues, como se ha mencionado, la cópula ata a las distribuciones marginales. En sí, el segundo capítulo de esta presentación versará sobre ello; aquí sólo se presenta un primer acercamiento.

En el siguiente resultado se presenta una importante relación de las cópulas con la estructura de dependencia notando que la ausencia de ella -es decir, la independencia- queda caracterizada en términos de una cópula que ya hemos mostrado con anterioridad.

**Teorema 1.4.2** Sean  $X$  y  $Y$  variables aleatorias con función de distribución conjunta  $H$  y marginales  $F$  y  $G$  respectivamente y copula  $C$ , entonces

$$C_{XY} = \Pi \Leftrightarrow X \perp Y \quad (1.10)$$

*Demostración:*

$$X \perp Y \Leftrightarrow H(x, y) = F(x)G(y)$$

por el Corolario 1.3.1

$$C(u, v) = H(F^{-1}(u), G^{-1}(v))$$

de donde se sigue que

$$X \perp Y \Leftrightarrow C(u, v) = F(F^{-1}(u))G(G^{-1}(v))$$

$$X \perp Y \Leftrightarrow C(u, v) = uv$$

□

Revisemos ahora otro importante resultado vinculado con las cotas de las cópulas  $W$  y  $M$  ahora en términos de dependencia.

### 1.4.1. Cotas de Frechét-Hoeffding para distribuciones conjuntas

Para  $X$  y  $Y$  variables aleatorias con función de distribución conjunta  $H$  y marginales  $F$  y  $G$ , tenemos:

$$\begin{aligned} W(x, y) &= \max\{F(x) + G(y) - 1, 0\} \\ &\leq H(x, y) = C[F(x), G(y)] \\ &\leq \min\{F(x), G(y)\} = M(x, y) \end{aligned} \quad (1.11)$$

La sencillez con que llegamos a estos resultados podría hacernos insensibles a su verdadero significado, para evitarlo, hagamos algunas breves observaciones que de hecho nos motivarán a explorar otras características de las cópulas.

La expresión en (1.11) implica que cualquiera que sea la distribución conjunta de las variables  $X$  y  $Y$  esta queda acotada en términos de sus marginales mientras que (1.10) muestra que la cópula producto caracteriza completamente la independencia.

A partir de esto, una interrogante natural y muy importante es: ¿qué clase de dependencia representaría que  $C_{XY}$  coincidiera con las cotas superior o inferior?.

**Teorema 1.4.3** Sean  $X, Y$  variables aleatorias con cópula  $C_{XY}$ , entonces :

$$i) C_{XY} = M \Leftrightarrow P[X > x, Y \leq y] = 0, \text{ o bien } P[X \leq x, Y > y] = 0$$

$$ii) C_{XY} = W \Leftrightarrow P[X > x, Y > y] = 0, \text{ o bien } P[X \leq x, Y \leq y] = 0$$

*Demostración:*

i) Por una parte :

$$\begin{aligned} H(x, y) &= P[X \leq x, Y \leq y] = P[X \leq x] - P[X \leq x, Y > y] \\ &= F(x) - P[X \leq x, Y > y] \end{aligned}$$

también

$$\begin{aligned} H(x, y) &= P[X \leq x, Y \leq y] = P[Y \leq y] - P[Y \leq y, X > x] \\ &= G(y) - P[Y \leq y, X > x] \end{aligned}$$

esto implica que

$$\begin{aligned} H(x, y) = F(x) &\Leftrightarrow P[X \leq x, Y > y] = 0 \\ &\Leftrightarrow G(y) > F(x) = H(x, y) \\ H(x, y) = G(y) &\Leftrightarrow P[Y \leq y, X > x] = 0 \\ &\Leftrightarrow G(y) \leq F(x) = H(x, y) \end{aligned}$$

por lo tanto

$$H(x, y) = \min\{F(x), G(y)\} \Leftrightarrow P[X > x, Y \leq y] = 0, \text{ ó } P[X \leq x, Y > y] = 0$$

ii) Del punto anterior

$$F(x) = H(x, y) + P[X \leq x, Y > y] = H(x, y) + (1 - G(y)) - P[Y > y, X > x]$$

por lo que

$$\begin{aligned} H(x, y) = F(x) + G(y) - 1 &\Leftrightarrow P[Y > y, X > x] = 0 \\ &\Leftrightarrow F(x) + G(y) - 1 > 0 \end{aligned}$$

también

$$\begin{aligned} H(x, y) = 0 &\Leftrightarrow P[X \leq x, Y \leq y] = 0 \\ &\Leftrightarrow F(x) + G(y) - 1 \leq 0 \end{aligned}$$

por lo tanto

$$H(x, y) = \max\{F(x) + G(y) - 1, 0\} \Leftrightarrow P[Y > y, X > x] = 0, \text{ o } P[X \leq x, Y \leq y] = 0$$

□

Detengámonos un poco para comprender qué significan los resultados anteriores. Tomemos por ejemplo la primera condición que garantiza que  $W$  sea la cópula de dos variables, es decir, que  $P[X > x, Y > y] = 0$  y consideremos que  $X$  y  $Y$  son uniformes en el intervalo unitario.

En palabras, para un par  $(x_0, y_0)$  fijo se garantiza que las variables aleatorias no pueden tomar simultáneamente valores más grandes, lo que equivale a la región  $\{(x, y) : x > x_0, y > y_0\}$ . Lo anterior se traduce en este caso en un rectángulo con vértices en  $(x_0, y_0)$ ,  $(x_0, 1)$ ,  $(1, y_0)$  y  $(1, 1)$ .

Como también se debe cumplir la condición  $P[X \leq x, Y \leq y] = 0$ , las variables no pueden tomar simultáneamente valores más pequeños que  $(x_0, y_0)$ , lo que ahora nos garantiza que las variables  $X$  y  $Y$  no tom valores en el rectángulo con vértices  $(x_0, y_0)$ ,  $(x_0, 0)$ ,  $(0, y_0)$  y  $(0, 0)$ .

De esta forma, para el par  $(x_0, y_0)$  fijo, sólomente los eventos 'cruzados'; es decir, los eventos  $(X > x_0 \text{ y } Y \leq y_0)$  y  $(X \leq x_0 \text{ y } Y > y_0)$  tienen probabilidad distinta de cero. Si ahora hacemos variar  $(x_0, y_0)$  vemos que los eventos con probabilidad distinta a cero deben estar comprendidos en la recta que une los puntos  $(0, 1)$  y  $(1, 0)$ .

Mediante razonamientos similares, notaremos que los eventos que tienen probabilidad distinta de cero son los cercanos a la recta que une los puntos  $(1, 0)$  y  $(0, 1)$ .

A continuación mostramos una representación de las cotas para marginales uniformes obtenidas mediante simulación.

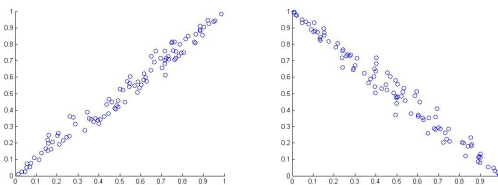


Figura 1.4: Representaciones de  $C_{X-X}$  y  $C_{X-X}$  para marginales uniformes obtenidas mediante simulación Monte Carlo

Como consecuencia de las observaciones anteriores, podemos probar que:

1.  $C_{X-X} = M$  puesto que  $P[X \leq x, X > x] = 0$
2.  $C_{X-X} = W$  dado  $P[X \leq x, -X \leq x] = 0$  y también  $P[X > x, -X > x] = 0$

Gracias a lo anterior, hemos resuelto casi por completo la interrogante planteada. Podemos pensar que (1) se da si  $Y$  es una función monótona creciente de  $X$  mientras que (2) se da cuando  $Y$  sea función monótona decreciente de  $X$ . Para formalizar estas aseveraciones, necesitamos presentar dos resultados que serán de utilidad cuando deseamos conocer la distribución conjunta de  $\alpha(X)$  y  $\beta(Y)$  una vez que conocemos la de  $X$  y  $Y$  con  $\alpha$  y  $\beta$  funciones monótonas.

El primero de los resultados muestra la invarianza de las cópulas ante transformaciones monótonas crecientes mientras que el segundo establece relaciones

entre las variables transformadas y la cópula de las variables previa su transformación.

**Teorema 1.4.4** Sean  $X, Y$  variables aleatorias continuas con cópula  $C_{XY}$ . Sean  $\alpha, \beta$  transformaciones estrictamente crecientes en  $\text{Ran}X$  y  $\text{Ran}Y$  respectivamente. Entonces  $C_{XY}$  es invariante con estas transformaciones y

$$C_{\alpha(X)\beta(Y)} = C_{XY}$$

*Demostración:*

Sean  $F_1, G_1, F_2, G_2$  las distribuciones de  $X, Y, \alpha(X), \beta(Y)$ ; por ser estrictamente crecientes  $F_2(x) = P[\alpha(X) \leq x] = P[X \leq \alpha^{-1}(x)] = F_1(\alpha^{-1}(x))$  y similar para  $G_2$

$$\begin{aligned} C_{\alpha(X)\beta(Y)}(F_2(x), G_2(y)) &= P[\alpha(X) \leq x, \beta(Y) \leq y] \\ &= P[X \leq \alpha^{-1}(x), Y \leq \beta^{-1}(y)] \\ &= C_{XY}(F_1(\alpha^{-1}(x)), G_1(\beta^{-1}(y))) \\ &= C_{XY}(F_2(x), G_2(y)) \end{aligned}$$

Como  $X$  y  $Y$  son continuas y  $\text{Ran}F_2 = \text{Ran}G_2 = \mathbf{I}^2$  se sigue que  $C_{\alpha(X)\beta(Y)} = C_{XY}$

□

**Teorema 1.4.5** Sean  $X, Y$  variables aleatorias continuas con cópula  $C_{XY}$ . Sean  $\alpha, \beta$  transformaciones monótonas en  $\text{Ran}X$  y  $\text{Ran}Y$  respectivamente. Entonces:

1. Si  $\alpha$  es estrictamente creciente y  $\beta$  es estrictamente decreciente

$$C_{\alpha(X)\beta(Y)}(u, v) = u - C_{XY}(u, 1 - v)$$

2. Si  $\alpha$  es estrictamente decreciente y  $\beta$  es estrictamente creciente

$$C_{\alpha(X)\beta(Y)}(u, v) = v - C_{XY}(1 - u, v)$$

3. Si  $\alpha$  y  $\beta$  son estrictamente decrecientes

$$C_{\alpha(X)\beta(Y)}(u, v) = u + v - 1 + C_{XY}(1 - u, 1 - v)$$

□

La prueba de este resultado sigue razonamientos similares al del Teorema 1.4.4 y puede encontrarse en Erdely. <sup>8</sup>

El siguiente resultado formaliza lo que veíamos gráficamente y su prueba se realiza en la misma línea de los dos teoremas anteriores.

---

<sup>8</sup>Erdely, Op. Cit. p22



**Corolario 1.4.1** *Sea  $\alpha$  una función monótona tal que  $\alpha(X) = Y$  casi seguramente<sup>9</sup>. Entonces:*

- i) Si  $\alpha$  es estrictamente creciente, entonces  $C_{XY} = M$*
- ii) Si  $\alpha$  es estrictamente decreciente, entonces  $C_{XY} = W$*

□

A partir de lo anterior podemos realizar una importante aseveración: como la cópula de dos variables continuas caracteriza de manera única la estructura de dependencia existente entre ellas y las cotas de las cópulas son  $M$  y  $W$ , los 'grados' de dependencia extremos se darán únicamente cuando una de las variables sea función monótona de la otra casi seguramente.

### 1.4.2. Orden

Además de que las cópulas permiten la caracterización mostrada con anterioridad, puede establecerse una relación de orden equivalente a la llamada dominación estocástica. Recordando esta característica, si el dominio de ambas coincide decimos que  $H_1$  domina a  $H_2$  si  $H_1 \leq H_2$  para todo elemento en  $Dom H_1$

De manera similar, definimos la relación de orden  $\prec$  entre cópulas como sigue:

$$C_1 \prec C_2 \Leftrightarrow C_1(u, v) \leq C_2(u, v), \text{ para todo } (u, v) \in \mathcal{I}^2$$

y decimos que la cópula  $C_1$  es menor que la cópula  $C_2$  (o bien que  $C_2$  es mayor que  $C_1$ ).<sup>10</sup>

Cuando no se pueda establecer esta relación de orden entre dos copulas, se dirá simplemente que no son comparables. Ejemplo de esto sería tratar de comparar  $\Pi$  con  $\alpha M + (1 - \alpha)\Pi$  (definida en el Ejemplo 1.2.2 con  $\beta = 0$ ); claramente habrá valores que hagan a  $\Pi$  mayor o menor que la otra cópula.

Una de las utilidades de esta relación radica en permitirnos reducir nuestro espacio de estudio cuando estamos indagando sobre alguna estructura de dependencia. Por ejemplo, si desconocemos a  $C_{XY}$  pero sabemos que verifica  $C_{XY} \prec \Pi$  sabremos que  $C_{XY}$  está 'más cercana' a  $W$  que a  $M$ , es decir, que la relación existente entre  $X$  y  $Y$  se asemeja más a una transformación creciente que a una decreciente y a partir de ello podemos elegir a una cópula que considere este tipo de dependencia. Por el contrario, si sabemos que  $C_{XY}$  es no comparable con  $\Pi$  realizaremos nuestro estudio dentro de las familias comprensivas.

Por el momento dejaremos hasta aquí esta observación ya que este tipo de características se revisarán en el siguiente capítulo.

Además, a partir de  $\prec$  podemos hacer la siguiente precisión: a una familia de cópulas parametrizada en  $\theta \{C_\theta\}$  se le llamará *ordenada positivamente* si  $C_\alpha \prec C_\beta$  cuando  $\alpha \leq \beta$  y *ordenada negativamente* si  $C_\alpha \succ C_\beta$  cuando  $\alpha \geq \beta$ .

<sup>9</sup>Para comprender este concepto bastará mencionar que la convergencia casi segura garantiza que la probabilidad de que la diferencia  $\alpha(X) - Y$  sea pequeña es casi cero. Para una revisión más profunda del concepto se recomienda consultar Bhat. B. *Modern probability theory*. 2a Ed. John Wiley & Sons

<sup>10</sup>Sabemos por ejemplo que:  $W \prec \Pi \prec M$  se cumple de manera directa.

### 1.4.3. Cópula supervivencia

En una gran cantidad de aplicaciones, las variables aleatorias representan 'el tiempo de vida' de algún objeto o incluso un individuo. Para ello, considerando la distribución conjunta  $H$ , la probabilidad de que ambos objetos sigan con 'vida' hasta al menos  $x$  y  $y$  respectivamente es  $P[X > x, Y > x] = \overline{H}(x, y)$ . Naturalmente  $\overline{H}(x, -\infty) = \overline{F}(x)$  y  $\overline{H}(-\infty, y) = \overline{G}(y)$ . Nótese que:

$$\begin{aligned} H(x, y) &= P[X \leq x, Y \leq y] = P[X \leq x] - P[X \leq x, Y > y] \\ &= P[X \leq x] + P[Y > y] - P[X > x, Y > y] \\ &= F(x) - (1 - G(y)) - \overline{H}(x, y) \end{aligned}$$

Lo anterior implica que :

$$\overline{H}(x, y) = \overline{F}(x) + \overline{G}(y) - 1 + H(x, y)$$

recordando el teorema de Sklar y haciendo  $u = \overline{F}(x)$ ,  $v = \overline{G}(y)$

$$\overline{H}(x, y) = u + v - 1 + C(1 - u, 1 - v)$$

definimos la *cópula supervivencia* como

$$\widehat{C}(u, v) = u + v - 1 + C(1 - u, 1 - v) \quad (1.12)$$

**Ejemplo 1.4.2** Para ejemplificar lo anterior, consideremos la cota superior cuya representación es bastante simple  $M(u, v) = \min(u, v)$  y de acuerdo con lo anterior:

$$\begin{aligned} \widehat{M}(u, v) &= u + v - 1 + C(1 - u, 1 - v) \\ &= u + v - 1 + \min(u, v) \end{aligned}$$

Otro ejemplo fácil de seguir es considerar la *cópula producto*, donde de hecho observamos que  $\Pi(u, v) = \widehat{\Pi}(u, v)$ .

Además de los ejemplos anteriores, a modo de comentario, es importante mencionar que la familia paramétrica Frank -que pertenece al conjunto de cópulas arquimedeanas sobre las que hablaremos más adelante- cumple con que su cópula supervivencia es ella misma:

$$C_\theta(u, v) = \widehat{C}_\theta(u, v) = -\frac{1}{\theta} \left[ 1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right] \quad (1.13)$$

### 1.4.4. Distribuciones condicionales a partir de cópulas

Al igual que el tiempo de supervivencia, en ocasiones se estará interesado en saber acerca de las distribuciones condicionales de las variables una vez que se conoce  $H$ . Puesto que las cópulas proveen la información acerca de cómo interactúan las variables en cuestión, las distribuciones condicionales pueden obtenerse fácilmente de  $C_{XY}$ . Antes de mostrar como se obtienen estas distribuciones, recordemos un importante resultado de probabilidad que nos será de utilidad de aquí en adelante. <sup>11</sup>

<sup>11</sup>Mood, A., Graybill, F., Boes, D. (1974) *Introduction to the Theory of Statistics*. McGraw-Hill. p202

**Teorema 1.4.6** (De la integral de probabilidad transformada)

Si  $X$  es una variable aleatoria con función de distribución  $F$  tenemos que  $F(X) \sim U(0, 1)$ , debido a esto, para una variable  $U$  uniforme  $F^{(-1)}(U) \sim X$ .

□

Dicho esto, supongamos sin pérdida de generalidad que nuestra función de distribución conjunta es precisamente  $C$  con marginales uniformes,

$$\begin{aligned} P[V \leq v \mid U = u] &= \lim_{\Delta u \rightarrow 0} \frac{P[V \leq v, U \leq u + \Delta u] - P[V \leq v, U \leq u]}{\Delta u} \\ &= \lim_{\Delta u \rightarrow 0} \frac{C(u + \Delta u, v) - C(u, v)}{\Delta u} \end{aligned}$$

luego,

$$P[V \leq v \mid U = u] = \frac{\partial}{\partial u} C(u, v)$$

y de manera similar,

$$P[U \leq u \mid V = v] = \frac{\partial}{\partial v} C(u, v)$$

**Ejemplo 1.4.3** Para la familia Frank por ejemplo:

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

y

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

Notamos que la obtención de las distribuciones condicionales es bastante simple. Esta es una de las ventajas de aislar la cópula de la distribución conjunta. Para evaluar la distribución condicional en general, una vez que se han obtenido las derivadas parciales para la cópula, bastará sustituir las marginales en las componentes y multiplicar por la densidad de la variable condicionada.

### 1.4.5. C-medida

Aunque los aspectos revisados hasta el momento pueden expresarse en términos de espacios probabilísticos, hasta ahora hemos evitado tomar este camino ya que los conceptos eran alcanzables sin necesidad de ello. Sin embargo, dada la importancia de los siguientes conceptos en el desarrollo de las secciones siguientes 1.5 y 1.6 tomaremos momentáneamente este rumbo ya que será de bastante utilidad posteriormente no sin antes presentar una aproximación intuitiva de los conceptos teóricos que se expondrán.

La teoría de la medida -como su nombre lo indica- se dedica al estudio de 'medidas'. Conceptualmente una medida es una función -aditiva y no negativa definida para todos los elementos de un conjunto- asignan valores determinados a cada uno de los elementos del conjunto donde está definida. La probabilidad es de hecho un ejemplo de medida que cumple con la característica adicional de estar acotada entre cero y uno.

En general, cualquier función de distribución de probabilidad induce una medida de probabilidad al asignar a cada elemento del dominio un valor en el intervalo unitario. De igual forma, una función de distribución bivariada induce una medida de probabilidad asignando a pares de eventos un valor numérico también en el intervalo unitario. Más aún, si vemos a las cópulas como funciones de distribución bivariadas con marginales uniformes <sup>12</sup>, las cópulas también inducen una medida de probabilidad.

Otro aspecto que será de interés es el concepto del soporte de una función de probabilidad. Intuitivamente, el soporte de una función 'carga' con todo el peso de la distribución y está comprendido por los 'lugares' que tienen valor para ella es decir, que tienen medida de probabilidad distinta de cero.

Formalmente, en  $(\Omega, F, P)$  espacio de probabilidad, el vector aleatorio  $(X, Y)$  es una función  $F$ -medible  $(X, Y) : \Omega \rightarrow \mathbb{R}^2$  e induce una medida de probabilidad  $P_{XY}$  en  $(\mathbb{R}^2, B(\mathbb{R}^2))$  espacio medible.

A partir de ello diremos que la función de distribución conjunta  $H$  induce una medida de probabilidad a través del volumen  $V_H$  y a su vez, una cópula  $C$  induce una medida de probabilidad en  $\mathbf{I}^2$  a través de  $V_C$  y la  $C$ -medida de un conjunto en  $\mathbf{I}^2$  quedará dada por su  $C$ -volumen.

Por otra parte, para una cópula  $C$  cualquiera, podemos escribir:

$$C(u, v) = A_C(u, v) + S_C(u, v)$$

donde

$$A_C(u, v) = \int_0^u \int_0^v \frac{\partial^2}{\partial s \partial t} C(s, t) dt ds$$

es la parte absolutamente continua <sup>13</sup> y

$$S_C(u, v) = C(u, v) - A_C(u, v)$$

es la parte singular de la cópula.

Al igual que como se define para una función de distribución, diremos que *el soporte de una cópula  $C$  es el complemento de la unión de todos los subconjuntos de  $\mathbf{I}^2$  que tengan  $C$ -medida cero*, o bien, el conjunto que incluye sólo subconjuntos con probabilidad positiva.

#### Ejemplo 1.4.4

Por ejemplo, para las cópulas  $W$  y  $M$  su soporte está dado por los conjuntos:

$$S_W = \{(u, v) \in \mathbf{I}^2 \mid u = v\} \text{ y } S_M = \{(u, v) \in \mathbf{I}^2 \mid v = 1 - u\}$$

lo cual es un tanto engorroso de probar pues hay que calcular varios  $C$ -volúmenes.

Sin embargo, independiente de esto, podemos afirmar que *son singulares* al notar que las segundas derivadas de estas cópulas son iguales a cero excepto en cada uno de los conjuntos especificados, es decir, el soporte de la cópula

<sup>12</sup>Ver ejemplo 1.3.2

<sup>13</sup>Todos aquellos puntos  $u$  y  $v$  donde su derivada no es cero i.e.  $\partial^2 C(u, v) / \partial u \partial v \neq 0$

quedará descrito completamente por los conjuntos señalados.

Ahora, examinemos el soporte de la cópula producto. A esta la podemos escribir como:

$$\Pi(u, v) = \int_0^u \int_0^v 1 \, dt ds = uv$$

en este caso diremos que la cópula tiene *soporte completo*, es decir, tiene como soporte a todo  $\mathbf{I}^2$  y además, que es *absolutamente continua* al poder recuperarse únicamente a partir de  $A_C$ .

### 1.4.6. Simulación

Para cerrar con esta sección, abordamos el tema de la simulación que es de suma importancia cuando se habla de variables aleatorias. En términos generales, la simulación es una técnica numérica que permite reproducir en forma aproximada algún evento real para su estudio.

En nuestro caso, en esta sección, estaremos interesados en obtener observaciones provenientes de una cópula, ya sea para modelar algún fenómeno o propiamente para estudiar el comportamiento de estas funciones.

Los métodos existentes tanto para la simulación de variables aleatorias como de observaciones pertenecientes a una cópula son variados. Para las primeras puede revisarse casi cualquier texto de simulación destacando el siguiente método que emplea el Teorema 1.4.6:

#### Algoritmo 1

Para obtener una observación de la variable aleatoria  $X$  con función de distribución  $F$  con cuasi-inversa  $F^{(-1)}$ :

1. Generar  $u$  observación uniforme estandar
2. Tomar  $x = F^{(-1)}(u)$
3.  $x$  es una observación de la variable aleatoria  $X$

Notamos que el algoritmo anterior asume que podemos generar observaciones provenientes de una distribución uniforme, lo cual no es nada abesado pues actualmente esto puede hacerse incluso a partir de una hoja de cálculo de Excel de Microsoft Office.

Para la simulación de observaciones provenientes de una cópula mostraremos un algoritmo que emplea las distribuciones condicionales.

#### Algoritmo 2

Para obtener un par  $(u, v)$  con cópula  $C$ <sup>14</sup>:

1. Generar  $u, t$  independientes y uniformes estandar
2.  $v = c_u^{(-1)}(t)$  con  $c_u^{(-1)}$  cuasi-inversa de  $c_u = \frac{\partial}{\partial u} C(u, v)$

---

<sup>14</sup>Nelsen, Op.Cit. p36

3. El par  $(u, v)$  es una observación de la cópula  $C$

Conceptualmente lo que hace el algoritmo, dado que se posee una observación uniforme, se genere otra con la característica de esté relacionada de la forma en que  $c_u$  indica. Revisemos un ejemplo.

**Ejemplo 1.4.5** *Si deseáramos simular observaciones provenientes de la familia Frank, restaría obtener la cuasi- inversa de alguna de las derivadas parciales en el ejemplo 1.4.2. Si hacemos esto con respecto a la derivada parcial de  $u$  obtenemos:*

$$c_u^{(-1)}(t) = -\frac{1}{\theta} \log\left[\frac{(1-t)e^{-\theta u}}{t} + \frac{e^\theta}{1 + e^{-\theta u} \frac{1-t}{t}}\right]$$

si se simulan  $u$  y  $t$  observaciones uniformes y se evalúan en la función anterior tomando  $v = c_u^{(-1)}(t)$ ,  $u$  y  $v$  siguen la distribución conjunta expresada en la ecuación (1.13).

## 1.5. Construcción de cópulas

Por el Teorema de Sklar sabemos que las características de la dependencia dentro de una distribución conjunta quedan descritas completamente a partir de las cópulas y como se mostrará en el siguiente capítulo también a partir de ellas es posible la construcción de diversas medidas de dependencia así como una reinterpretación de algunas características de dependencia en términos de cópulas. Debido a esto y en parte también por su utilidad en la generación de nuevas funciones de distribución conjuntas, resulta de interés y ventajoso contar con una gama de cópulas entre las que podamos elegir aquella que se ajuste mejor a nuestras necesidades.

En esta sección se busca responder a esta inquietud ilustrando básicamente dos estrategias para la construcción de cópulas. La primera de ellas ya ha sido explorada con anterioridad y corresponde al método de extracción o inversión presentado en los Ejemplos 1.3.1 y 1.4.1 que explota el Teorema de Sklar y su Corolario. El siguiente corresponde a un método geométrico para construir cópulas con soporte o secciones prescritas y también mediante sumas convexas. Finalmente, a partir de este último método inducimos la presentación de otro tipo de cópulas que por sus características tan especiales dejamos para la siguiente sección. Cabe mencionar que estas son solo algunas de las posibilidades existentes para la construcción de cópulas y si se desea profundizar en ello se recomienda la revisión del tercer capítulo de Nelsen <sup>15</sup>.

### 1.5.1. Método de inversión

Siguiendo la ruta de los ejemplos mencionados, presentamos otro más en el que se explota la definición de la cópula supervivencia.

#### Ejemplo 1.5.1

Consideremos ahora un artículo con dos componentes cuya descompostura repercute directamente en la falla del artículo y sean  $X$  y  $Y$  las variables aleatorias

---

<sup>15</sup>Nelsen Op.Cit. pp 45 a 82

que describen el tiempo de vida del componente 1 y el 2 respectivamente. Modelaremos las fallas sobre los componentes a partir de tres procesos Poisson independientes con parámetros  $\lambda_1, \lambda_2, \lambda_{12}$ .<sup>16</sup> A su vez, modelaremos los tiempos de ocurrencia de las averías anteriores con las variables aleatorias exponenciales  $Z_1, Z_2$  y  $Z_{12}$  respectivamente. La probabilidad de que los componentes 1 y 2 lleguen sin fallos a los tiempos  $x$  y  $y$  respectivamente quedará dada por:

$$\bar{H}(x, y) = P[X > x, Y > y] = P[Z_1 > x, Z_2 > y, Z_{12} > \max(x, y)]$$

que de acuerdo con nuestros supuestos podemos reescribir como:

$$\begin{aligned} \bar{H}(x, y) &= P[Z_1 > x] P[Z_2 > y] P[Z_{12} > \max(x, y)] \\ &= P[Z_1 > x] P[Z_2 > y] P[Z_{12} > x + y - \min(x, y)] \\ &= \exp[-\lambda_1 x - \lambda_2 y - \lambda_{12}(x + y - \min(x, y))] \end{aligned}$$

de donde podemos observar que:

$$\bar{F}(x) = \exp[-(\lambda_1 + \lambda_{12})x] \text{ y } \bar{G}(y) = \exp[-(\lambda_2 + \lambda_{12})y]$$

son las distribuciones de supervivencia marginales y tienen distribución exponencial. A partir de lo anterior, podemos reexpresar lo anterior como:

$$\begin{aligned} \bar{H}(x, y) &= \bar{F}(x) \bar{G}(x) \exp[(\lambda_{12} \min(x, y))] \\ &= \bar{F}(x) \bar{G}(x) \min[\exp(\lambda_{12} x), \exp(\lambda_{12} y)] \end{aligned}$$

Como las cuasi-inversas de  $\bar{F}$  y  $\bar{G}(x)$  son:

$$\bar{F}^{-1}(u) = \log u^{-(\lambda_1 + \lambda_{12})^{-1}} \text{ y } \bar{G}^{-1}(v) = \log v^{-(\lambda_2 + \lambda_{12})^{-1}}$$

respectivamente y usando el Corolario 1.3.1 se tiene que

$$\hat{C}(u, v) = \bar{H}(\bar{F}^{-1}(u), \bar{G}^{-1}(v)) = uv \min[u^{-\frac{\lambda_{12}}{\lambda_1 + \lambda_{12}}}, v^{-\frac{\lambda_{12}}{\lambda_2 + \lambda_{12}}}]$$

es la cópula supervivencia de  $X$  y  $Y$ .

Haciendo

$$\alpha = \frac{\lambda_{12}}{\lambda_1 + \lambda_{12}} \text{ y } \beta = \frac{\lambda_{12}}{\lambda_2 + \lambda_{12}}$$

llegamos a la llamada *familia Marshall-Olkin* o *familia Cuadras-Auge generalizada*

$$C(u, v) = uv \min[u^{-\alpha}, v^{-\beta}] \quad (1.14)$$

De esta forma hemos ilustrado nuevamente el funcionamiento de este método para la extracción de una cópula dentro de una distribución conjunta.

<sup>16</sup>El primer parámetro es sobre la avería de únicamente el primer componente, el segundo de ellos si se inutiliza sólo el segundo y el tercero para cuando fallan ambos simultáneamente.

### 1.5.2. Métodos geométricos

Como se mencionó, lejos de emplear el teorema de Sklar, la construcción geométrica de una cópula retoma la definición de cópula. Se construyen funciones que cumplan con ella a partir de información conocida sobre secciones diagonales, verticales u horizontales o sobre su soporte. Además, se explora la construcción a partir de sumas convexas no lineales. A continuación se muestran diversos ejemplos que muestran como se realiza la construcción.

#### Ejemplo 1.5.2 Cópulas con soporte dado.

Sean  $U$  y  $V$  variables aleatorias uniformes y  $\theta \in (0, 1)$  tales que  $P[V = U + \theta] = 1$  y sea  $C$  la cópula de estas variables.

Nótese que si  $U$  toma valores entre 0 y  $1 - \theta$ ,  $V$  tomará valores entre  $\theta$  y 1 con probabilidad 1; si  $U$  toma valores entre  $1 - \theta$  y 1,  $V$  tomará valores entre 0 y  $\theta$ . El soporte de  $C$  quedará descrito por las rectas en la figura 1.5.

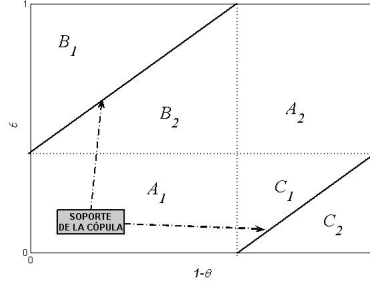


Figura 1.5: Soporte de la cópula de variables  $U$  y  $V$  tales que  $P[V = U + \theta] = 1$

Ahora, con el fin de obtener C-volumenes no nulos, siguiendo la partición de  $\mathbf{I}^2$  descrita en la figura anterior y utilizando las condiciones de frontera que debe cumplir una cópula:

- Sean  $(u, v) \in C_1$ , con  $0 > u + v - 1$ ,  
 $V_{[0,u] \times [0,v]} = 0 \Rightarrow C(u, v) = 0$
- Sean  $(u, v) \in C_2$  con  $0 \leq u + v - 1$ ,  
 $V_{[u,1] \times [v,1]} = 0 \Rightarrow C(u, v) = u + v - 1$
- Para  $u, v$  en la región  $A_1$ , con  $u \leq 1 - \theta$   
 $V_{[0,u] \times [0,v]} = V_{[0,u] \times [0,1]} = u \Rightarrow C(u, v) = u$
- Sean  $(u, v) \in A_2$ , con  $u > 1 - \theta$   
 $V_{[0,u] \times [0,v]} = V_{[0,1-\theta] \times [0,1]} = 1 - \theta \Rightarrow C(u, v) = 1 - \theta$
- Para  $u, v$  en la región  $B_1$ , con  $u - (1 - \theta) \leq v$   
 $V_{[0,u] \times [0,v]} = V_{[1-\theta,v] \times [0,1]} = u - (1 - \theta) \Rightarrow C(u, v) = u - (1 - \theta)$
- Para  $u, v$  en la región  $B_2$ , con  $u - (1 - \theta) > v$   
 $V_{[0,u] \times [0,v]} = V_{[0,1] \times [0,v]} = v \Rightarrow C(u, v) = v$



De lo anterior se deduce que :

$$C(u, v) = \begin{cases} \min[u - 1 + \theta] & (u, v) \in [0, 1 - \theta] \times [\theta, 1]; \\ \min[u, \theta - v] & (u, v) \in [1 - \theta, 1] \times [0, \theta]; \\ W(u, v) & \text{otro caso} \end{cases}$$

**Ejemplo 1.5.3** *Cópulas con secciones dadas*

Para ilustrar este procedimiento, supongámos que deseamos conocer las cópulas que sean lineales en  $u$ ; es decir, aquellas que cumplan con:

$$C(u, v) = a(v)u + b(v)$$

Empleando las condiciones de frontera,

$$C(0, v) = 0 \Rightarrow b(v) = 0; \text{ y también}$$

$$C(1, v) = v \Rightarrow a(v) = v; \text{ con lo que}$$

$$C(u, v) = uv = \Pi(u, v)$$

Si ahora aplicamos el mismo razonamiento para identificar las cópulas lineales en  $v$  se llega al mismo resultado lo que indica que la única cópula con sección vertical u horizontal lineal es la cópula producto.

Ahora considérese aquellas con sección cuadráticas en  $u$ , es decir las que pueden expresarse como:

$$C(u, v) = a(v)u^2 + b(v)u + c(v).$$

De manera similar que en la construcción anterior,  $c(v) = 0$  y  $a(v) + b(v) = v$ . Haciendo  $\psi(v) = -a(v)$ , tenemos  $b(v) = v - \psi(v)$ .

$$C(u, v) = -\psi(v)u^2 + (v - \psi(v))u = uv + \psi(v)(u(1 - u))$$

Con lo que podemos afirmar que las cópulas con sección vertical cuadrática quedarán descritas por las funciones  $\psi$  que cumplan con hacer el C-volumen no nulo y las condiciones de frontera. Una de las soluciones es

$$\psi(v) = \theta v(1 - v)$$

que da origen a la *familia Ferlie-Gumbel-Morgesten (FMG)* que incluye todas las cópulas con secciones cuadráticas en  $u$  y  $v$  simultaneamente.

$$C(u, v) = uv + \theta uv(1 - u)(1 - v) \tag{1.15}$$

De esta manera, pueden prescribirse diferentes secciones y resolver las ecuaciones funcionales resultantes para hallar las cópulas relacionadas. Analicemos ahora la generación por medio de sumas convexas.

**Ejemplo 1.5.4** *Sumas convexas y mezclas*

De acuerdo con la Proposición 1.2 es posible generar nuevas cópulas a partir de una suma convexa de cópulas. Ahora, consideraremos una variante sobre familias de cópulas parametrizadas en  $\theta$ . Para ello, asumamos que  $\theta$  corresponde a una

observación de la variable aleatoria real y continua  $\Theta$  con función de distribución paramétrica  $\Lambda_\alpha$ . Entonces la función

$$C\alpha(u, v) = \int_{\mathbf{R}} C_\theta(u, v) d\Lambda_\alpha(\theta)$$

es una cópula. Una variedad de funciones  $\Lambda$  pueden ser propuestas con el objeto de obtener nuevas familias de cópulas. Por ejemplo, la función de distribución conjunta:

$$H(u, v) = \int_0^\infty F^\theta(u)G^\theta(v)d\Lambda(\theta)$$

es una mezcla de funciones.

Si además consideramos a la transformada de Laplace

$$\psi(t) = \int_0^\infty e^{-\theta t} d\Lambda(\theta)$$

y las funciones de distribución marginales:

$$F(u) = \exp[-\psi^{-1}(u)] \text{ y a } G(v) = \exp[-\psi^{-1}(v)]$$

tenemos:

$$H(u, v) = \psi[\psi^{-1}(u) + \psi^{-1}(v)]$$

Puede probarse que las marginales son uniformes y entonces llegamos a que

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

es una cópula que, como se observa depende únicamente de la función univariada  $\psi$  además de ser simétrica. La función  $\psi$  es tan solo una de las tantas funciones que permiten definir de manera similar una cópula. En la siguiente sección se exploran estas funciones y las *cópulas Arquimedeanas* generadas a partir de ellas.

## 1.6. Cópulas Arquimedeanas

En esta última sección revisaremos las características de las Cópulas Arquimedeanas, que, como se mencionó con anterioridad tienen -entre varias características importantes- la peculiaridad de poder describirse en términos de una función univariada. El tema es bastante extenso y nos limitaremos a presentar los resultados que son de mayor trascendencia para nuestros fines.

Revisamos primero las características que debe tener un *generador*<sup>17</sup> para poder definir una cópula, las características de la cópula generada -incluyendo la función de distribución de probabilidad de esta- y finalmente algunos resultados mediante los cuales generaremos algoritmos para la simulación de observaciones también un listado de las familias pertenecientes a este grupo.

Antes de iniciar directamente con el tema de interés, necesitamos hacer una precisión sobre lo que denominaremos pseudo-inversa de una función real pues su empleo en adelante será necesario.

<sup>17</sup>Así le denominaremos a nuestra función univariada

**Definición 1.6.1** Sea  $\phi: I \rightarrow [0, \infty)$  continua,  $\phi(x_1) > \phi(x_2) \forall x_1 < x_2$  y además  $\phi(1) = 0$ . Su pseudo-inversa es  $\phi^{(-1)}: [0, \infty) \rightarrow I$  con

$$\phi^{(-1)}(t) = \begin{cases} \phi^{-1}(t) & 0 \leq t \leq \phi(0); \\ 0 & \phi(0) \leq t \leq \infty \end{cases} \quad (1.16)$$

donde  $\phi^{-1}(t)$  es la inversa de la función.

Con esto, podemos ver que si  $\phi(0) \rightarrow \infty$  entonces  $\phi^{(-1)} = \phi^{-1}$ . Definamos la siguiente función considerando a  $\phi$  y a  $\phi^{(-1)}$  como arriba

$$C(u, v) = \phi^{(-1)}[\phi(u) + \phi(v)] \quad (1.17)$$

### 1.6.1. Propiedades generales

Ahora, siguiendo un procedimiento similar al empleado en la sección 1.5.2 podemos preguntarnos que otras condiciones debe cumplir  $\phi$  de modo que  $C$  sea una cópula. Como se verá, nuevamente la característica 2-creciente es la menos accesible. Partiendo de la definición, necesitamos que:

1.  $\text{Dom}C = I^2$ ; que se verifica precisamente por como se han definido  $\phi$  y su cuasi-inversa
2.  $C$  fijada en cero; debemos tener que

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

como tanto  $\phi(u) + \phi(0)$  como  $\phi(0) + \phi(v) \in [\phi(0), \infty)$  por ser  $\phi$  positiva, con lo que efectivamente

$$C(u, 0) = C(0, v) = 0$$

3. Que  $C$  sea 2-creciente, i.e.

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

Para asegurar este último punto procederemos primero a observar las características de  $C$ , que implícitamente contempla las de  $\phi$  por su construcción. Si asumimos que  $C$  es 2-creciente, podemos ver que particularmente

$$V_C[u_1, u_2] \times [v, 1] \geq 0$$

es decir, que

$$C(u_2, v) - C(u_1, v) \leq u_2 - u_1 \quad \text{para toda } u_1, u_2, v$$

Ahora, tomando dos puntos cualesquiera que satisfagan  $v_1 \leq v_2$  notamos que

$$C(0, v_2) = 0 \leq v_1 \leq v_2 \leq C(1, v_2)$$

Si consideramos que  $v_2$  es un punto fijo,  $C(t, v_2)$  es una función de variable real y también es continua al ser  $\phi$  y su pseudo-inversa continuas. Luego, por el Teorema del Valor Medio existe  $t_0$  tal que  $C(t_0, v_2) = v_1$ . Por lo anterior :

$$C(u_2, v_1) - C(u_1, v_1) = C(u_2, C(t_0, v_2)) - C(u_1, C(t_0, v_2))$$

$$\begin{aligned} &= C(C(u_2, v_2), t_0) - C(C(u_1, v_2), t_0) \\ &\leq C(u_2, v_2) - C(u_1, v_2) \end{aligned}$$

y por tanto  $C$  es 2-creciente.

Los razonamientos anteriores nos dicen que  $C$  definida en (1.17)

$$C \text{ es 2-creciente} \Leftrightarrow C(u_2, v) - C(u_1, v) \leq u_2 - u_1$$

Ahora, en términos de  $\phi$ ,  $C$  esto se da si y solo si:

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

o de manera equivalente:

$$\phi^{(-1)}(a_1) + \phi^{(-1)}(a_2 + b) \leq \phi^{(-1)}(a_2) + \phi^{(-1)}(a_1 + b) \quad (1.18)$$

tomando  $\phi(u_i) = a_i$  y  $\phi(v) = b$  con  $a_1 \leq a_2$  y  $b \geq 0$ . La relación anterior nos llevan a identificar una característica importante de  $\phi$ :

Por una parte si (1.18) se cumple y hacemos  $a_1 = (s+t)/2$ ,  $a_2 = s$  y  $b = (s-t)/2$  llegamos a que:

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

con lo que  $\phi^{(-1)}$  es convexa al punto medio; por ser continua  $\phi^{(-1)}$  es convexa en todo su dominio y como consecuencia  $\phi$  también lo es.

Por otra, si asumimos que  $\phi$  es convexa y definimos  $\alpha = (a_1 - a_2)/(a_1 - a_2 + b)$  y de acuerdo a esto  $a_1 = (1 - \alpha)a_2 + \alpha(a_1 + b)$  y  $a_2 + b = \alpha(a_2 + (1 - \alpha)(a_1 + b))$  lo que nos lleva a:

$$\phi^{(-1)}(a_1) \leq (1 - \alpha)\phi^{(-1)}(a_2) + \alpha\phi^{(-1)}(a_1 + b)$$

y también:

$$\phi^{(-1)}(a_2 + b) \leq \alpha\phi^{(-1)}(a_2 + (1 - \alpha)\phi^{(-1)}(a_1 + b))$$

y claramente sumando ambas desigualdades llegamos a (1.18). Formalizando, presentamos el siguiente :

**Teorema 1.6.1** *Sea  $\phi$  continua, estrictamente decreciente y  $\phi(1) = 0$  con pseudo inversa definida en la igualdad (1.16). Entonces la función*

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

*es una cópula si y sólo si  $\phi$  es convexa.*

□

Diremos que  $\phi$  es el *generador* de  $C$  y si  $\phi^{(-1)} = \phi^{-1}$  que  $C$  es *estricta*. Revisemos ahora un resultado interesante gracias al cual las cópulas que pueden escribirse como en el teorema anterior resiven un nombre muy específico.

**Proposición 1.3** *Sea  $C$  una cópula generada por  $\phi$ . Defínase a las potencias de  $u$  recursivamente por*

$$u_C^1 = u \quad y \quad u_C^{n+1} = C(u, u_C^n) \\ \Rightarrow \forall u, v \in I \exists n \in \mathbf{N} \text{ tal que } u_C^n < v$$

*Demostración:*

Nótese que  $u_C^n = \phi^{(-1)}(n\phi(u))$ .

Por la propiedad de Arquímedes para  $\phi(u)$ ,  $\phi(v)$  números reales existe un entero  $n$  tal que  $n\phi(u) > \phi(v)$  entonces se sigue que  $\phi^{(-1)}[n\phi(u)] < v$ .

□

Es precisamente por esta característica que se les denomina cópulas Arquimedeanas. Para resumir otras características interesantes de ellas se presenta el siguiente resultado.

**Proposición 1.4** *Sea  $C$  una cópula arquimedea con generador  $\phi$ . Entonces:*

1.  $C$  es simétrica
2.  $C$  es asociativa
3.  $C_\phi = C_{c\phi}$
4.  $u \geq \delta_C(u) = \phi^{(-1)}(2\phi(u)) \forall u \in [0, 1]$   
y en particular  $\delta_C(u) = u \Rightarrow (u = 1 \text{ ó } u = 0)$

*Demostración:*

1. Es claro que

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

2. También se tiene

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

3. Para probar este punto es necesario observar que para  $\phi_1(s) = c\phi(s)$  la pseudo-inversa esta dada por  $\phi_1^{(-1)}(t) = \phi^{(-1)}(c^{-1}t)$ ; a partir de esto es claro que

$$C_{\phi_1}(u, v) = C_{c\phi}(u, v) = \phi^{(-1)}[c^{-1}\{c\phi(u) + c\phi(v)\}] = C_\phi(u, v)$$

4. Es fácil ver que

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

el resto se deduce de aplicar la definición de pseudo-inversa y notar que  $\phi(u) = 2\phi(u) = 0$   $2\phi(u) \in [0, \phi(0)]$  y  $\phi(u) = \phi(0)$  en cualquier otro caso.

□

**Ejemplo 1.6.1**

a) La cópula producto y la cota inferior son cópulas Arquimedeanas puesto que para  $\phi_1(t) = -\log t$  con  $t \in [0, 1]$  y  $\phi_2(t) = 1 - t$  con  $t \in [0, 1]$  son ambos generadores estrictos. Las inversas son  $\phi_1^{-1}(t) = e^{-t}$  y  $\phi_2^{-1}(t) = \max(1 - t, 0)$  con  $t \in [0, 1]$  para ambas inversas.

Claramente las funciones anteriores son convexas y a partir de (1.17) vemos que:

$$\Pi(u, v) = C_1(u, v) = \exp[-(-\log u - \log v)] = uv$$

y también

$$W(u, v) = C_2(u, v) = \max[1 - (1 - u + 1 - v), 0] = \max[u + v - 1, 0]$$

b) La cota superior  $M$  no es una cópula Arquimedeanas. Es fácil probar esto a partir de la proposición anterior; según la negación lógica de la última parte de (iv)  $u \neq 0$  y  $u \neq 1 \Rightarrow \delta_C(u) \neq u$  para una cópula arquimedeanas. Sin embargo observamos que:

$$M(u, u) = \min(u, u) = u = \delta_C(u) \quad \forall u \in (0, 1)$$

y por tanto puede concluirse que  $M$  no es arquimedeanas.

**1.6.2. C-medida y curvas de nivel**

Como se revisó en la segunda sección de este capítulo, las curvas de nivel o conjuntos de nivel son útiles para visualizar el comportamiento de las cópulas. Para una cópula arquimedeanas se tiene

$$A_t(C) = \{(u, v) \in \mathbf{I}^2 | \varphi(u) + \varphi(v) = \varphi(t)\}$$

o de manera equivalente, los puntos en la curva de nivel  $v = L_t(u) = \varphi^{[-1]}(t) - \varphi^{[-1]}(u)$ .

Tomando los puntos  $u = t$  y  $u = 1$ , puede observarse que esta curva une los puntos  $(1, t)$  y  $(t, 1)$ .

Por otra parte, al conjunto  $A_0(C)$  se le denota como  $Z(C)$  y se nombra *conjunto cero* y para algunas cópulas arquimedeanas este conjunto consiste en los segmentos  $\{0\} \times \mathbf{I}$  y  $\mathbf{I} \times \{0\}$  mientras que para otras, este tiene un área positiva. En estos casos a la cota  $v = L_0(u)$  se le denomina *curva cero*.

Puede demostrarse que las curvas de nivel de una cópula arquimedeanas son convexas y la C-medida de estas está dada por

$$\phi(t) \left[ \frac{1}{\phi'(t^-)} - \frac{1}{\phi'(t^+)} \right]$$

con  $\phi'(t^+)$  y  $\phi'(t^-)$  las derivadas por derecha e izquierda respectivamente.

Presentamos ahora otros resultados relacionados con la C-medida (que como mencionamos es una medida de probabilidad) inducida por una cópula arquimedeanas.

**Teorema 1.6.2** Sea  $C$  una cópula arquimedea generada por  $\phi$ . Sea  $K_C(t)$  la  $C$ -medida del conjunto  $\{(u, v) \in \mathbf{I}^2 | C(u, v) \leq t\}$ . Entonces para toda  $t \in \mathbf{I}$

$$K_C(t) = t - \frac{\phi(t)}{\phi'(t^+)} \quad (1.19)$$

*Demostración:*

Sea  $n \in \mathbf{N}$ . Considérese la partición  $\{0, w/n, \dots, kw/n, \dots, w = \phi(t)\}$  y también

$\{t = t_0, t_1, \dots, t_n = 1\}$  donde  $t_{n-k} = \phi^{(-1)}(kw/n)$ .

Bajo esta partición obsérvese que:

$$\begin{aligned} C(t_j, t_k) &= \phi^{[-1]}(\phi(t_j) + \phi(t_k)) \\ &= \phi^{[-1]}((n-j)w/n + (n-k)w/n) \\ &= \phi^{[-1]}(w(2 - (j+k)/n)) \end{aligned}$$

de donde se sigue que  $C(t_j, t_{n-j}) = t$

La  $C$ -medida del conjunto  $\{(u, v) \in \mathbf{I}^2 | C(u, v) \leq t\}$  corresponde a calcular el  $C$ -volumen del área debajo de  $L_t(u)$ .

Definamos a  $R_k$  como el rectángulo  $[t_{k-1}, t_k] \times [0, t_{n-k+1}]$  se tiene

$$\begin{aligned} V_C(R_k) &= C(t_k, t_{n-k+1}) - C(t_k, 0) - C(t_{k-1}, t_{n-k+1}) + C(t_{k-1}, 0) \\ &= C(w(1 - 1/n) - 0 - t + 0) = \phi^{(-1)}(w - w/n) - \phi^{(-1)}(w) \end{aligned}$$

La  $C$ -medida puede aproximarse como la suma de  $V_C([0, t] \times [0, 1]) = t$  y  $V_C \bigcup_{k=1}^n (R_k)$  y de hecho:

$$\begin{aligned} K_C(t) &= t - \lim_{n \rightarrow \infty} \sum_{k=1}^n V_C(R_k) \\ &= t - \lim_{n \rightarrow \infty} -w \frac{\phi^{(-1)}(w) - \phi^{(-1)}(w - w/n)}{w/n} \\ &= t - \frac{\phi(t)}{\phi'(t^+)} \end{aligned}$$

□

*Observación:* podemos afirmar que la distribución de probabilidad de  $C(U, V)$  para una cópula arquimedea queda en términos del generador puesto que la  $C$ -medida es lo mismo que  $P[C(U, V) \leq t]$  con  $C$  vista como función de distribución con marginales uniformes. Todo esto será de utilidad en el capítulo de ajuste de cópulas.

Para ilustrar el funcionamiento del resultado anterior, revisemos el siguiente:

**Ejemplo 1.6.2** Consideremos la familia Clayton cuyo generador es  $\phi(t) = \frac{1}{\theta}(t^{-\theta} - 1)$ ; derivando con respecto de  $t$  obtenemos:  $\frac{d}{dt}\phi(t) = -\frac{t^{-\theta}}{t}$ ; de lo anterior se sigue que para esta familia

$$K_C(t) = \frac{1}{\theta}t(\theta + 1 - t^\theta)$$

A continuación presentamos un resultado útil cuando se trabaja con cópulas arquimedeanas absolutamente continuas que será útil para probar el último teorema de este capítulo.

**Proposición 1.5** *Sea  $C$  una cópula arquimedeaana estricta absolutamente continua y con generador  $\phi$ . Entonces se cumple que:*

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

*Demostración:*

$$\begin{aligned} \frac{\partial^2}{\partial u \partial v} C(u, v) &= \frac{\partial}{\partial u} [\phi^{[-1]'}(\phi(u) + \phi(v))\phi'(v)] \\ &= \frac{\partial}{\partial u} \frac{\phi'(v)}{\phi'[\phi^{[-1]}(\phi(u) + \phi(v))]} \\ &= \frac{\partial}{\partial u} \left[ \frac{\phi'(v)}{\phi'[C(u, v)]} \right] \\ &= - \frac{\phi'(u)\phi'(v)\phi''(C(u, v))}{[\phi'(C(u, v))]^3} \end{aligned}$$

□

**Teorema 1.6.3** *Sean  $U, V, K_C$  como en (1.17) y defínanse  $S = \frac{\phi(U)}{\phi(U) + \phi(V)}$  y  $T = C(U, V)$ . Entonces la distribución conjunta de  $S$  y  $T$  está dada por*

$$H(s, t) = sK_C(t)$$

*Demostración:*

Supondremos que  $C$  es absolutamente continua<sup>18</sup>. Para hallar la densidad conjunta de  $S$  y  $T$  usaremos el *Teorema de transformación*<sup>19</sup>. Primeramente, necesitamos  $\frac{\partial(u, v)}{\partial(s, t)}$  el Jacobiano de las transformaciones asociadas a  $S$  y  $T$ .

Para  $s$  y  $t$  tenemos que

$$s[\phi(u) + \phi(v)] = \phi(u) \quad \text{y también} \quad \phi(v) = \phi(t) + \phi(u)$$

De donde

$$s\phi(t) = \phi(u) \quad \text{luego} \quad \phi(v) = (1 - s)\phi(t)$$

y por lo tanto

$$J = \frac{\phi(t)\phi'(t)}{\phi'(u)\phi'(v)}$$

Como

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

<sup>18</sup>La demostración del caso general puede encontrarse en Muriel, Nelson. (2004) *Sobre cópulas y dependencia de variables aleatorias* Tesis. UNAM, México

<sup>19</sup>Mood, et.al Op.Cit. p209



entonces, por la proposición anterior

$$\begin{aligned} h(s, t) &= |J| \frac{\partial^2}{\partial u \partial v} C(u, v) \\ &= \frac{\phi(t) \phi''(t)}{[\phi'(t)]^2} \end{aligned}$$

De donde se sigue que

$$\begin{aligned} H(s, t) &= \int_0^s \int_0^t \frac{\phi(y) \phi''(y)}{[\phi'(y)]^2} dy dx \\ &= s \left[ t - \frac{\phi(t)}{\phi'(t)} \right] \\ &= s K_C(t) \end{aligned}$$

□

*Observación:* El teorema anterior indica que las variables  $S$  y  $T$  son independientes, además de que  $S \sim U(0, 1)$

### 1.6.3. Simulación

Presentamos ahora dos algoritmos útiles para la simulación de observaciones de cópulas arquimedeanas. Primeramente, basándonos en el teorema anterior, notemos que para variables  $s$  y  $t$  uniformes si hacemos:  $w = K_C^{[-1]}(t)$   $w$  se distribuye como  $C(U, V)$  mientras que si tomamos  $u = \phi^{[-1]}(s\phi(w))$  y  $v = \phi^{[-1]}((1-s)\phi(w))$  se distribuirán como observaciones de  $U$  y  $V$  respectivamente.

Lo que nos permite proponer el siguiente algoritmo.

#### Algoritmo 3

1. Generar dos variables aleatorias  $s$  y  $t$  uniformes e independientes
2. Hacer  $w = K_C^{[-1]}(t)$
3.  $u = \phi^{(-1)}(s\phi(t))$
4.  $u = \phi^{(-1)}((1-s)\phi(t))$
5. el par  $(u, v)$  es una observación de  $C$

Ahora, retomando el Algoritmo 2 de la sección 1.4, en términos de cópulas arquimedeanas podemos obtener:

**Algoritmo 4**

1. Generar dos variables aleatorias  $u$  y  $t$  uniformes e independientes
2. Hacer  $w = \phi'(-1)(\phi(u)/t)$ ;
3. Consideramos  $v = \phi^{(-1)}(\phi(w) - \phi(u))$ ;
4. el par  $(u, v)$  es una observación de  $C$

En términos generales la elección del algoritmo para realizar la simulación dependerá de la dificultad que implique cada uno en su implementación. Desde luego el primero de ellos exige la existencia de la primera derivada de  $\phi$  que existe casi en todas partes al tratarse de una función continua y convexa.

De acuerdo con el Teorema 1.6.1 podemos generar tantas cópulas arquimedeanas como funciones convexas podamos identificar. Para finalizar este capítulo listaremos algunas de las familias más conocidas. Para un listado más extenso se recomienda Nelsen.<sup>20</sup>

Familia	$\phi_\theta(t)$	$C_\theta(u, v)$	$\theta \in$
Ali-Mikhail-Haq	$\log \frac{1 - \theta(1-t)}{t}$	$\frac{uv}{1 - \theta(1-u)(1-v)}$	$(0, \infty)$
Clayton	$\frac{1}{\theta}(t^{-\theta} - 1)$	$\max[(u^{-\theta} + v^{-\theta} - 1), 0]$	$[-1, \infty),$ $\theta \neq 0$
Gumbel-Hougaard	$(-\log t)^\theta$	$\exp\{-[(-\log u)^{-\theta} + (-\log v)^{-\theta}]^{1/\theta}\}$	$[1, \infty]$
Frank	$-\log \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$	$-\frac{1}{\theta} \log[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)}]$	$[-\infty, \infty),$ $\theta \neq 0$
Gumbel Barnett	$-\log(1 - \theta \log t)$	$uv \exp[-\theta \log u \log v]$	$(0, 1]$

Cuadro 1.1: Algunas familias arquimedeanas

---

<sup>20</sup>Nelsen, Op.Cit pp 94 a 97

## Capítulo 2

# Dependencia

La ciencia actuarial continuamente se enfrenta al reto del modelado de la dependencia estocástica dentro de aplicaciones financieras, estadísticas e incluso económicas. Es un cuestionamiento inneludible que, como se verá, puede ser resuelto completamente a través de las cópulas, al proveernos estas de estructuras, medidas y herramientas para su caracterización.

Como parte del estudio del fenómeno que se desea modelar es conveniente realizar pruebas sobre la independencia, monotonicidad y en general del comportamiento estocástico de él de tal manera que se identifiquen sus características esenciales a nivel intuitivo y también a nivel práctico. Es a partir de aquí que podemos diferenciar dos aspectos de las herramientas empleadas con el propósito de caracterizar la dependencia existente en el fenómeno: por una parte, están las medidas numéricas, que en general reportan un grado de dependencia, y por otro, los modos de dependencia (como la monotonicidad y dependencia de cola) que nos dicen cómo se da esta.

El coeficiente de correlación lineal, las medidas de correlación de rangos (que veremos son 'medidas de concordancia') la prueba de independencia Chi forman parte del primer grupo. Como ventaja de las dos primeras, está el hecho de que son medidas libres de escala, están acotadas y sus casos límites tienen una interpretación simple. Sus limitantes son incuestionables: la validez del primero se encuentra limitado al caso en que exista una relación lineal entre las variables o se trabaje con distribuciones elípticas como es el caso de la Normal; la segunda no caracteriza por completo la independencia (i.e. para la rho de Spearman  $X \perp Y \Rightarrow \rho_{XY} = 0$  pero  $\rho_{XY} = 0 \not\Rightarrow X \perp Y$ ) mientras que la última se verá afectada por la categorización que se haga.

En contraste, los modos de dependencia no tienen ninguno de estos problemas al resumir las características de dependencia en términos no cuantitativos además de proveer información distinta a la concordancia y la relación lineal entre variables. Como desventaja, para probar la existencia de un modo de dependencia generalmente es necesario recurrir a las distribuciones condicionales lo que en términos muestrales no suele ser simple.

Al incorporar las cópulas en nuestro estudio seremos capaces de inducir medidas tales que  $X \perp Y \Leftrightarrow \delta_{XY} = 0$  útiles en todo caso y caracterizaciones de la dependencia a partir de las particularidades propias de la cópula  $C$  asociada a las variables y otras a partir de la relación de orden  $\succ$  presentada en el capítu-

lo anterior. Además de esto, veremos que medidas conocidas como la tau de Kendall y la rho de Spearman pueden ser expresadas únicamente en términos de cópulas.

La estructura de este capítulo es la siguiente: en la primer sección se realiza la revisión correspondiente a los modos de dependencia. Se abordan solamente tres formas de ella y se presentan algunos resultados importantes que permiten incorporar las cópulas en este estudio. En la segunda sección, y dado que es casi inevitable referirnos a ella cuando hablamos de dependencia, hablaremos de la correlación y el coeficiente de Pearson, presentando algunas de las características que limitan su uso y otras que por el contrario rescataremos en las siguientes dos secciones. Posteriormente, en la sección 2.3 revisaremos las denominadas medidas de concordancia como un acercamiento a la medición de la dependencia aunque no como único propósito. Como se verá en el capítulo siguiente, particularmente la  $\tau$  de Kendall es de gran utilidad en la estimación de los parámetros de cópulas arquimedeanas. En la sección 2.4 revisamos las características de las medidas de dependencia y presentamos una de ellas cuya implementación muestral es relativamente simple. Finalmente, en 2.5 presentamos a las Cópulas Empíricas, sus propiedades y su empleo en el cálculo de las medidas muestrales de dependencia y concordancia.

## 2.1. Modos de dependencia

En el capítulo anterior mostramos que las cópulas conforman la estructura de dependencia de una distribución conjunta. Bajo esta aseveración es necesario cuestionarnos qué otras características de la dependencia pueden leerse en ellas. Las cotas de Frechét-Hoeffding y la cópula producto representan la dependencia en términos de funciones monótonas y la independencia respectivamente. Pero ¿qué podemos decir sobre las cópulas que gráficamente se encuentran entre éstas? ¿es posible dar una interpretación en términos de dependencia a estas formas geométricas?

En esta primera sección mostraremos que lo anterior es posible en algunos casos a partir de los modos de dependencia. Al igual que la relación de orden  $\prec$ , los modos de dependencia no caracterizan todas las formas de dependencia existentes, sin embargo proven criterios formales sobre ella. Para nuestra exposición presentaremos primero algunas definiciones tomadas directamente de Joe<sup>1</sup> para después vincularlas con las cópulas como lo hace Erdely.<sup>2</sup>

### 2.1.1. Dependencia de cuadrante

Empecemos por el modo de dependencia denominado 'Dependencia de cuadrante'. Intuitivamente, el concepto básico es comparar la distribución de probabilidad conjunta con aquella que se obtendría bajo el supuesto de independencia. En términos de esta comparación, habrá una diferencia positiva o negativa, que son las dos versiones de la dependencia de cuadrante. Hagamos la revisión formal.

<sup>1</sup>Joe, H.(1997) *Multivariate Models and Dependence Concepts* Chapman & Hall.

<sup>2</sup>Erdely, A. OpCit. pp 54 a 69

Decimos que las variables aleatorias  $X, Y$  son *Dependientes en Cuadrante Positivamente*, y lo denotamos por  $DCP(X, Y)$ , si para todo  $(x, y) \in \mathbb{R}^2$  se cumple

$$P[X \leq x, Y \leq y] \geq P[X \leq x]P[Y \leq y],$$

y decimos que dicha dependencia es estricta si la desigualdad es estricta para al menos un  $(x, y) \in \mathbb{R}^2$ .

Nótese que si consideramos que  $X$  y  $Y$  tienen distribución conjunta  $H$  y marginales  $F$  y  $G$  respectivamente :

$$DCP(X, Y) \Leftrightarrow H(x, y) \geq F(x)G(y) \Leftrightarrow F(x|y) \geq F(x) \Leftrightarrow G(y|x) \geq G(y)$$

En palabras, existe una influencia positiva entre las variables: los sucesos en una afectan a la otra al aumentar su probabilidad de ocurrencia. Si consideramos ahora a  $X$  y  $Y$  como los tiempos de vida de dos componentes,  $DCP$  implicaría que la probabilidad de que ambos sean grandes (pequeños) simultáneamente es superior a la que tendrían individualmente, es decir, la supervivencia de uno empuja a la supervivencia del otro.

Ahora, de acuerdo a la segunda desigualdad y asumiendo que  $X$  y  $Y$  tienen por cópula a  $C$ , siguiendo el Teorema de Sklar y la definición de orden del capítulo anterior:

$$DCP(X, Y) \Leftrightarrow C \succ \Pi$$

De acuerdo a esto,  $DCP(X, Y)$  implica que  $C$  está más cerca a  $M$  que a  $W$ , como ya se había hecho notar en la sección 1.4.2 cuando hablábamos de la relación de orden  $\succ$ . De manera similar definimos  $DCN(X, Y)$  ( $X$  y  $Y$  *Dependientes en Cuadrante Negativamente*) invirtiendo el sentido de las desigualdades anteriores y llegando a

$$DCN(X, Y) \Leftrightarrow H(x, y) \leq F(x)G(y) \Leftrightarrow F(x|y) \leq F(x)$$

y a

$$DCN(X, Y) \Leftrightarrow C \prec \Pi$$

para las cópulas.

En este caso, en términos de probabilidad,  $DCN$  implicaría que siendo  $X$  y  $Y$  tiempos de vida, la probabilidad de supervivencia de los componentes en conjunto será menor que si se consideran por separado, de ahí el término de  $DCN$  pues se influyen negativamente cuando interactúan.

### 2.1.2. Monotonicidad de cola

Revisemos ahora otro modo de dependencia basándonos en el anterior. Si condicionamos sobre  $X$ , podemos decir que  $DCP$  se da si:

$$P[Y \leq y, X \leq x] \geq P[Y \leq y|X \leq \infty]$$

de donde podemos notar que una condición más fuerte que  $DCP$  para  $x_1 \leq x_2$  y para toda  $y$  es que:

$$P[Y \leq y|X \leq x_1] \geq P[Y \leq y|X \leq x_2]$$

Es claro que si ésta última condición se cumple, también se cumple *DCP* pero no necesariamente se cumple en el otro sentido. Revisando la última expresión, vemos que se está pidiendo que la probabilidad condicional de  $Y$  dada  $X$  sea una función no decreciente de  $x$ .

En el contexto de que  $X$  y  $Y$  denoten tiempos de vida de dos componentes significa que la probabilidad de que el tiempo de vida  $Y$  sea corto no se incrementa aunque la de  $X$  si lo haga. Descomponiendo la probabilidad condicional notamos que la desigualdad anterior implicaría también:

$$\frac{P[Y \leq y, X \leq x_1]}{P[X \leq x_1]} \geq \frac{P[Y \leq y, X \leq x_2]}{P[X \leq x_2]}$$

Como las funciones de distribución son no decrecientes, para  $x_1 \leq x_2$  sabemos que se debe cumplir que  $P[X \leq x_1] \leq P[X \leq x_2]$ . De esta forma vemos que la probabilidad conjunta  $P[Y \leq y, X \leq x_1]$  debe incrementarse menos que la probabilidad del numerador al considerar  $x_2$  en lugar de  $x_1$ .

En términos de esto, si tomamos valores de  $x_1$  grandes vemos por una parte que los denominadores serán muy parecidos y la desigualdad se justificará únicamente por la distribución conjunta. Por lo anterior, notamos que la probabilidad conjunta decrece -y de hecho decrece de forma monótona- en la cola de la distribución.

Expresemos ahora formalmente dos de las definiciones vinculadas con la monotonidad de cola:

1. *Y es Decreciente de Cola Izquierda en X, y lo denotamos por DCI(Y|X), si  $P[Y \leq y|X \leq x]$  es función monótona decreciente de  $x$  para todo  $y$ ;*  
Si en lugar de las distribuciones acumuladas tomamos las funciones de supervivencia basándonos en *DCP* podemos llegar a:
2. *Y es Creciente de Cola Derecha en X, y lo denotamos por CCD(Y|X), si  $P[Y > y|X > x]$  es función monótona creciente de  $x$  para todo  $y$ ;*

De manera similar a como se hizo con *DCP* pero ahora empleando *DCN* podemos ver que :

$$P[Y \leq y, X \leq x] \leq P[Y \leq y|X \leq \infty]$$

de donde podemos notar que una condición más fuerte que *DCN* sería que:

$$P[Y \leq y|X \leq x_1] \leq P[Y \leq y|X \leq x_2]$$

con  $x_1 \leq x_2$  para toda  $y$ . En esta ocasión, en términos de los tiempo de vida, la probabilidad de que  $Y$  sea pequeño si se incrementa si el de  $X$  aumenta. Si nuevamente descomponemos la probabilidad condicional y tomamos valores de  $x_1$  grandes, notamos que la probabilidad conjunta crece también de forma monótona.

Lo anterior introduce los términos *Creciente Cola Izquierda (CCI)* y *Decreciente en Cola Derecha (DCD)* únicamente intercambiando las palabras 'decreciente' y 'creciente' en los enunciados anteriores.

Ahora, si consideramos  $F, G, H$  y  $C$  como lo hemos hecho anteriormente y de acuerdo al Teorema de Sklar, podemos notar que:

$$DCI(Y|X) \Leftrightarrow P[Y \leq y|X \leq x] \geq P[Y \leq y|X \leq x+h], h > 0$$

$$\begin{aligned} &\Leftrightarrow \frac{H(x, y)}{F(x)} \leq \frac{H(x + h, y)}{F(x + h)} \\ &\Leftrightarrow \frac{C(u, v)}{u} \leq \frac{C(u + h^*, v)}{(u + h^*)} \\ &\Leftrightarrow C(u, v) \end{aligned}$$

es función monótona decreciente en  $u$  para cualquier  $v$

De hecho, a partir de razonamientos similares podemos establecer los siguientes puntos:

1.  $DCI(Y|X) \Leftrightarrow C(u, v)$  es función monótona decreciente en  $u$  para cualquier  $v$ .
2.  $CCD(Y|X) \Leftrightarrow 1 - u - v + C(u, v)/(1 - u)$  es monótona creciente en  $u$ , o equivalentemente, si  $v - C(u, v)/(1 - u)$  es monótona creciente en  $u$ .

### 2.1.3. Monotonidad estocástica

Reproduciendo la forma en que se motivaron las definiciones para la monotonidad de cola, partiendo de  $CCD$  podemos encontrar una condición más fuerte si en vez de  $P[Y > y|X > x]$  exigimos la monotonidad -creciente o decreciente- en  $P[Y > y|X = x]$ .

De esta manera, diremos que:

1.  $Y$  es Estocásticamente Creciente en  $X$ , y lo denotamos por  $EC(Y|X)$ , si  $P[Y > y|X = x]$  es función monótona creciente de  $x$ , para todo  $y$ ;
2.  $Y$  es Estocásticamente Decreciente en  $X$ , y lo denotamos por  $ED(Y|X)$ , si  $P[Y > y|X = x]$  es función monótona decreciente de  $x$ , para todo  $y$ .

De manera similar que como hicimos anteriormente,  $EC$  implicará que la probabilidad de que  $Y$  sea grande crecerá conforme observemos incrementos puntuales en la vida de  $X$ .

Revisemos ahora como podemos interpretar esto en términos de cópulas. De la sección 1.4.11 sabemos que las distribuciones condicionales sobre  $u$  y  $v$  para la cópula  $C(u, v)$  están dadas por las derivadas parciales de estos argumentos.

Entonces, directamente podemos afirmar que para toda  $v$  y para casi toda  $u \in \mathbf{I}$ :

$$EC(Y|X) \Leftrightarrow \frac{\partial C(u, v)}{\partial u} \text{ es monótona decreciente en } u$$

Para finalizar con esta sección presentamos un cuadro con las relaciones entre los modos de dependencia presentados.<sup>3</sup>

<sup>3</sup>Erdely. Op.Cit p73

$$\begin{array}{ccc}
EC(X | Y) \Rightarrow & CCD(Y | X) & \\
\Downarrow & \Downarrow & \\
DCI(Y | X) \Rightarrow & DCP(X, Y) & \Leftarrow CCD(X | Y) \\
& \Uparrow & \Uparrow \\
& DCI(X | Y) & \Leftarrow EC(X | Y)
\end{array}$$

Cuadro 2.1: Relaciones entre los modos de dependencia

## 2.2. Correlación

El término correlación es continuamente malinterpretado; muy comúnmente nos referimos a éste como si fuera una medida de dependencia siendo que este tan sólo describe un aspecto de ella, a saber, una relación lineal entre variables. En Embrechts <sup>4</sup> se exploran detalladamente las propiedades y desventajas del coeficiente de correlación en eventos donde el supuesto de normalidad multivariada no se cumple.

En esta sección únicamente exponemos algunas de las ventajas y desventajas que supone el uso de este estadístico con el objeto de motivar otro tipo de medidas útiles, no solo para medir la relación lineal entre variables sino la dependencia existente entre ellas, sin necesidad de realizar supuestos tan fuertes como los de normalidad.

Un detalle de notación: al coeficiente de Pearson lo denotaremos como  $\rho^P$  para diferenciarlo del coeficiente de Spearman ( $\rho$ ).

### 2.2.1. Coeficiente de Pearson

*El coeficiente de correlación lineal o coeficiente de correlación de Pearson para las variables aleatorias  $X$  y  $Y$  queda definido como*

$$\rho^P = \text{corr}_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

*siempre que las varianzas  $\sigma_X^2$  y  $\sigma_Y^2$  sean finitas.* El coeficiente y la covarianza serán positivos si tanto  $X - \mu_X$  como  $Y - \mu_Y$  tienen el mismo signo con alta probabilidad y negativos cuando tienen diferente signo. La ventaja del primero sobre la covarianza es que al dividirse entre las desviaciones de las variables se elimina el efecto de ellas.

Ahora, de acuerdo con el lema de Hoeffding, bajo el supuesto de que las esperanzas de los valores absolutos de  $X$ ,  $Y$  y  $XY$  son finitas y siendo  $F$ ,  $G$  distribuciones marginales y  $H$  distribución conjunta de  $X$  y  $Y$  tenemos que:

$$\text{Cov}(X, Y) = \int \int_{\mathbb{I}^2} [H(x, y) - F(x)G(y)] dx dy$$

de donde podemos hacer notar: a) que  $DCP$  implica necesariamente  $\rho^P \geq 0$  y b) la covarianza -y por ende el coeficiente de correlación- dependen directamente del comportamiento marginal de las variables y no solamente de la estructura de dependencia en la distribución conjunta  $H$ .

<sup>4</sup>Embrechts, P., McNeil, A., Straumann, D. (1998) *Correlation and dependence in risk management: properties and pitfalls*. Swiss Federal Institute of Technology Zurich



### 2.2.2. Limitantes de $\rho^P$

A partir de esta definición podemos observar una gran limitante: las varianzas deben ser finitas, de otra manera esta medida queda indefinida. Para los actuarios, quienes trabajan muy frecuentemente con distribuciones de colas pesadas es un problema significativo pues este reportará resultados erróneos.

Otra limitante del coeficiente de correlación es que para dos variables aleatorias  $X$  y  $Y$  se cumple que  $X \perp Y \Rightarrow \rho^P = 0$  pero  $\rho^P = 0 \not\Rightarrow X \perp Y$  a menos de que las variables sigan una distribución normal bivariada. Un ejemplo de esto, es considerar  $Y = g(X)$  con  $g(t) = t^2$  y  $X$  tal que  $E(X) = 0$  y  $\sigma_X^2 = 1$ ; como todos los momentos impares de  $X$  serán cero, se sigue que  $\rho^P = 0$  a pesar de que evidentemente no hay independencia ya que podemos conocer toda la información de  $Y$  a partir de una transformación de  $X$ .

En un trabajo reciente, D.Sandor y J. Mango <sup>5</sup> exponen otra limitante más del coeficiente de Pearson. Con el objeto de probar un procedimiento para la obtención de pérdidas correlacionadas, simularon primero una muestra normal estandar bivariada para variables  $(X, Y)$  con un  $\rho^P$  fijo, para posteriormente generar las variables lognormales  $U = \exp(X)$  y  $V = \exp(Y)$ . Se observó que  $\rho_{U,V}^P$  resultó ser bastante inferior que el coeficiente  $\rho_{X,Y}^P$  inicialmente propuesto debido tanto a la pérdida de exactitud en la simulación -que es bastante pequeña- como a la transformación exponencial.

En este sentido, puede probarse que:

$$\rho^P(T_1(X), T_2(Y)) \neq \rho^P(X, Y) \quad \text{para } T_i \text{ transformaciones no lineales}$$

Un problema más de  $\rho^P$  radica la incertidumbre de lo que significa que  $\rho^P \in (0, 1)$ . Para el caso de correlación perfecta a través de la siguiente expresión, podemos ver que si  $Y = aX + b$  casi seguramente entonces  $\rho^2 = \pm 1$

$$\rho^2 = \frac{\sigma_Y - \min_{a,b} E[(Y - (aX + b))^2]}{\sigma_Y}$$

Podemos considerar que si  $\rho^P \rightarrow \pm 1$  existe una correlación casi perfecta, pero ¿qué pasa si por ejemplo  $\rho^P = \pm 0,5$  ?

Estamos en una posición intermedia que nos dice poco sobre el fenómeno estudiado pues sólo podemos saber que una de las variables explica linealmente aproximadamente la mitad del comportamiento de la otra. Desde luego el análisis de regresión podrá decirnos si debemos considerar como una buena opción este modelo o no, y es en este último caso en el que nos veremos obligados a pensar en otra alternativa tanto para la medición de la dependencia como para su modelación.

### 2.2.3. Ventajas de $\rho^P$

Sin embargo, no debemos pensar que este es un indicador inútil. Una ventaja incuestionable es la manera tan simple en que se puede calcular en términos muestrales, basta para ello la obtención de los primeros y segundos momentos. Además tiene la ventaja de que se trata de una medida estandarizada al

<sup>5</sup>Mango, D., Sandor, J.( ) *Dependence Models and the Portfolio Effect*

cumplirse que:

$$-1 \leq \rho^P \leq 1$$

Otro aspecto positivo radica en el hecho de que es invariante a transformaciones lineales al satisfacer:

$$\rho^P[a_1X + b_1, a_2Y + b_2] = \text{signo}(a_1a_2)\rho_\ell[X, Y]$$

Además, si se trabaja con distribuciones elípticas

$$X \perp Y \Leftrightarrow \rho^P = 0$$

caso claro de ello es la distribución Normal Bivariada Estandar:

$$F(u, v) = \int_{-\infty}^v \int_{-\infty}^u \frac{1}{2\pi\sqrt{1-\rho^2}} \times \exp\{x_2 - 2\rho^P xy + y^2\} dx dy$$

donde  $\rho_\ell = 0$  necesariamente nos remite a la independencia.

Finalmente, rescatamos dos observaciones muy importantes; para ello consideremos  $F$  y  $G$  distribuciones marginales y  $H$  distribución conjunta:

1. si se trabaja con distribuciones elípticas entonces con la información de las distribuciones marginales  $F$  y  $G$  y su correlación  $\rho^P$  podemos recuperar a  $H$  por completo
2. si se trabaja con distribuciones no elípticas entonces con la información de las distribuciones marginales  $F$  y  $G$  y su correlación  $\rho^P$  en general no podemos saber cómo es  $H$ .

Conceptualmente esto es un poco difícil de entender, sin embargo, para aclararlo mostramos la Figura 2.1. En el primer recuadro se muestran observaciones de la

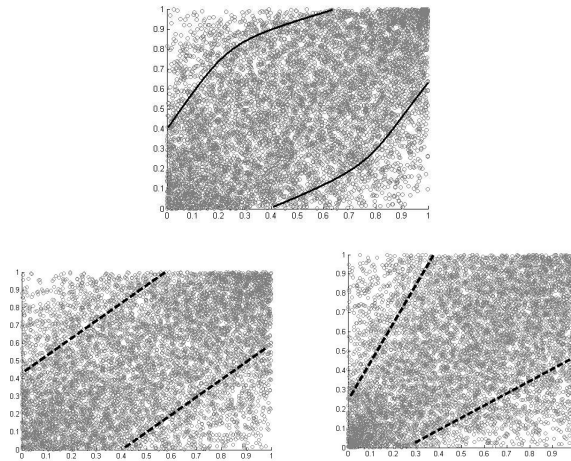


Figura 2.1: Observaciones de cópula Gauss, Frank y Ali-Mikhail con  $\rho^P = 0.42$

cópula gaussiana y en las siguientes a la cópula Frank y Ali-Mikhail-Haq todas ellas con marginales uniformes y coeficiente de correlación  $\rho^P = 0,42$ .

La diferencia entre las estructuras señaladas radica principalmente en las dispersión comprendida entre las esquinas superior izquierda y derecha inferior. Mientras que la cópula gaussiana tiene una forma elíptica y en los bordes superior derecho e inferior izquierdo define pequeños triángulos, la cópula Frank parece extenderse a lo largo de una banda lineal teniendo una menor concentración de puntos hacia los extremos aunque también formando dos triángulos de mayor tamaño que los de la cópula gaussiana. Por su parte la cópula Ali-Mikhail-Haq tiene una mayor concentración de observaciones en el borde inferior izquierdo y se proyecta hacia uno de mayor tamaño en el extremo opuesto.

Bajo estas observaciones es claro que un coeficiente de correlación -aún no tan pequeño como el de la simulación- no garantiza nada respecto a la forma de dependencia explícita de las variables de interés aún suponiendo un comportamiento marginal fijo para ellas. Si bien *DCP* (*DCN*) garantiza un coeficiente de Pearson positivo (negativo), no hay mucho más que podamos saber a menos que se asuma normalidad. Plantemos ahora una pregunta que tiene todo el objeto de cuestionar precisamente el supuesto de normalidad para una muestra bivariada cuando se ha concluido normalidad en cada una de las marginales.

Supongamos que poseemos muestra bivariada y las pruebas de bondad de ajuste sobre cada una arroja resultados que no permiten rechazar la normalidad en cada caso; además se obtiene un coeficiente de correlación  $\rho_{12}$ . ¿A partir de lo anterior es correcto asumir un comportamiento normal bivariado con los parámetros estimados y coeficiente de correlación  $\rho_{12}$ ? Aún más, ¿si  $\rho_{12} = 0$  podemos estar seguros de que hay independencia? Obviamente en ambos casos estaríamos inclinados a pensar que sí, sin embargo, la respuesta debe ser NO en ambos casos.

Para justificarlo, notemos primeramente que no se ha hecho nada por especificar qué tipo de dependencia hay entre las variables. Aunque sepamos que estas son normales, bien puede suceder que en efecto no tengan relación alguna -en cuyo caso sería correcto concluir independencia- o bien, que haya una relación entre ellas -si esta es no lineal, se justifica el hecho de que  $\rho_{12} = 0$ - y la conclusión de independencia es totalmente errónea.

Para aclarar esto, regresemos a la Figura 2.1 donde veíamos que para marginales uniformes existen diferentes representaciones de dependencia que permiten obtener el coeficiente de correlación señalado. ¿Habrá algo que nos impida hacer esto mismo para marginales normales? De hecho la respuesta es NO y además hay algo muy importante : *esto puede hacerse para cualquier par de marginales no necesariamente de la misma familia*. De esta forma, vemos que el comportamiento marginal y el coeficiente de correlación no permiten garantizar una forma explícita del comportamiento conjunto <sup>6</sup>, para lo cuál es necesario realizar pruebas sobre la estructura de dependencia.

Como conclusión podemos decir que el coeficiente de correlación de Pearson es una herramienta útil para la descripción de la dependencia sólo que en su empleo debemos tener cuidado y recordar que los resultados obtenidos a partir de él siguen supuestos de linealidad o normalidad de los datos en conjunto y no solamente sobre el comportamiento marginal.

<sup>6</sup>Aunque esto sucede también para las medidas de asociación que se presentan más adelante, veremos que éstas últimas no son sensibles al comportamiento marginal a diferencia del coeficiente de Pearson.

### 2.2.4. Características deseables en una medida de dependencia

Para finalizar este primer apartado, con base en lo expuesto anteriormente, listamos algunas de las características que nos gustaría conservar en una medida de dependencia :

1. Estandarizada y en sus límites tenga una interpretación clara (como en el caso de  $\rho^P$  que en sus extremos -1 y 1 simboliza una relación lineal perfecta).
2. Invariante a transformaciones lineales y no lineales
3. Capaz de caracterizar por completo la dependencia y a partir de ella podamos concluir si existe o no.
4. Útil sin necesidad de supuestos como los de normalidad o linealidad.
5. Sensible a las estructuras de dependencia existentes y evite la influencia del comportamiento marginal

## 2.3. Medidas de Concordancia

Como alternativa inmediata al coeficiente de correlación, existen las llamadas *medidas de asociación* que engloban tanto a las medidas de dependencia -que son objeto de estudio en la siguiente sección- como las medidas de concordancia. Estas últimas son medidas no paramétricas que calculan el grado de asociación entre dos variables aleatorias tomando en consideración algo bastante simple: ¿para valores 'grandes' ('pequeños') de una de ellas corresponden valores 'grandes' ('pequeños') de la otra? .

Existen una variedad de medidas con esta característica, sin embargo, tanto por su popularidad como por el impacto que tendrán más adelante -en cuanto a medidas de dependencia y ajuste de cópulas se refiere- destinaremos nuestro estudio a la tau de Kendall ( $\tau$ ) y la rho de Spearman ( $\rho$ ) que como se anticipó, estas quedarán expresadas únicamente en términos de la cópula que asocia a las variables aleatorias presentes. Adicionalmente se presenta una definición formal de lo que es una medida de concordancia con la finalidad de mostrar qué puntos han quedado cubiertos de nuestro listado en la sección anterior.

Definamos primero formalmente la concordancia.

**Definición 2.3.1** Para  $(x_i, y_i), (x_j, y_j)$  observaciones de  $(X, Y)$  vector de variables continuas, diremos que este par:

- i) es concordante si  $x_i < x_j$  y  $y_i < y_j$  o bien  $x_i > x_j$  y  $y_i > y_j$ .
- ii) es discordante si  $x_i < x_j$  y  $y_i > y_j$  o bien  $x_i > x_j$  y  $y_i < y_j$ .

Sin ningún problema podemos notar que las siguientes afirmaciones son equivalentes:

- i) es concordante si  $(x_i - x_j)(y_i - y_j) > 0$
- ii) es discordante si  $(x_i - x_j)(y_i - y_j) < 0$ .

A partir estas definiciones podemos presentar una función que se encargue de medir la concordancia y discordancia como sigue:

$$Q := P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] \quad (2.1)$$

Como se observa,  $Q$  es la diferencia entre las probabilidades de concordancia. Dejaremos para más adelante una revisión de sus características generales. Por el momento, notemos que:

$$P[(X_1 - X_2)(Y_1 - Y_2) < 0] = 1 - P[(X_1 - X_2)(Y_1 - Y_2) > 0]$$

y también que:

$$P[(X_1 - X_2)(Y_1 - Y_2) > 0] = P[X_1 > X_2, Y_1 > Y_2] + P[X_1 < X_2, Y_1 < Y_2]$$

por lo que la igualdad en la definición puede replantearse como:

$$Q := 2(P[X_2 < X_1, Y_2 < Y_1] + P[X_1 < X_2, Y_1 < Y_2]) - 1 \quad (2.1b)$$

Mostramos ahora un importante resultado sobre la función  $Q$  y que será de gran ayuda cuando hablemos de la  $\tau$  de Kendall y la  $\rho$  de Spearman.

**Teorema 2.3.1** Sean  $(X_1, Y_1)$  y  $(X_2, Y_2)$  vectores aleatorios independientes de variables aleatorias continuas con funciones de distribución conjunta  $H_1$  y  $H_2$ , con marginales  $F$  para  $X_1$  y  $X_2$  y  $G$  para  $Y_1$  y  $Y_2$ . Además supóngase que  $C_1$  y  $C_2$  son las cópulas de  $H_1$  y  $H_2$  respectivamente y sea  $Q$  como se definió anteriormente.

Entonces:

$$Q = Q(C_1, C_2) = 4 \int \int_{I^2} C_2(u, v) dC_1(u, v) - 1 \quad (2.2)$$

*Demostración:*

Integrando sobre la distribución del vector  $(X_1, Y_1)$  y considerando que  $P[x < X_2, y < Y_2] = 1 - F(x) - G(y) + H_2(x, y)$  de acuerdo con probabilidad condicionada:

$$P[X_1 < X_2, Y_1 < Y_2] = \int \int_{R^2} [1 - F(x) - G(y) + H_2(x, y)] dH_1(x, y)$$

Ahora, sabemos por el Teorema de Sklar que:

$$= \int \int_{R^2} [1 - F(x) - G(y) + C_2(F(x), G(y))] dC_1(F(x), G(y))$$

y si hacemos la sustitución  $u = F(x), v = G(y)$  llegamos a:

$$\begin{aligned} &= \int \int_{I^2} [(1 - u - v) + C_2(u, v)] dC_1(u, v) \\ &= \int \int_{I^2} [(1 - u - v)] dC_1(u, v) + \int \int_{I^2} C_2(u, v) dC_1(u, v) \end{aligned}$$

Como  $C_1$  es la función de distribución del vector  $(U, V)$  de variables uniformes  $(0,1)$ , la primer doble integral corresponde a  $1 - E[U] - E[V] = 0$  y

$$P[X_1 < X_2, Y_1 < Y_2] = \int \int_{I^2} C_2(u, v) dC_1(u, v)$$

Para  $P[X_1 > X_2, Y_1 > Y_2]$  siguiendo un procedimiento similar se llega a:

$$P[X_1 > X_2, Y_1 > Y_2] = \int \int_{I^2} C_2(u, v) dC_1(u, v)$$

entonces, de acuerdo con (2.1b):

$$Q = Q(C_1, C_2) = 4 \int \int_{I^2} C_2(u, v) dC_1(u, v) - 1$$

□

*Nota: si se emplea  $\widehat{C}_1$  y  $\widehat{C}_2$  en lugar de  $C_1$  y  $C_2$  se llega exactamente al mismo resultado.*

De esta forma, notamos que la función definida con anterioridad puede expresarse únicamente en términos de las cópulas de los vectores involucrados.

### Ejemplo 2.3.1

Ilustremos el funcionamiento del teorema y la medida  $Q$  con las cotas de Fréchet-Hoeffding y la cópula producto. Recordemos de (1.4.5) que los soportes de  $W$  y  $M$  son

$$S_W = \{(u, v) \in I^2 \mid u = v\} \text{ y } S_M = \{(u, v) \in I^2 \mid v = 1 - u\}$$

respectivamente. Por tanto, para una función  $g : I^2 \mapsto I^2$  integrable tenemos

$$\int \int_{I^2} g(u, v) dM(u, v) = \int_0^1 g(u, u) du$$

y

$$\int \int_{I^2} g(u, v) dW(u, v) = \int_0^1 g(u, 1 - u) du$$

Por lo anterior,

$$Q(M, M) = 4 \int \int_{I^2} \min(u, v) dM(u, v) - 1 = 4 \int_0^1 u du - 1 = 1 ;$$

$$Q(M, \Pi) = 4 \int \int_{I^2} uv dM(u, v) - 1 = 4 \int_0^1 u^2 du - 1 = 1/3 ; \text{ y}$$

$$Q(M, W) = 4 \int \int_{I^2} \max(u + v - 1, 0) dM(u, v) - 1 = 4 \int_{1/2}^1 (2u - 1) du - 1 = 0 .$$

De manera similar:

$$Q(W, \Pi) = -1/3 \quad \text{y} \quad Q(W, W) = -1$$

Ahora, para la cópula producto sabemos que  $d\Pi = dudv$  con lo que

$$Q(\Pi, \Pi) = 4 \int \int_{I^2} uvdudv - 1 = 0$$

Finalmente, recapitemos algunas de las características de la medida  $Q$  y veamos qué puntos se han cubierto de las características deseables en una medida de dependencia.

1. La medida  $Q$  no se ve afectada por las marginales al estar únicamente en términos de las cópulas de las variables aleatorias.
2.  $|Q| \leq 1$
3.  $Q(C_1, C_2) = 1$  solo si solo hay pares concordantes;
4.  $Q(C_1, C_2) = -1$  solo si solo hay pares discordantes;
5.  $Q(C_1, C_2) = Q(C_2, C_1)$ ;
6. Para  $C_1 \succ C'_1$  y  $C_2 \succ C'_2$  tenemos que  $Q(C_1, C_2) \geq Q(C'_1, C'_2)$
7.  $Q(C_1, C_2) = Q(\widehat{C}_1, \widehat{C}_2)$

A partir de este listado, observamos que la medida  $Q$  cubre los puntos 1, 4 y 5 de forma total y el punto 2 para transformaciones monótonas.<sup>7</sup>

### 2.3.1. Tao de Kendall

Vimos ya un ejemplo de la aplicación del Teorema 2.3.1. Ahora lo emplearemos para reexpresaremos las medidas de Spearman y Kendall en términos de cópulas. Comenzaremos por esta última medida por tener una expresión más natural en términos de  $Q$  que la rho de Spearman. En términos poblacionales la tao de Kendall se calcula como:

$$\tau = \tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]$$

siendo  $(X_1, Y_1)$  y  $(X_2, Y_2)$  vectores aleatorios independientes de variables aleatorias continuas con función de distribución conjunta  $H$  y para la medida muestral como

$$t = \frac{c - d}{c + d} = (c - d)/(n)$$

donde  $n$  es el número de observaciones del vector  $(X, Y)$  y por tanto las combinaciones de  $n$  en 2 el número total de pares de observaciones diferentes,  $c$  el número de pares concordantes y  $d$  el número de discordantes.

A partir de (2.1) y (2.2) es fácil ver que:

$$\tau_{X,Y} = \tau_C = Q(C, C) = 4 \int \int_{I^2} C(u, v) dC(u, v) - 1 \quad (2.3)$$

<sup>7</sup>Recordemos que las cópulas son invariantes a transformaciones monótonas, por ende, la medida  $Q$  -que está en términos de ellas- también lo es.

Ahora, para revisar otro aspecto de esta medida, podemos ver a  $C(U, V)$  como función del vector  $(U, V)$  donde sus componentes son variables aleatorias uniformes  $(0,1)$ , con lo que tenemos:

$$\tau_{X,Y} = \tau_C = 4E(C(U, V)) - 1 \quad (2.4)$$

y para una cópula arquimedea con generador  $\phi$ , considerando (1.17), donde  $K_C(t)$  es la función de distribución de  $C(U, V)$ , realizando integración por partes

$$\tau_C = 4E(C(U, V)) = 4 \int_0^1 t dK_C(t) - 1 = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt \quad (2.5)$$

**Ejemplo 2.3.2** Para la familia Gumbel-Hougaard con generador  $\phi(t) = (-\ln t)^\theta$  observamos que para  $\theta \geq 1$ ,  $\phi'(t) = -\theta(-\ln t)^{\theta-1}t^{-1}$  con lo que  $\frac{\phi(t)}{\phi'(t)} = \theta^{-1}t \ln t$  y luego

$$\tau_\theta = \theta^{-1}(\theta - 1)$$

Lo anterior nos da una fuerte herramienta para la obtención de  $\tau$  cuando trabajamos con cópulas arquimedeanas, notando que en general este dependerá únicamente del parámetro de la cópula.<sup>8</sup> En general nos abocaremos al trabajo con este tipo de cópulas, sin embargo, existe un importante resultado para calcular el tao de Kendall cuando el producto de sus derivadas parciales es integrable, el cual resumimos en la siguiente identidad.

$$\tau_C = 1 - 4 \int \int_{I^2} \frac{\partial}{\partial u} C(u, v) \frac{\partial}{\partial v} C(u, v) dudv \quad (2.6)$$

La prueba puede encontrarse en Nelsen.<sup>9</sup>

**Ejemplo 2.3.3** Para la familia de cópulas Marshall-Olkin  $0 < \alpha, \beta < 1$

$$C_{\alpha,\beta}(u, v) = uv \min[u^{-\alpha}, v^{-\beta}]$$

encontrar  $\tau$  resultará más simple mediante (2.6) que seguir (2.4). Hay que notar también que las derivadas parciales no existirán en  $u^\alpha = v^\beta$ ; en otros términos, en estos puntos no hay probabilidad<sup>10</sup>, por ello tenemos:

$$\frac{\partial C_{\alpha,\beta}(u, v)}{\partial u} \frac{\partial C_{\alpha,\beta}(u, v)}{\partial v} = (1-\alpha)u^{1-2\alpha}v \quad u^\alpha > v^\beta; \quad (1-\beta)uv^{1-2\beta} \quad u^\alpha < v^\beta;$$

y empleando (2.6) llegamos a :

$$\tau_{\alpha,\beta} = \frac{\alpha\beta}{\alpha - \alpha\beta + \beta}$$

<sup>8</sup>Esto será de gran utilidad en el capítulo de ajuste. Al final de esta presentación se muestra un listado de algunas familias y su tao de Kendall

<sup>9</sup>Nelsen, R. Op.Cit. p131

<sup>10</sup>Se recomienda regresar al capítulo primero a la sección 1.4.5 donde se aborda la C-medida. A partir de ello es fácil ver que el soporte de ésta cópula se localiza precisamente en  $u^\alpha \neq v^\beta$



### 2.3.2. Rho de Spearman

Revisemos ahora la  $\rho$  de Spearman cuyo significado es similar a la  $\tau$  de Kendall salvo por que se construye con la información de tres vectores en lugar de dos. Además de esto, precisamente por el empleo de otro vector independiente, la medida quedará en términos de una comparación con la cópula independencia. En términos poblacionales la rho de Spearman se calcula como:

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

es decir, la probabilidad de concordancia menos la probabilidad de discordancia entre los vectores  $(X_1, Y_1)$  y  $(X_2, Y_3)$ , donde  $(X_1, Y_1)$ ,  $(X_2, Y_2)$  y  $(X_3, Y_3)$  son tres vectores aleatorios independientes con función de distribución conjunta  $H$ , marginales  $F$  y  $G$ .

Nótese que la función de distribución conjunta de  $(X_1, Y_1)$  es  $H(x, y)$  mientras que la de  $(X_2, Y_3)$  es  $F(x)G(y)$  por ser  $X_2$  y  $Y_3$  independientes, además, en lugar de éste último puede emplearse el vector  $(X_3, Y_2)$  sin problema. Como resultado de lo anterior y si asumimos que las variables aleatorias  $X$  y  $Y$  tienen cópula  $C$ :

$$\begin{aligned} \rho &= \rho_C = 3Q(C, \Pi) \\ &= 12 \int \int_{I^2} uv dC(u, v) - 3 \\ &= 12 \int \int_{I^2} [C(u, v) - uv] dudv \\ &= 12 \int \int_{I^2} C(u, v) dudv - 3 \end{aligned} \quad (2.7)$$

*Observación:* por la construcción anterior, puede verse a  $\rho$  como un promedio de DCP y DCN estudiadas al principio del capítulo. Por lo anterior, si DCP (DCN) puede concluirse que  $\rho \geq 0$  ( $\rho \leq 0$ )

Retomando los cálculos hechos para la función de discordancia  $Q$  vemos que también  $|\rho| \leq 1$ , puesto que los límites de ésta quedan dados por  $3Q(\Pi, M) = 1$  y  $3Q(\Pi, W) = -1$ , además la rho de Spearman al igual que la  $\tau$  de Kendall hereda el resto de las propiedades de  $Q$ .

Ahora bien, a la rho de Spearman también se conoce como 'coeficiente de correlación de rangos'. Veamos por que:

$$\begin{aligned} \rho &= 12 \int \int_{I^2} uv dC(u, v) - 3 \\ &= 12E[UV] - 3 = \frac{E[UV] - 1/4}{1/12} \\ &= \frac{E[UV] - E[U]E[V]}{\sqrt{\text{var}[U]}\sqrt{\text{var}[V]}} \end{aligned}$$

donde  $U = F(X)$  y  $V = G(Y)$  son variables aleatorias uniformes en el intervalo unitario.

Si consideramos una muestra bidimensional de tamaño  $n$  con rangos  $R_i(x)$  y  $R_i(y)$

$$r = r^P(R_i(x), R_i(y))$$

y toma la forma:

$$r = 1 - 6(n^3 - n)^{-1} \sum_{i=1}^n d_i^2$$

con  $d_i = R_i(x) - R_i(y)$

**Ejemplo 2.3.4** Para la familia de cópulas Marshall-Olkin:

$$\begin{aligned} \rho &= 12 \int \int_{I^2} C(u, v) dudv - 3 \\ &= \frac{3\alpha\beta}{2\alpha - \alpha\beta + 2\beta} \end{aligned}$$

*Observación:* De forma general, tanto la tau de Kendall como la rho de Spearman quedarán en términos de los parámetros de la cópula; es decir, podemos ver a estos como funciones de los parámetros. Como mencionamos con anterioridad, esto será de gran utilidad cuando se deseen estimar parámetros de una cópula a partir de una muestra. Simbólicamente, tendremos:

$$\tau = \tau(\theta) \text{ y } \rho = \rho(\theta)$$

### 2.3.3. Algunas relaciones entre la tau de Kendall y la rho de Spearman

Numéricamente los valores de  $\tau$  y  $\rho^S$  serán en general bastante distintos y esto precisamente por la forma en que se construyen. Mientras que analíticamente la segunda de ellas puede verse como una distancia promedio a la cópula producto y también como el volumen debajo de la gráfica de la cópula por la igualdad (2.7), la primera corresponde a la norma con la cópula producto.

Sin embargo, al estar determinadas por la función  $Q$  existen una variedad de relaciones entre las cuales están <sup>11</sup>:

1.  $-1 \leq 3\tau - 2\rho \leq 1$
2.  $(1 + \rho)/2 \geq [(1 + \tau)/2]^2$
3.  $(1 - \rho)/2 \geq [(1 - \tau)/2]^2$

Ahora, presentamos la definición formal de medida de concordancia, que, como se observa, únicamente es débil en el sentido de no satisfacer el punto (3) de las características deseables para una medida de dependencia. Por lo demás, podemos asegurar que una medida que cumpla con estas características es bastante potente aunque no suficiente para caracterizar la dependencia y es un buen acercamiento a la medida que perseguimos.

<sup>11</sup>Nelsen, Op.Cit pp 141-145

**Definición 2.3.2** Una medida numérica de asociación  $\kappa$  entre dos variables aleatorias  $X, Y$  con cópula  $C$  se denomina medida de concordancia si:

1.  $\kappa$  está definida para todo par  $X, Y$  de variables aleatorias continuas;
2.  $\kappa_{X,-X} = -1 \leq \kappa_{X,Y} \leq 1 = \kappa_{X,X}$ ;
3.  $\kappa_{X,Y} = \kappa_{Y,X}$ ;
4. si la cópula de  $X$  y  $Y$  es  $\Pi$ ,  $\kappa_{X,Y} = \kappa_{\Pi} = 0$  ;
5.  $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$ ;
6. Para  $C_1 \prec C_2$  cópulas,  $\kappa_{C_1} \leq \kappa_{C_2}$ ;
7. Si  $\{(X_n, Y_n)\}$  es una sucesión de variables aleatorias continuas con cópulas  $C_n$ , y si  $\{C_n\}$  converge a  $C$  entonces  $\lim_{n \rightarrow \infty} \kappa_{C_n} = \kappa_C$

Cerramos ahora con un resultado que engloba la mayoría de los aspectos revisados en esta sección.

**Teorema 2.3.2** Tanto  $\tau$  definida en (2.3) como  $\rho^S$  definida en (2.7) son medidas de concordancia.

*Demostración:* El hecho de que estas cumplan las primeras seis condiciones se sigue directamente del empleo de la función  $Q$  para su construcción. La última puede probarse a partir de la condición de Lipchitz expresada en la Proposición 1.1, según la cuál la convergencia de  $\{C_n\}$  a  $C$  es uniforme.

□

## 2.4. Medidas de dependencia

Hasta el momento se han explorado medidas que proveen nociones sobre la forma en que se relacionan dos variables aleatorias  $(X, Y)$ . El coeficiente de correlación lineal y las medidas de concordancia -desde luego en un nivel diferente- facilitan la caracterización de la dependencia aunque por otra parte, no proveen una herramienta que permita probar categóricamente que dos variables aleatorias no se encuentran relacionadas, es decir, que son independientes. Sin más preámbulo y retomando para ello las características deseables en una medida de dependencia presentadas en (2.1.4) presentamos la siguiente definición.<sup>12</sup>

**Definición 2.4.1** Una medida de dependencia para un par de variables aleatorias  $X, Y$  es un número que denotaremos  $\delta_{X,Y}$  que satisface las siguientes propiedades:

1.  $\delta_{X,Y}$  está definida para todo par de variables aleatorias continuas  $X, Y$ ,
2.  $\delta_{X,Y} = \delta_{Y,X}$ ;
3.  $0 \leq \delta_{X,Y} \leq 1$ ;
4.  $\delta_{X,Y} = 0$  si y sólo si  $X, Y$  son independientes;

---

<sup>12</sup>Nelsen, Op.Cit p170

5.  $\delta_{X,Y} = 1$  si y sólo si tanto  $X$  como  $Y$  son funciones monótonas casi seguramente de la otra;
6. Si  $\alpha$  y  $\beta$  son funciones estrictamente monótonas en  $\text{Ran} X$  y  $\text{Ran} Y$ , respectivamente, entonces  $\delta_{\alpha(X),\beta(Y)} = \delta_{X,Y}$
7. Si  $\{(X_n, Y_n)\}$  es una sucesión de variables aleatorias continuas con cópulas  $C_n$ , y si  $\{C_n\}$  converge a  $C$  entonces  $\lim_{n \rightarrow \infty} \delta_{C_n} = \delta_C$

Notamos que cualquier medida  $\delta$  que satisfaga las características de la definición cubre con los requerimientos intuitivos que habíamos hecho con anterioridad. Presentamos a continuación algunas medidas que cumplen con la definición 2.4.1 haciendo énfasis en la relación directa que guardan con las cópulas.

### 2.4.1. Distancias $L_p$

De acuerdo con Schweizer y Wolff <sup>13</sup>, cualquier medida normalizada de las distancias entre las superficies  $z' = C(u, v)$  y  $z = uv = \Pi(u, v)$  puede convertirse en una medida semiparamétrica de dependencia en términos de la definición anterior. Estas se conocen como distancias  $L_p$  entre  $C$  y  $\Pi$  y en general para cualquier  $1 \leq p < \infty$  ésta se calcula como:

$$(k_p \int \int_{I^2} |C(u, v) - uv|^p dudv)^{1/p} \quad (2.8)$$

donde  $k_p$  es una constante de normalización.

### 2.4.2. Algunas medidas de dependencia

Para ejemplificar lo anterior, tomaremos la medida más sencilla inducida a partir de las distancias  $L_p$ . Con  $p = 1$  toma la forma

$$\sigma_{X,Y} = \sigma_C = 12 \int \int_{I^2} |C_{XY}(u, v) - uv| dudv \quad (2.9)$$

Como puede notarse, esta guarda gran similitud con la  $\rho$  de Spearman con la salvedad de que  $\sigma$  es una distancia 'absoluta'. Este cambio, que pareciera ser menor, dota de gran potencia a esta última medida de acuerdo con el punto cuarto de la Definición 2.3.1. Esbozaremos aquí la prueba siguiéndonos el orden de los puntos en la definición e interesándonos particularmente en que  $\sigma_{XY} = 0$  garantice la independencia.

**Teorema 2.4.1** *La expresión definida en (2.9) conocida como  $\sigma$  de Schweizer y Wolff es una medida de dependencia*

*Demostración:*

1. Se sigue por el Teorema de Sklar pues para un par de variables aleatorias continuas  $X$  y  $Y$  existe una única cópula  $C_{XY}$ .
2. Se siguen del hecho de que  $C_{XY}$  es uniformemente continua y empleando el Corolario al Teorema de Sklar, lo que nos permitirá intercambiar los órdenes de integración en (2.8).

<sup>13</sup>Schweizer, B. y Wolff, E. (1981) 'On non parametric measures of dependence for random variables', *Ann. Statist.* **9**, pp 870-885

3. Se sigue de aplicar las cotas de Frechét-Hoeffding al integrando de (2.8)
4. Por el Teorema 1.4.2, si  $X \perp Y \Rightarrow C_{XY} = \Pi$  por lo que  $\sigma_{XY} = 0$ . En el otro sentido, si  $\sigma_{XY} = 0$ , entonces  $|C_{XY}(u, v) - uv| = 0$  casi seguramente. Sin embargo, al ser continuas  $C_{XY}$  y  $\Pi$   $|C_{XY}(u, v) - uv| = 0$  para todo  $(u, v) \in \mathbf{I}^2$  y por el Teorema de Sklar -al ser única la cópula-  $C_{XY} = \Pi$  por lo que empleando nuevamente el Teorema 1.4.2 tenemos  $X \perp Y$ .
5. Se sigue del Corolario 1.4.1
6. Se sigue de los teoremas 1.4.4 y 1.4.5
7. Se sigue del hecho de que las cópulas cumplen la condición de Lipchitz.

□

También basándonos en (2.8) y sustituyendo  $p = 2$  y elevando al cuadrado toda la expresión, tenemos:

$$\Phi_{X,Y}^2 = 90 \int_{\mathbf{I}^2} |C(u, v) - uv|^2 dudv \quad (2.10)$$

Al que se le conoce como 'índice de dependencia' y su raíz cuadrada cumple con ser una medida de dependencia en el sentido de la Definición (2.4.1).

Otra alternativa es el caso límite con  $p = \infty$  tomando la forma

$$\lambda_{X,Y} = \lambda_C = 4 \sup_{u,v \in \mathbf{I}} |C(u, v) - uv| \quad (2.11)$$

que cumple con las condiciones para ser medida de dependencia con excepción de la quinta.

Como se observa la medidas presentadas con anterioridad representan un gran avance para determinar si existe dependencia o no entre dos variables aleatorias. Retomando la lista en 2.2.4, veremos que los requerimientos ahí expresados se satisfacen completamente.

Todas estas características permitirán contrastar las hipótesis  $H_o : C = \Pi$ ,  $H_o : C = W$  o  $M$  al ser equivalentes con las hipótesis  $H_o : \delta_{XY} = 0$  y  $H_o : \delta_{XY} = 1$  respectivamente y todo esto sin necesidad de emplear otro supuesto que no sea el de continuidad para la función de distribución conjunta.

## 2.5. Cópulas empíricas

Si observamos con cuidado los resultados de este capítulo, observaremos que en general consideramos a una cópula  $C$  continua y dada. Sin embargo, ¿qué pasa cuando deseamos analizar la dependencia existente en una muestra?

Con lo revisado hasta ahora, tendríamos que limitarnos a considerar las medidas de concordancia poblacionales y el coeficiente de correlación líneal. Características propias de las medidas de dependencia de la sección anterior y otras relacionadas con los modos de dependencia quedarían completamente aisladas.

En esta última sección resolvemos esta limitante mediante la presentación de

una función no paramétrica denominada 'cópula empírica' y que nos permitirá emplear todo lo desarrollado hasta el momento cuando trabajemos con una muestra de dos variables y más adelante el contraste de diferentes hipótesis para la bondad de ajuste.

Presentaremos por una parte las características esenciales de esta, que incluyen la convergencia a una cópula, y por otra las reexpresiones de medidas de concordancia y dependencia en términos de ella.

Como se mencionó, mostraremos aquí un estadístico para el contraste de la hipótesis  $H_o : C = \Pi$  para una muestra y un algoritmo que responde a la necesidad de hallar valores críticos de él. Iniciemos inmediatamente con su construcción.

Si consideramos una muestra  $\{(x_k, y_k)\}_{k=1}^n$  del vector aleatorio  $(X, Y)$ ; con estadísticos de orden  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$  y  $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$  veremos que las funciones de distribución empírica quedan dadas por:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(x_i \leq x)$$

o bien, en términos de los estadísticos de orden,

$$F_n(x) = \begin{cases} 0 & x < X_{(1)} \\ k/n & X_{(k)} \leq x \leq X_{(k+1)} \\ 1 & X_{(n)} \leq x \end{cases}$$

y algo similar para  $G_n$ . Para la distribución empírica conjunta tenemos:

$$H_n(x, y) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(x_i \leq x, y_i \leq y)$$

Por el Teorema de Sklar sabemos que dada una función de función de distribución existe una cópula  $C_n$ <sup>14</sup> tal que:

$$H_n(x, y) = C_n(F_n(x), G_n(y))$$

y por su Corolario

$$C_n(u, v) = H_n[F_n^{(-1)}(u), G_n^{(-1)}(v)]$$

En términos de los estadísticos de orden, las cuasi-inversas de  $F$  y  $G$  estarán dadas por:

$$F_n^{(-1)}(k/n) = X_{(k)} \text{ y } G_n^{(-1)}(k/n) = Y_{(k)} \text{ respectivamente.}$$

De aquí que:

$$C_n(k_1/n, k_2/n) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(X_{(i)} \leq X_{(k_1)}, Y_{(i)} \leq Y_{(k_2)}) \quad (2.12)$$

con  $k_1, k_2 = 1, 2, \dots, n$

<sup>14</sup>Que en este caso no es única en todo  $\mathbf{I}^2$  dado que la función de distribución no es continua.

También a partir del Teorema de Sklar sabemos que (2.12) es única en el conjunto

$$I_n := \{k_1/n, k_2/n\} \quad \text{con } k_1, k_2 = 1, 2, \dots, n$$

Aquí cabe señalar algo muy importante: la función en (2.1.2) es de hecho una 2 sub-cópula por las anotaciones a la definición 1.2.1; de esta forma, podemos extenderla a una cópula -que como dijimos no es única- con el Lema 1.3.2, lo que nos beneficia en el sentido de extraer en forma funcional la estructura de dependencia de una muestra.

Esta extensión en general no es simple pues como habíamos mencionado antes, se trata de estirar la función de modo que esté definida en todo el cuadrado unitario de forma que se preserven las propiedades iniciales. Más adelante revisamos otras características interesantes de ella, pero primero formalicemos:

**Definición 2.5.1** *Al estimador en (2.12) se le denomina cópula empírica asociada a la muestra  $\{(x_k, y_k)\}_{k=1}^n$ .*

Entre las ventajas de la cópula empírica se encuentra el hecho de que es una estimación no paramétrica de  $C$ , es de distribución libre al no depender de las distribuciones marginales  $F$  y  $G$ .

Por otra parte, las distribuciones empíricas marginales  $F_n, G_n$  sólo dependen de  $F$  y  $G$  y no de  $C$ . Por lo anterior, dadas  $C_n$  y las marginales  $F_n$  y  $G_n$  tenemos que  $H_n$  se determina de modo biunívoco. Las observaciones anteriores nos permiten formular la siguiente:

**Proposición 2.1** *La cópula empírica  $C_n$  definida en (2.12) es un estimador suficiente de  $C$*

□

Además de esta característica, existe un resultado que muestra que este estimador resulta ser uniformemente convergente a una cópula  $C$  el cual enunciamos a continuación.<sup>15</sup>

**Teorema 2.5.1** *Cualquier estimador  $C_n$  definido como en (2.12) sobre  $I_n$  converge uniformemente a  $C$ , es decir:*

$$\lim_{n \rightarrow \infty} \sup_{(u,v) \in P} |C_n(u,v) - C(u,v)| = 0 \quad \text{c.s.} \quad (2.13)$$

También a partir de lo anterior podemos ver que  $C_n$  es un estimador consistente de  $C$  y con ello podemos emplear todos los resultados presentados mediante el empleo de la cópula empírica.

### 2.5.1. Reexpresión de medidas de asociación en términos de $C_n$

Para ilustrar su empleo presentamos las siguientes relaciones muestrales para la tau de Kendall  $t$ , rho de Spearman  $r$  y la sigma de Schweizer y Wolff  $s$ .

---

<sup>15</sup>Muriel, Op.Cit. p105

1.  $t = \frac{2n}{n-1} \sum_{i=2}^n \sum_{j=2}^n \{C_n[\frac{i}{n}, \frac{j}{n}]C_n[\frac{i-1}{n}, \frac{j-1}{n}] - C_n[\frac{i}{n}, \frac{j-1}{n}]C_n[\frac{i-1}{n}, \frac{j}{n}]\}$
2.  $r = \frac{12}{n^2-1} \sum_{i=1}^n \sum_{j=1}^n \{C_n[\frac{i}{n}, \frac{j}{n}] - \frac{i}{n} \cdot \frac{j}{n}\}$
3.  $s = \frac{12}{n^2-1} \sum_{i=1}^n \sum_{j=1}^n \{| C_n[\frac{i}{n}, \frac{j}{n}] - \frac{i}{n} \cdot \frac{j}{n} |\}$

Como podemos ver, el cálculo de las expresiones 1 y 2 puede resultar probablemente más costoso que el empleo de las medidas muestrales revisadas en 2.3.3 y 2.3.4. Sin embargo, para  $s$  no teníamos una expresión en términos muestrales.

### 2.5.2. Elaboración de estadísticos de prueba a partir de la cópula empírica $C_n$

Como mencionamos, a partir de la cópula empírica y gracias a su convergencia uniforme, podemos crear diversos estadísticos de prueba. Empleando el estadístico  $s$  anterior podemos realizar una prueba de independencia o bien emplear estadísticos de la forma:

$$Tn_C = \sup_{u,v \in \mathbf{I}} | C_n(u, v) - C(u, v) | \quad (2.14)$$

que nos permiten contrastar la hipótesis  $H_o : C_n = C$ , de ahí la preferencia por emplear como subíndice de  $Tn$  a  $C$ .

La justificación para este planteamiento se basa en la convergencia de  $C_n$  y en el hecho de que ésta es libre de distribución. Supongamos que tomamos una muestra de  $C$ , es decir, efectivamente  $C_n = C$ . Al obtener nuestro estadístico  $Tn_C$  este se anulará casi seguramente por el Teorema 2.5.1.

Si por el contrario, nuestra muestra no es de  $C$ , también gracias a este teorema, sabremos que  $C_n$  divergerá casi seguramente, con lo que  $Tn_C$  reportará un valor distinto de cero casi seguramente.

Ahora bien, ¿para una muestra dada necesariamente debe cumplirse que  $Tn_C$  sea cero para no rechazar  $H_o$ ? ¿con que nivel de confianza podemos afirmar esto?

Para resolver estas interrogantes podemos emplear el *bootstrapping*<sup>16</sup> ya que no conocemos la distribución a la que converge  $Tn_C$  y su búsqueda es costosa en términos analíticos además de no garantizar un resultado sencillo.

El método de bootstrapping para la obtención de valores críticos consiste en aproximar la distribución del estadístico mediante simulación y tiene su justificación en el Teorema de Glivenko-Cantelli<sup>17</sup> y en nuestro caso particular, también en el Teorema 2.5.1.

De esta manera, a partir de la generación de una gran cantidad de evaluaciones del estadístico podemos obtener una aproximación de la función de distribución real.

Resumiendo, para hallar los valores críticos de  $Tn_C$  :

<sup>16</sup>Esta técnica numérica está aún en desarrollo aunque cuenta ya con varias aplicaciones para estimación de parámetros, prueba de hipótesis y estimación de valores críticos. Para una revisión de ella se recomienda: Davison. A., Hinkley. D. (1997) *Bootstrap Methods and their Application* Cambridge University.

<sup>17</sup>Mood, Op.Cit. p507



**Algoritmo 5**

1. Generar  $N$  muestras de  $C$  y  $C_n$
2. Calcular  $Tn_C(i)$  con  $i = 1, 2, \dots, N$
3. Ordenar en forma asendente  $Tn_C(i)$  y nombrar  $Tn_C[i]$
4. El estadístico ordenado número  $[(1 - \alpha)/N]$  es el valor crítico con nivel de confianza  $\alpha$  del contraste
5. El cuantil del estadístico  $Tn_C$  obtenido en una muestra cualquiera está dado por

$$\frac{1}{n} \sum_{i=1}^N \mathbf{1}(Tn_C \leq Tn_C(i))$$

A partir de lo anterior, diremos que  $H_o$  se rechazará siempre que  $Tn_C \geq Tn_C[(1 - \alpha)/N]$ , o bien, cuando el cuantil de  $Tn_C$  sea menor que  $\alpha$ .

Claramente el estadístico para realizar el contraste sobre cualquier cópula, por ejemplo, para la cópula producto ( $Tn_{\Pi}$ ) o las cotas de Frechet-Hoeffding ( $Tn_W$  y  $Tn_M$ ).

El algoritmo anterior también puede emplearse para hallar los valores críticos de la  $s$  de Schweizer y Wolff y de hecho, en cualquier estadístico sobre el que pueda demostrarse convergencia uniforme. A pesar de ello, muy probablemente encontraremos mucho más fácil de implementar una prueba de hipótesis de independencia a partir de  $Tn_{\Pi}$  por requerir de menos cálculos.

Para finalizar este capítulo, presentamos algunos de los valores críticos del estadístico  $Tn_{\Pi}$  obtenidos a partir de simulación Monte Carlo para diferentes tamaños de muestra.

Tamaño de la muestra	Cuantil		
	90 %	95 %	99 %
100	0.0736	0.0799	0.0921
200	0.0530	0.0575	0.0660
450	0.0358	0.0386	0.0449

Cuadro 2.2: Valores críticos de  $Tn_{\Pi}$  para prueba de independencia bivariada mediante 10,000 ensayos Monte Carlo

## Capítulo 3

# Ajuste de cópulas bivariadas

En los dos capítulos anteriores revisamos algunas de las características esenciales de las cópulas así como su utilidad para la modelación y medición de dependencia. Con excepción de este último punto, todo lo demás ha sido abordado a nivel teórico y poco se ha dicho sobre la aplicación en una muestra.

En este sentido, hagamos una breve descripción del procedimiento a seguir una vez que se nos presenta una muestra bivariada y deseamos indagar sobre el comportamiento del fenómeno descrito en ella. Para esto, asumamos poseemos una muestra  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$  donde desconocemos tanto las distribuciones marginales  $F$  y  $G$  como su distribución conjunta  $H$ .

Por el Teorema de Sklar podemos afirmar que existe una cópula  $C$  tal que  $H(x, y) = C(F(x), G(y))$  única en todo  $\mathbf{I}^2$  si las distribuciones anteriores son continuas; y única sólo en su dominio si éstas no lo son. Por simpleza, supongamos que tanto las distribuciones marginales como la conjunta son continuas y que la cópula  $C$  tiene un solo parámetro.<sup>1</sup>

A partir de lo anterior, en un primer nivel de estudio, aplicaríamos precisamente alguna (o varias) de las pruebas de independencia presentadas en el capítulo anterior. Si a partir de ello observamos que existe independencia con cierto nivel de confianza, podremos modelar este fenómeno sin problemas a partir del producto de las distribuciones marginales. En caso de rechazarse la independencia, procederíamos a la descripción del perfil estocástico a partir la información a priori que poseamos sobre el fenómeno y también empleando las definiciones de la sección 2.1 .

Una vez realizada esta revisión, es posible cuestionarnos sobre la cópula y las distribuciones marginales que describen el fenómeno. En cuanto a las marginales, no abordaremos su estimación como tópico principal pues existen ya una variedad de alternativas para esto mientras que para la estimación de la cópula, revisaremos tres aspectos de interés.

---

<sup>1</sup>Este último supuesto facilita la estimación de parámetros de manera significativa y aunque reduce nuestro espacio de posibilidades, la existencia de una variedad de cópulas uniparamétricas nos garantiza en cierta forma una descripción adecuada de la muestra.

El primero de ellos tiene por objeto reducir nuestra búsqueda y se relaciona con un espacio de alternativas al que puede pertenecer la cópula que describe nuestra muestra, pues como se intuye, no todas permiten modelar todos los perfiles estocásticos. Los dos puntos restantes corresponden a la estimación y las pruebas de bondad de ajuste útiles para seleccionar dentro de este espacio, la cópula que mejor se ajuste.

Para clarificar el primer punto, notemos que la cópula Frank y la cópula Clayton son comprensivas y explican todos los niveles de dependencia y que la cópula Gumbel abarca solo el espacio comprendido entre la cópula producto y la cota superior; por ello, si nuestro fenómeno exhibe *DCN*, definitivamente la cópula Gumbel no es parte de las alternativas para la estimación y sólo realizaríamos la estimación para las dos primeras familias mencionadas. Como se notará este tópico podría llevar a una nueva investigación, sin embargo, con la finalidad de brindar un acercamiento útil a él y que nos permita continuar en nuestra exposición, presentamos una tabla en el apéndice donde se indica que modos de dependencia se pueden modelar a partir de algunas familias de cópulas. El comportamiento preciso de cada una de ellas queda fuera del alcance de esta presentación, realizando con ello la selección de alternativas únicamente bajo el criterio anterior en conjunto con sugerencias en la literatura ya existente sobre este tópico.

Sobre el segundo punto, la estimación, podemos encontrar en la literatura básicamente dos vertientes: la estimación paramétrica y la no paramétrica; ambas hechas bajo el supuesto de que la cópula que describe la muestra es arquimedean. En esta presentación damos holgura al supuesto al bastar que la cópula sea derivable o el parámetro de ella quede descrito como función de una medida de concordancia para lograr una estimación paramétrica. En cuanto a la estimación no paramétrica, retomamos los resultados para cópulas arquimedeanas y describimos brevemente como puede hacerse una estimación de este tipo para familias no arquimedeanas.

Finalmente abordamos el tema de la bondad de ajuste de cópulas como criterio de selección de la cópula que mejor describe a la muestra; y esto no sin antes mostrar algunas alternativas para la revisión visual de los resultados, mismas que se presentan con la finalidad de ayudar a la selección de la cópula y por otra parte, facilitar la comprensión de los resultados teóricos que les preceden. Ya en la revisión analítica, mostramos primero algunas de las desventajas de los estadísticos conocidos Kolmogorov-Smirnov (K-S) y Ji-cuadrada para las pruebas de bondad de ajuste y mostramos dos más -ligados completamente con KS- que tienen la ventaja de ser exclusivos para las cópulas.

En resumen, el capítulo tiene la siguiente estructura: en la primera y segunda sección se muestran los procedimientos de estimación paramétrica y no paramétrica respectivamente. En la sección 3.3 se presentan las alternativas para una revisión gráfica y finalmente en 3.4 la revisión analítica con estadísticos conocidos y con los nuevos propuestos a partir de los llamados procesos empíricos.

### 3.1. Estimación paramétrica

Rescatemos nuestro supuesto fundamental para la estimación y ajuste de cópulas: dada una muestra  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$  asumimos que poseen marginales continuas  $F(x)$  y  $G(y)$  y siguen una distribución conjunta desconocida  $H(x, y)$  con cópula  $C$  uniparamétrica, no necesariamente arquimedea pero derivable o con medida de concordancia en términos de su parámetro. Observando esto podemos tomar caminos distintos para una estimación paramétrica. Intuitivamente, a partir de la muestra podemos:

- Estimar tanto los parámetros de las distribuciones marginales como de la cópula en un solo paso, o bien,
- Estimar las distribuciones marginales y el parámetro de la cópula por separado

El segundo camino es más accesible por 'separar' los problemas a los que nos enfrentamos -estimación de parámetro de distribuciones marginales y parámetro de la cópula-, sin embargo, al realizar esta división es muy posible que se pierda información sobre la cópula o sobre las marginales. La estimación en ambos casos puede hacerse mediante máxima verosimilitud. Revisemos brevemente las alternativas con este método, asumiendo que  $\theta$  es el parámetro de la cópula y  $\alpha_X, \alpha_Y$  son los parámetros (o vectores de parámetros de las marginales):

1. *Máxima verosimilitud sobre los todos los parámetros en un sólo paso.* Esta supondría la maximización de la función:

$$L(\theta, \alpha_X, \alpha_Y; X, Y)$$

que en términos generales plantea un gran problema por depender de tres parámetros: el correspondiente a la cópula y los de las distribuciones marginales  $F$  y  $G$ .

2. *Máxima verosimilitud en dos pasos (conocido como método de funciones de inferencia para marginales o IFM por sus siglas en inglés).* Supone maximizar primero la función

$$L(\alpha_X, \alpha_Y; X, Y)$$

y posteriormente ya con los parámetros de las marginales estimados  $\hat{\alpha}_X$  y  $\hat{\alpha}_Y$ , estimar el parámetro de la cópula a partir de:

$$L(\theta; X, Y, \hat{\alpha}_X, \hat{\alpha}_Y)$$

3. *Máxima verosimilitud sobre el parámetro  $\theta$  en un paso.* La diferencia esencial respecto al anterior es una estimación no paramétrica de las distribuciones marginales mediante sus funciones empíricas  $F_n$  y  $G_n$  respectivamente, y después la maximización de:

$$L(\theta; X, Y, F_n, G_n)$$

Como se observa una estimación como se propone en el primer punto es compleja incluso en términos numéricos, no se hable ya sobre una solución analítica. Por

otra parte, la estimación en el segundo punto conlleva diversos errores precisamente por la partición de la información. Con base en esto, la tercera opción parece ser la más adecuada, ya que, por una parte, se evita la pérdida de la información existente en la segunda alternativa y por otra, simplifica mucho el cálculo propuesto en la primera. Esta alternativa es la primera que revisaremos en esta sección además de la basada en las medidas de concordancia que en general es mucho más simple y se basará en la observación hecha al final de la sección 2.3.4.

### 3.1.1. Máxima Verosimilitud Canónica (MVC)

Al método descrito en el tercer punto se le denomina máxima verosimilitud canónica y es de hecho un método de estimación semiparamétrico y se ha preferido presentarlo en esta sección debido a que la estimación de  $\alpha$  es paramétrica, a través de la función de verosimilitud. Iniciemos con la construcción del estimador para una cópula con parámetro  $\theta \in \Theta$  espacio de posibles valores.

De acuerdo al Teorema de Sklar y considerando que la cópula que describe a  $H$  es derivable en casi todos lados <sup>2</sup>:

$$h(x, y) = f(x)g(y)C_{12}(F(x), G(y))$$

con  $C_{12}(u, v) = \frac{\partial}{\partial u} \frac{\partial}{\partial v} C(u, v)$  que en general es una función del parámetro  $\theta$ .

Definimos:

$$U_i := \frac{n}{n+1} F_n(X_i) \text{ y } V_i := \frac{n}{n+1} G_n(Y_i) \quad (3.1)$$

donde  $(X_i, Y_i)$   $i = 1, \dots, n$  son observaciones en la muestra. Tanto  $U_i$  como  $V_i$  serán observaciones aproximadamente uniformes por el Teorema 1.4.6 de la integral de probabilidad transformada y el Teorema de Glivenko-Cantelli.

Luego, la función de verosimilitud quedará dada por:

$$L(\theta; U, V) = \prod_{i=1}^n C_{12}(U_i, V_i)$$

y el estimador de máximo verosimil de  $\theta$  queda dado por:

$$\hat{\theta}_{MVC} = \operatorname{argmax} \sum_{i=1}^n \log C_{12}(U_i, V_i) \quad \theta \in \Theta \quad (3.2)$$

Para simplificar los cálculos, es factible la siguiente reexpresión de  $U_i$  y  $V_i$  como:

$$U_i := \frac{1}{n+1} \sum_{j=1}^n 1(X_j \leq X_i) \text{ y } V_i := \frac{1}{n+1} \sum_{j=1}^n 1(Y_j \leq Y_i) \quad (3.3)$$

que en general representa una mejor alternativa al no requerir expresamente la construcción de las funciones empíricas de distribución.

Como se habrá observado, se ha hecho una ligera modificación sobre las funciones

<sup>2</sup>Esto es posible precisamente por el supuesto de continuidad en conjunto con las características de las cópulas expuestas en el primer capítulo.

empíricas al multiplicarlas por un factor de  $\frac{n}{n+1}$ . La ventaja de hacer esto es más clara si se observa que conforme  $U_i$  o  $V_i$  se acerquen a 1, (3.2) puede anularse y con ello llevarnos a sesgos en la estimación.

El estimador resultante cumple con ser consistente y asintóticamente normal al igual que los obtenidos a partir de máxima verosimilitud. A esta alternativa de estimación también se le conoce como 'pseudo verosimilitud' y la prueba de las características mencionadas puede encontrarse en Genest & Ghoudi<sup>3</sup>

Recalcamos que este primer método es útil siempre que  $C_{12}$  exista, lo que no nos restringe al uso de cópulas arquimedeanas. Para resumir tenemos el siguiente:

#### Algoritmo 6

1. Obtención de las funciones  $F_n$  y  $G_n$
2. Determinación de  $C_{12}$
3. Obtención de  $U_i$  y  $V_i$  a partir de (3.1) o (3.3)
4. Evaluar  $C_{12}(U_i, V_i)$
5. Maximizar  $\sum_{i=1}^n C_{12}(U_i, V_i)$

#### 3.1.2. Estimación a partir de medidas de concordancia

Como se hizo notar anteriormente, en general las medidas de concordancia quedarán en términos de los parámetros de las cópulas. A partir de las expresiones (2.3) y (2.7) que retomamos adelante por comodidad, esto es claro pues al realizar las integrales únicamente quedan los parámetros de las cópulas.

$$\tau = Q(C, C) = 4 \int \int_{I^2} C(u, v) dC(u, v) - 1 \quad (2.3)$$

$$\rho = 12 \int \int_{I^2} [C(u, v) - uv] dudv \quad (2.7)$$

En estos términos podremos expresar las medidas de concordancia anteriores como funciones del parámetro de la cópula :

$$\tau = z_1(\theta) \text{ y } \rho = z_2(\theta)$$

y por consiguiente si son invertibles de forma analítica -o incluso numéricamente invertibles- podremos expresar a  $\theta$  en términos de las medidas de concordancia como:

$$\theta = z_1^{-1}(\tau) \text{ y } \theta = z_2^{-1}(\rho)$$

para funciones  $z_1^{-1}$  y  $z_2^{-1}$  inversas.

Ahora, si hemos estimado las medidas de concordancia a partir de  $\hat{\tau}$  y  $\hat{\rho}$ , que son estadísticos consistentes, podemos llegar a estimadores de  $\theta$  mediante:

$$\hat{\theta}_\tau = z_1^{-1}(\hat{\tau}) \text{ y } \hat{\theta}_\rho = z_2^{-1}(\hat{\rho}) \quad (3.4)$$

<sup>3</sup>Genest, C., Goudi, K., Rivest, L-P. (1995). 'A semiparametric estimation procedure for parameters in multivariate families of distribution.' *Biometrika*, Vol.82, Issue 3, pp543-552.

Para ejemplificar esto, considerese a una cópula de la familia Farlie-Gumbel-Morgesten. Por el apéndice (NUM) sabemos que:

$$\rho = \theta/3 \text{ y } \tau = 2\theta/9$$

a partir de esto podemos ver que:

$$\theta = 3\rho \text{ y } \theta = 9\tau/2$$

y si calculamos las medidas muestrales  $\hat{\rho}$  y  $\hat{\tau}$  a partir de (3.4) tenemos:

$$\hat{\theta}_\rho = 3\rho \text{ y } \hat{\theta}_\tau = 9\hat{\tau}/2$$

obtendremos estimadores del parámetro.

En general, podríamos emplear cualquiera de las dos medidas de concordancia para lograr una estimación, sin embargo, en adelante preferiremos la tau de Kendall por ser más accesible en su cálculo muestral y porque en general la función  $z_1$  como la describimos arriba es más sencilla y fácil de invertir y también guarda una relación estrecha con la estimación no paramétrica que presentaremos en el siguiente apartado.

Para finalizar, mostramos la consistencia del estimador para diferentes familias. La prueba de ello se ha realizado en forma numérica ya que una exploración analítica se apetece difícil dadas las diferentes características de las funciones inversas.

Cópula		$n=100$		$n=200$		$n=1000$	
Familia	$\theta$	Media	Desv. Est.	Media	Desv. Est.	Media	Desv. Est.
AM	-0.5	-0.5441	0.3404	-0.4996	0.2642	-0.5005	0.1127
	0.5	0.5154	0.1762	0.4853	0.1473	0.4956	0.0707
Cl	-0.5	-0.4978	0.0692	-0.498	0.0478	-0.4999	0.0213
	0.5	0.5193	0.2151	0.5101	0.1498	0.5013	0.0655
GH	1.5	1.5154	0.1446	1.5052	0.1011	1.5017	0.0445
	6	6.0975	0.7938	6.0444	0.5465	6.0079	0.2346
Fk	-5	-5.0539	0.8237	-5.0324	0.5721	-5.0074	0.2513
	5	5.0588	0.8147	5.0299	0.5695	5.0019	0.2478
FM	0.8	0.7976	0.2847	0.7993	0.1972	0.7994	0.0878
		-0.8011	0.2849	-0.7977	0.2002	-0.7998	0.089
CA	0.1	0.0912	0.1286	0.0958	0.0895	0.0995	0.0395
	0.9	0.8988	0.0324	0.8994	0.0225	0.8999	0.01

Cuadro 3.1: Consistencia del estimador de Genest-Rivest para diferentes tamaños de muestra  $n$  con  $N= 10,000$  ensayos MonteCarlo.

Explicemos un poco el cuadro 3.1 basándonos en un ejemplo. Para la familia Clayton con parámetro  $\theta = -0,5$  y tamaño de muestra 100, se han simulado 10,000 muestras y se ha calculado el parámetro  $\hat{\theta}$  a partir de (3.4) para cada una de ellas. Sobre estos 10,000 estimadores se ha calculado su media y desviación estandar, valores que se tabulan en las columnas correspondientes.

Con base en el experimento anterior podemos afirmar que el estimador es insesgado al coincidir su esperanza con el valor real del parámetro. Además, conforme

se incrementa el tamaño de la muestra observamos que la desviación estandar disminuye. En conjunto, podemos afirmar que se trata de un estimador consistente. En el apéndice se presenta un listado que muestra a  $\tau_0$  en términos del parámetro de la cópula para algunas familias.

## 3.2. Estimación no paramétrica

Para realizar una estimación no paramétrica de  $C$  la primer alternativa es desde luego el empleo de la cópula empírica definida en (2.12). Al ser esta convergente de manera uniforme, prové un estimador natural. Aquí no dedicaremos más espacio a esta alternativa por tratarse de algo bastante simple y que tomará mayor importancia en la última sección, donde anticipamos, que este estimador nos permitirá hacer una de las pruebas de bondad de ajuste.

Además de esta posibilidad, pueden emplearse también los Polinomios de Bernstein y 'Checkerboard Approximation' para lograr una estimación consistente de la estructura de dependencia existente en nuestra muestra y que por requerir la introducción de otros conceptos teóricos no se abordará, se recomienda Durrleman <sup>4</sup> para una revisión introductoria de éstas.

Como mencionamos al principio del capítulo, existe un método para la estimación no paramétrica de cópulas arquimedeanas. Este fue propuesto por Genest y Rivest <sup>5</sup> y lo desarrollamos a continuación.

### 3.2.1. Estimación no paramétrica para cópulas arquimedeanas

Para orientar nuestro procedimiento, retomemos el resultado del Teorema 1.6.7. De acuerdo con él, para una cópula arquimedea  $C$ , la  $C$ -medida del conjunto  $\{C(u, v) \leq t\}$  es la función univariada  $K_C(t)$  que queda en términos del generador de  $C$ , explícitamente:

$$K_C(t) = t - \frac{\phi(t)}{\phi'(t)} \quad (1.17)$$

Ahora, directamente de (1.17) vista como una ecuación diferencial y resolviendo para  $\phi$  tenemos:

$$\phi(t) = \exp\left\{\int_{t_0}^t \frac{1}{\mu(t)} dt\right\} \quad (3.5)$$

con  $\lambda(t) = t - K_C(t)$ , lo que exhibe una relación muy explícita entre  $K_C(t)$  y el generador  $\phi$ . A partir de esto, es claro que si conseguimos construir un estimador no paramétrico de  $K_C$ , también podremos obtener un estimador para  $\phi$ .

*Observación:*  $\phi$  definida arriba es de hecho un generador de una cópula arquimedea al satisfacer las condiciones del Teorema 1.6.2 cuando  $\lambda(t) < 0$ .

<sup>4</sup>Durrleman, V., Nikegbali, A., Roncalli, T. (2000) *Copulas approximation and new families*. Groupe de Recherche Opérationnelle, Credit Lyonnais.

<sup>5</sup>Genest, C., Rivest, L-P. (1993) 'Statistical Inference Procedures for Bivariate Archimidean Copulas. *Journal of the American Statistical Association*, Vol.88, pp1034-1043



Además de esto, si vemos que en realidad podemos encontrar a  $K_C(t)$  en términos de la  $C$ -medida del conjunto  $\{C(u, v) \leq t\}$  para cualquier cópula  $C$ , mediante (3.5) podemos definir algo similar a una proyección de casi cualquier cópula a una cópula arquimedea.

Para deducir de donde obtendremos el estimador de  $K_C$ , recordemos ahora que la  $C$ -medida del conjunto es equivalente a  $P[C(U, V) \leq t]$  con  $C$  vista como función de distribución con marginales uniformes en el intervalo unitario. Si utilizamos el Teorema de Sklar para  $X$  y  $Y$  variables aleatorias continuas con funciones de distribución  $F$  y  $G$  marginales y  $H$  conjunta y el Teorema de la Integral de Probabilidad Transformada:

$$K_C(t) = P[C(U, V) \leq t] = P[C(F(X), G(Y)) \leq t] = P[H(X, Y) \leq t]$$

*Observación:*  $X$  y  $Y$  son variables aleatorias y no debe confundirse con  $x$  y  $y$  valores

Si poseemos una muestra bivariada  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$  es factible la obtención de  $F_n$ ,  $G_n$  y  $H_n$  funciones empíricas de  $F$ ,  $G$  y  $H$  respectivamente. Intuitivamente, podríamos seguir un procedimiento similar al empleado en la obtención de la cópula empírica para lograr una aproximación de  $H(X, Y)$  en términos de  $H_n$ . Teniendo claro esto, definamos

$$W_i = \frac{1}{n-1} \sum_{j=1}^n \mathbf{1}[X_j < X_i, Y_j < Y_i] \quad (3.6)$$

que como observamos es aproximadamente  $H_n$  y en términos generales, cuando  $n$  sea grande serán iguales. Se ha preferido definir de esta para ser congruentes con Genest y Rivest <sup>6</sup> en vez de emplear directamente  $H_n$ .

A partir de lo anterior, también podemos definir:

$$K_n(w) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}[W_i < w] \quad (3.7)$$

que es un estimador no paramétrico de  $K_C$  sobre el cual mostraremos sus características y su relación con medida de concordancia de Kendall a continuación.

### 3.2.2. Consistencia del estimador y Procesos de Kendall

Para evidencia la relación de esta estimación con la  $\tau$  de Kendall muestral, defínase:

$$I_{ij} = \begin{cases} 1 & \text{si } X_j < X_i, Y_j < Y_i \\ 0 & \text{otro caso} \end{cases}$$

entonces tenemos que:

$$W_i = \sum_j \frac{I_{ij}}{n-1}$$

---

<sup>6</sup>Genest y Rivest Op.Cit. pp 1037

y también que el número de pares concordantes de una muestra de tamaño  $n$  de variables aleatorias continuas es:

$$c = \sum_i \sum_j I_{ij} = (n-1) \sum_i W_i$$

y el número de pares discordantes

$$d = \frac{n(n-1)}{2} - (n-1) \sum_i W_i$$

luego, podemos expresar la medida de concordancia de Kendall para una muestra como sigue:

$$\begin{aligned} t &= \frac{2(c-d)}{n(n-1)} \\ &= 4 \frac{\sum_i W_i}{n-1} \\ &= 4\bar{W} - 1 \end{aligned}$$

*Observación:* Si se define  $V_i = \sum_j \frac{I_{ji}}{n-1}$ , llegamos a  $\tau = 4\bar{V} - 1$  y la distribución empírica de  $V_i$  serviría para modelar la función de supervivencia de  $(X, Y)$ .

Dada la relación de las  $W_i$ 's definidas con la  $\tau$  de Kendall, a

$$\alpha_n(t) = \sqrt{n}\{K_n(t) - K(t)\}$$

se le denomina 'proceso de Kendall'. En el artículo *On Kendall's Process*<sup>7</sup> se muestran diferentes propiedades de él, destacando su convergencia (cuando  $n$  tiende a infinito) a un proceso de ruido blanco bajo ciertas condiciones de regularidad.

Lo anterior entre otras cosas, permite la construcción de estadísticos para la realización de pruebas de bondad de ajuste -uno de los cuales mostramos en la última sección de este capítulo- y probar la consistencia del estimador en (3.7); resultado que enunciamos a continuación.

**Proposición 3.1** *Para  $H(x, y)$  absolutamente continua en sentido de Lebesgue y una muestra de  $H \{(x_k, y_k)\}_{k=1}^n$ , bajo condiciones de regularidad débiles, la distribución de las  $W_i$  definidas en (3.6) converge a la de  $K_C(t) = P[H(X, Y) \leq t]$  y la distribución empírica  $K_n(t)$  de ellas definida por (3.7) es un estimador  $\sqrt{n}$ -consistente*

Presentamos a continuación un algoritmo que permite la obtención del estimador  $K_n(t)$  a partir de una muestra y -en caso de ser arquimedean- también la del generador  $\hat{\phi}$ .

<sup>7</sup>Barbe, P. Genest, C. (1996) 'On Kendall's Process.' *Journal of Multivariate Analysis*, **58**, pp197- 229

**Algoritmo 7. Obtención de  $K_n$  y  $\hat{\phi}$** 

Para obtener un estimador de  $\phi$  dada una muestra :

1. Obténgase  $W_i$  a partir de (3.6)
2. Con estas obténgase  $K_n$  mediante (3.7)
3. Obténgase la distribución empírica  $\lambda_n$  y a partir de ella la distribución empírica continua .
4. Obténgase la aproximación de  $\phi$  a partir de (3.5)

Antes de iniciar con la revisión de la bondad de ajuste, es pertinente plantear la siguiente pregunta. Dada una muestra ¿cuál de los métodos es mejor para estimar su parámetro?.

De acuerdo con lo que revisamos anteriormente, MVC se empleará cuando la cópula sea derivable, la estimación por medidas de concordancia cuando la expresión que vincule a tau de Kendall con el parámetro sea simple, y el método no paramétrico explicado al final sólo cuando trabajemos con familias arquimedeanas. Pero ¿qué método elegir en caso de que todos sean factibles?.

Para esto no hay una respuesta sencilla. Por una parte, MVC provee un estadístico consistente y asintóticamente normal. Por otra, la obtención del estadístico a partir de la tau de Kendall también es bastante acertada y mucho más simple.

Roberto De Matteis <sup>8</sup> por ejemplo, realiza la estimación a partir de ambos métodos cuando es posible para varias familias uni- paramétricas arquimedeanas y en las pruebas de bondad de ajuste da prioridad a aquellas cuyos parametros estimados a partir de los diferentes métodos son más parecidos. En lo que resta de la presentación seguiremos esta misma ruta cuando llegemos a la parte aplicativa.

### 3.3. Revisión gráfica

Al igual que en la elección del método de estimación, surgen también algunas dificultades en la revisión de la bondad de ajuste. De Matteis sugiere un conjunto de pruebas, partiendo de una revisión gráfica y pasando por los contrastes K-S y Ji cuadrada y el empleo del criterio de Akaike, optando finalmente por aquella familia con mejores resultados globales. Valentyn Panchenko <sup>9</sup> tras el empleo de MVC propone realizar pruebas sobre la bondad de ajuste mediante formas bilineales. Fermanian y Scaillet <sup>10</sup> construyen y emplean estadístico asintóticamente distribuido como una Ji cuadrada a partir del kernell de la cópula empírica.

Aquí mostraremos primero una aproximación gráfica al problema de la bondad de ajuste por considerar que provee información importante previa la aplicación

<sup>8</sup>De Matteis, R. (2001). *Fitting copulas to data*. Diploma thesis. Institute of Mathematics of the University of Zurich.

<sup>9</sup>Panchenko, V. (2005) *Goodness-of-fit test for copulas*. University of Amsterdam

<sup>10</sup>Fermanian, J-D. Scaillet, O. (2004) *Some Statistical Pitfalls in Copula Modeling for Financial Applications* FAME - International Center for Financial Asset Management and Engineering, Research Paper 108. Universidad de Génova. Italia

de algún método analítico, que dicho sea de paso, dejamos para la siguiente sección. Con excepción de la última alternativa, las opciones aquí provistas para la visualización de los ajustes que son básicamente las que propone De Matteis<sup>11</sup> y nos limitaremos a describir la herramienta ya que en el último capítulo mostramos ejemplos reales de estos gráficos.

### 3.3.1. Distribución condicional

Retomando lo dicho en la sección 1.4.11, la distribución condicional de  $Y$  dada  $X$

$$H_{Y|X}(x, y) = C_1(F(x), G(y))$$

con  $C_1(u, v) = \frac{\partial}{\partial u} C(u, v)$  y recordando que esta última será también función del parámetro  $\theta$ . De acuerdo a Klugman y Parsa<sup>12</sup>.  $H_{Y|X}$  sigue una distribución uniforme. A partir de esto, si en lugar de  $F$  y  $G$  se valúan las distribuciones empíricas,  $C_1(F_n(x), G_n(y))$  tendrá una distribución casi uniforme. Entonces si graficamos los cuantiles de una distribución uniforme y  $C_1(F_n(x), G_n(y))$  -evaluando esta también con  $\hat{\theta}$  obtenido a partir de (3.4) o (3.2)- el gráfico tenderá a ser una línea con pendiente uno y que corta el origen si el ajuste es bueno.

### 3.3.2. Distribución de $K_C$

Siguiendo un razonamiento similar al anterior,

$$K_C(C(F_n(x), G_n(y))) = K_C(C(U, V))$$

sigue una distribución casi uniforme. Basándonos en esto, si se realiza el gráfico de  $K_\theta(C(F_n(x), G_n(y)))$  para los diversos valores de la muestra contra cuantiles uniformes, un ajuste exacto mostrará también una línea recta con pendiente 1.

### 3.3.3. Comportamiento de $K_n$

Aquí es preciso graficar  $K_n$  y  $K_C$  valuado en  $\hat{\theta}$ . Un ajuste exacto mostrará dos curvas completamente encimadas. Debido a la construcción de  $\lambda$ , de realizarse un análisis simultáneo se obtendrán gráficos distintos y de hecho el de  $\hat{\lambda}$  será más sensible a desviaciones.

### 3.3.4. Diagonal empírica

Una vez que se ha obtenido la cópula empírica, podemos obtener también la diagonal empírica de la cópula basándonos en la definición 1.2.2. A esta la obtendremos como:  $\delta_n(t) = C_n(t, t)$  y podemos compararla con  $\delta_C(t)$  mediante un gráfico como el siguiente. Es claro que este gráfico no es determinante para la selección de una cópula, pero si puede serlo para descartarla y esto dicho bajo el supuesto de que las diagonales de dos cópulas diferentes pueden parecerse, pero si las diagonales no son parecidas, es muy probable que no se trate de la misma cópula.

<sup>11</sup>De Matteis, Op. Cit. pp 39-41

<sup>12</sup>Klugman, S., Parsa, R. (1997) *Fitting bivariate loss distributions with copulas*. Papers from the Conference on Insurance: Mathematics and Economics held at the University of Amsterdam. pp139-148

### 3.4. Revisión analítica

En esta sección se exploran dos tipos de pruebas de bondad de ajuste. Las primeras corresponden a dos pruebas conocidas (K-S y Ji cuadrada) que como se verá son de uso cuestionable cuando se trabaja con cópulas. La idea central detrás de su empleo está plasmada en las sugerencias gráficas que se presentaron anteriormente y cuyo logro fundamental es transformar el problema de realizar una prueba para datos divariados a una univariada.

Complementario a esto, se realiza la presentación de estadísticos generados exclusivamente para el ajuste de cópulas a partir de los llamados 'procesos empíricos' (donde tienen su origen estadísticos como K-S y Cramér Von Misses) sobre los cuales será necesario hablar un poco para justificar su empleo. Además de esto, se muestran algoritmos para la obtención de los valores críticos de estos y desde luego una tabla con los resultados para algunas familias de cópulas.

#### 3.4.1. Pruebas a partir de Kolmogorov-Smirnov

La prueba Kolmogorov-Smirnov (K-S) tiene por objeto determinar si una muestra proviene de una población con una distribución continua específica  $F$  a partir de la función de distribución empírica  $F_n$ . Formalmente:

$$H_0 : F(x) = F_n(x) \forall x \quad H_1 : F(x) \neq F_n(x) \text{ para algún } x$$

El estadístico de contraste se basa en la distancia entre  $F$  y la función empírica:

$$D = \max_i |F_n(x_i) - F(x_i)|$$

La hipótesis nula se rechazará si  $D$  es mayor que el valor crítico en tablas para este contraste, mismas que pueden encontrarse en casi cualquier libro de inferencia estadística. Este estadístico tiene la conveniencia de ser consistente, libre de distribución <sup>13</sup> y reportar siempre el mismo valor para una muestra a diferencia del de la prueba Ji cuadrada que adelante se describe. Como desventajas esta el hecho de que únicamente es útil para ajustar distribuciones continuas y completamente especificadas, es decir, si se estiman parámetros de esta última a partir de la muestra, los valores críticos en tablas no serán válidos para nuestro ejercicio.

#### 3.4.2. Pruebas Ji-cuadrada

Esta se basa en el estadístico:

$$X^2 = \sum_{i=1}^k \frac{[f_i - np(x_i)]^2}{np(x_i)}$$

donde  $k$  es el número de clases,  $f_i$  la frecuencia absoluta de la muestra en la clase  $i$  y  $np(x_i)$  es la frecuencia teórica de la clase  $i$ . Si el estadístico  $X^2$  es mayor a los valores críticos de una distribución  $\chi_{k-c}^2$  donde  $c$  es el número de parámetros estimados en el modelo, se rechaza la hipótesis nula planteada. Esta prueba es sensible a la estructura y número de clases hechas pero tiene la ventaja

<sup>13</sup>Es decir, que los valores críticos son siempre los mismos independientemente de la  $F$  que se desee ajustar.

de que puede aplicarse a cualquier distribución no necesariamente continua y con parámetros no especificados.

### 3.4.3. Empleo en el ajuste de cópulas

A partir de lo anterior, podemos realizar las pruebas tanto K-S como Ji cuadrada para las representaciones gráficas Q-Q, sin embargo, debido a la limitante en cuanto a la arbitraria partición de clases, se dará preferencia a K-S en este aspecto. Para el ajuste de  $K_n$  se empleara la prueba Ji cuadrada por requerir la previa estimación del parámetro  $\theta$ . Hasta ahora no ha sido posible mostrar que efectivamente el estadístico  $X^2$  converge a una distribución  $\chi^2$  cuando el parámetro es estimado.<sup>14</sup> De igual manera, la absoluta validez de los valores críticos de K-S queda en entredicho por requerir la estimación implícita de un parámetro para evaluar tanto  $C_1(F_n(x), G_n(y))$  como  $K_\theta(C(F_n(x), G_n(y)))$ .

### 3.4.4. Pruebas basados en procesos empíricos

Las observaciones anteriores nos obligan a buscar otras alternativas analíticas para probar la bondad de ajuste de las cópulas. Como se comentó, este es aún un campo en desarrollo y sobre el que se pueden encontrar propuestas desde las más intuitivas, pasando por pruebas como las que se plantean en la sección anterior, llegando hasta conceptos teóricos fuera del alcance de esta presentación como es el caso de las propuestas de Panchenko y Fermanian y Scarlet.

Teniendo esto en cuenta, como alternativa para la obtención de estadísticos confiables para el ajuste de cópulas se eligieron aquellos procedentes de los denominados 'procesos empíricos' debido a la cercanía con los conceptos ya revisados.

La prueba K-S tiene su origen en este tipo de procesos, el proceso de Kendall  $\alpha(t)$  definido la sección 3.2.2 es parte de ellos y  $|C_n - C|$  también guarda ciertas similitudes; de ahí que se haya preferido esta ruta. Con lo anterior se logrará un replanteamiento no tanto de los estadísticos sino de los valores críticos y evitar así una carga teórica mayor.

Se denomina proceso empírico de una muestra  $X_1, X_2, \dots, X_n$  de la distribución  $F$  a la función  $b_n(x) = \sqrt{n}(F(x) - F_n(x))$  donde  $F_n = \frac{1}{n} \sum_{i=1}^n 1(X_i \leq x)$  es la función empírica de la muestra. La característica esencial de estos es su convergencia bajo condiciones de regularidad a un proceso gaussiano o de ruido blanco.

Para entender esta característica considérense las siguientes observaciones sobre el proceso cuando la muestra es de tamaño 1 :

1. el proceso se reduce a  $b_X(x) = 1(X \leq x) - F(x)$
2.  $E[b_X(x)] = E[1(X \leq x) - F(x)] = 0$
3.  $E[b_X(x)b_X(y)] = E[1(X \leq x)1(X \leq y) - 1(X \leq x)F(y) - F(x)1(X \leq y) + F(x)F(y)] = F(x \wedge y) - F(x)F(y)$
4.  $b_n(x) = \sum_{i=1}^n b_{X_i}(x)$

<sup>14</sup>Savu, C. , Trede, M (2004). *Goodness-of-fit tests for parametric families of Archimedean copulas*. Institute for Econometrics, University of Münster.

Por una parte si se recuerda el teorema central del límite afirma que la suma estandarizada de variables aleatorias converge a  $N(0,1)$ , de manera similar, la expresión de (4) convergerá a un proceso normal con la salvedad de que en este caso se habla de un caso multivariado. Si observamos (2) y (3) nos daremos cuenta de que el proceso tiene media cero y una covarianza definida lo cual nos habla de una características esencial de los procesos brownianos. Adicionalmente, en términos formales, se dice que  $b_n$  converge a un puente browniano asociado a  $F$  y se denota por  $b^F$ .

Ahora, ¿cómo es que estas características nos ayudan a construir estadísticos y valores críticos para probar la hipótesis  $h_0 : F = F_n \equiv F_n - F = 0$  contra la alternativa  $h_1 : F \neq F_n$ ?

Por lo dicho con anterioridad, si cumple  $h_0$  entonces el proceso  $b_n$  definido será convergente mientras que si se cumple  $h_1$  el proceso diverge casi seguramente.

Más aún, supóngase que el proceso no es convergente, es decir siempre tendremos incertidumbre (a veces más o menos, ni siquiera eso sabemos) sobre la diferencia entre lo que conocemos (la distribución empírica  $F_n$ ) y lo que creemos sobre la muestra (que proviene de una distribución  $F$ ). Bajo este escenario no es posible establecer bandas que nos permitan decidir si  $F_n$  se parece a  $F$ .

Con esto aclarado, presentamos a continuación dos estadísticos que permiten contrastar la hipótesis :

$h_0$  : La cópula asociada a la muestra bivariada pertenece a una familia con parámetro  $\theta$  no determinado de antemano

$h_1$  : La cópula asociada a una muestra bivariada no pertenece a esta familia con parámetro no determinado de antemano .

Básicamente los estadísticos presentados serán más adecuados que los de la sección anterior para el contraste de esta hipótesis al permitir la estimación del parámetro de la cópula y además, al ser construidos en términos de los procesos empíricos son consistentes y suficientes. La obtención de los valores críticos nos remite nuevamente al bootstrapping, comentado previa la exposición del Algoritmo 5.

### 3.4.5. Kolmogorov-Smirnov modificado

La forma original del estadístico K-S para funciones continuas y univariadas se presentó a principios de esta sección. Para el caso en que nuestra función de distribución empírica es  $C_n$  y nuestra distribución continua es  $C$ , considerando los rangos de las observaciones tendremos, dada la convergencia presentada en 2.5.5:

$$Tn_C = \sup_{(u_i, v_i)} | (C_n(u_i, v_i) - C(u_i, v_i)) | \quad (3.8)$$

y si consideramos una muestra, podemos utilizarlo haciendo uso de la igualdad 3.1 para obtener las observaciones  $(u_i, v_i)$ .

De manera similar, para el ajuste de  $K_n$  se llega a:

$$Tn_K = \sqrt{n} \sup_t | K_n(t) - K_C(t) | \quad (3.9)$$

recordando que  $K_C(t) = K(t; \theta)$ .

Este último estadístico es prácticamente idéntico al mostrado por Genest <sup>15</sup> difiriendo únicamente en los términos evaluados. Para ser más precisos, evalúa  $\frac{j}{n}$  y  $\frac{j+1}{n}$  con  $i = 0, 1$  y  $j = 1, 2, \dots, n-1$  en lugar de valores diferentes de  $t \in (0, 1)$ . Con ello se emplean únicamente las observaciones presentadas de modo puntual, mientras que el acercamiento a través de  $t$  permite extender la distribución de  $K_n$  de forma continua -bajo la obtención de la distribución empírica continua- y a partir de ello obtener las distancias bajo la partición del intervalo unitario que se desee.

En cuanto al primer estadístico, su empleo lo sugiere Muriel <sup>16</sup> como estadístico de prueba para la hipótesis planteada al principio de esta sección, pero únicamente se implementa para el contraste de independencia ya comentado en la sección anterior. Respecto a esto, la implementación del estadístico para cópulas paramétricas difiere al de la cópula producto precisamente por la necesidad de incorporar la estimación del parámetro dentro de la obtención de los valores críticos.

Debido a esto, la obtención de los valores críticos a través de simulación debe seguir el siguiente algoritmo construido bajo el sustento de bootstrapping en dos pasos de Genest y Rémillard: <sup>17</sup>

#### Algoritmo 8. Obtención de valores críticos y cuantiles

1. Generar una muestra grande de  $C_\theta$
2. Obtener un estimador consistente  $\hat{\theta}$  de  $\theta$
3. Generar  $N$  muestras de tamaño  $n$  de  $C_{\hat{\theta}}$
4. Calcular el estadístico deseado para cada una de las muestras y ordenarlos de menor a mayor.  $T_{(1)} \leq T_{(2)} \leq \dots \leq T_{(N)}$
5. El estadístico  $T_{\lfloor (1-\alpha)N \rfloor}$  es el valor crítico con nivel  $\alpha$  de la prueba.
6. El cuantil para el estadístico observado ( $Tn$ ) quedará dado por

$$\frac{1}{N} \sum_{i=1}^N \mathbf{1}(T_{(i)} \geq Tn)$$

*Nota: la hipótesis nula se rechaza siempre que  $Tn > T_{\lfloor (1-\alpha)N \rfloor}$ , o bien, cuando el cuantil de  $Tn$  es menor que  $\alpha$ .*

A continuación presentamos una tabla obtenida a partir del algoritmo anterior para la cópula Frank para muestras de diferentes tamaños, dos niveles de confianza y algunos valores del parámetro.

<sup>15</sup>Genest, C., Quessy, J-F., Remillard, B. (2005) 'Goodness-of-fit Procedures for Copula Models Based on the Probability Integral Transformation.' *Board of the Foundation of the Scandinavian Journal of Statistics*, Blackwell, Vol 32.

<sup>16</sup>Muriel, Op. Cit. p107

<sup>17</sup>Genest, C., Rémillard, B. (2005). 'Validity of the Parametric Bootstrap for Goodness of Fit Testing in Semiparametric Models'. *Les Cahiers du GERAD*, G-2005-51.



$\theta$	Cuantil		
	90 %	95 %	99 %
-50	0.0337	0.0361	0.0437
-25	0.0445	0.0475	0.0549
-10	0.0680	0.0729	0.0817
-5	0.0715	0.0779	0.0888
-2	0.0887	0.0958	0.1101
2	0.0654	0.0702	0.0793
5	0.0538	0.0581	0.0671
25	0.0436	0.0471	0.0545

Cuadro 3.2: Valores críticos de  $T_{n_C}$  para prueba cópula Frank mediante 10,000 ensayos Monte Carlo con tamaño de la muestra  $n=100$

### 3.5. Problemas en el ajuste

Hemos mostrado algunos criterios para la selección y el ajuste de una cópula a una muestra bivariada. En esta última sección presentamos algunas de las limitantes identificadas tanto en el proceso de estimación como en el de ajuste.

Por una parte, dentro de la estimación, asumimos que la cópula tiene derivada, o bien, su parámetro queda en términos de alguna de sus medidas de concordancia. Para las cópulas que presentamos en el último capítulo no hay problema en este respecto, ya que la estimación para todas es posible mediante al menos alguno de los métodos. Sin embargo es posible que en determinado momento requiramos emplear o construir una cópula que no tenga derivada y cuya medida de concordancia guarde una relación complicada con su parámetro, lo que nos muestra que los procedimientos aquí presentados no son infalibles.

Otro aspecto importante es el tamaño de la muestra. Con anterioridad mostramos numéricamente la consistencia del estimador obtenido a partir de la tau de Kendall y un ejercicio similar puede hacerse para el de máxima verosimilitud canónica evidenciando la importancia de muestras grandes para una estimación más exacta. Para los primeros, notamos una desviación estandar importante para muestras con cien observaciones, misma que se vuelve alarmante si se requiere trabajar con poca información.

Ya en la parte del ajuste, para la obtención de los estadísticos de contraste, es indispensable el empleo de herramientas computacionales tanto para la obtención del estadístico muestral, como para la identificación de la región crítica a partir del Algoritmo 8, lo que en términos globales se traduce en una inversión importante de tiempo pues es necesario realizar este proceso para cada uno de los modelos y parámetros considerados.

Para finalizar, hacemos cuenta de la limitante más importante que representa el empleo de cópulas y que tiene origen precisamente en la cantidad de modelos existentes y que pueden crearse a partir de lo expuesto en las secciones 1.5 y 1.6. La cópula que seleccionemos muy probablemente no es la cópula con mejor ajuste que existe y lo que es peor, no hay manera de probar lo contrario. El único respaldo con el que contamos es el de realizar la mejor selección dentro de un grupo de alternativas diversas, lo que a su vez nos lleva al problema de

cómo conformarlo, para lo cuál tampoco hay una respuesta definitiva.

Si bien este es un problema grave, el empleo de los estadísticos modificados reduce su impacto de una manera crucial. Al usarlos, tendremos la certeza de que la cópula seleccionada es suficiente para modelar nuestra muestra aunque no podemos asegurar que es la mejor que existe. En el siguiente capítulo se presenta un caso en el que efectivamente se 'encuentra' una cópula con mejor ajuste que el logrado a partir de un grupo inicial, sin embargo, los resultados obtenidos son muy similares.

## Capítulo 4

# Aplicación de las cópulas: Valor en Riesgo

En este capítulo finalmente llegamos a la parte práctica de la presentación. Como se podrá imaginar existe una variedad de posibles aplicaciones de toda la teoría revisada con anterioridad. Si bien ésta es de utilidad para atacar casi cualquier problema en que estén inmiscuidas distribuciones de probabilidad conjuntas, dependerá mucho de las herramientas que tengamos a nuestra disposición y, desde luego del problema que se trate.

Teniendo esto en consideración, además de su indiscutible prevalencia en la actualidad, se ha preferido abordar algunas aplicaciones financieras por mantenerse en un justo medio entre aplicación práctica y requerimientos teóricos, ya que por una parte permiten el empleo de una carga teórica importante y por otra, no dejan de ser ejemplos ilustrativos y útiles.

Revisaremos varias aplicaciones: el cálculo del Valor en Riesgo (VaR por sus siglas en inglés ) para un portafolio de dos activos, indicadores de riesgo para un portafolio en términos de la estructura de dependencia existente, el cálculo de la prima de reaseguro, el cálculo de la prima de un seguro de vida conjunta, valuación de opciones con dos activos subyacentes, regresión y finalmente procesos estocásticos.

Como se observará cada uno de estos temas tiene un perfil distinto aunque mantienen en común la necesidad de modelar dependencia bajo supuestos no normales. Cabe mencionar que únicamente desarrollamos de manera íntegra la primera de las aplicaciones listadas con el propósito de emplear gran parte de los conceptos teóricos presentados. Para el resto de ellas, únicamente se esboza la motivación para el empleo de las cópulas y además de algunas consideraciones sobre su uso.

De esta manera, las tres primeras secciones corresponden al cálculo del VaR de dos portafolios integrado por acciones de empresas que cotizan en la Bolsa Mexicana de Valores. Tratamos como primer punto algunas de las alternativas ya existentes para este cálculo y sus diferencias fundamentales con el empleo de las cópulas.

Posteriormente realizamos la elección de la cópula a partir de los procesos de

estimación, ajuste, revisión gráfica y modo de dependencia de la muestra. Finalmente, presentamos una comparación de los resultados obtenidos a partir de la implementación de la cópula elija con los hechos bajo supuestos clásicos. Por último, la sección 4.4 quedará destinada para la revisión del resto de las aplicaciones.

## 4.1. Valor en Riesgo

La cuantificación del riesgo es por sí misma un tópico de gran interés y podríamos asegurar que para un par de instrumentos (activos, derivados, créditos, etcétera) existe una infinidad de factores que influyen directamente sobre el riesgo que representa su tenencia (temporalidad, volatilidad, tasas de interés, política, etcétera). Si se nos pregunta por ejemplo, cuál es la pérdida máxima que podríamos sufrir en una inversión, la única respuesta genuina que podríamos dar desafortunadamente es *todo* aunque posiblemente esto suceda con una probabilidad baja.

Un acercamiento a la cuantificación del riesgo proviene de la teoría moderna de portafolios al encapsularlo en las desviaciones sufridas por el rendimiento del portafolio (desviación estandar) que prevalece aún con la diversificación (riesgo sistemático). Actualmente existen una variedad de alternativas para realizar esta medición: el análisis de sensibilidad, cash flow en riesgo y el valor en riesgo son algunas de ellas.

### 4.1.1. Conceptos importantes

El VaR surge en las últimas dos décadas; su difusión y aceptación ha sido relativamente rápida debido a los esfuerzos internacionales por proveer una base sólida para regulación de riesgos en Basilea I y Basilea II; la reunión del Grupo de los Ocho en la década de los noventa y desde luego a la introducción de la aplicación RiskMetrics de J.P Morgan que automatiza el procedimiento requerido.

Conceptualmente, el VaR refina nuestra respuesta inicial la interrogante de la pérdida máxima al cuantificar la probabilidad de ocurrencia. Podemos decir que el valor en riesgo es el monto máximo que se perderá en un portafolio en un determinado horizonte de tiempo con una probabilidad definida. La elección más común para esta probabilidad es de 90 %, 95 %, ó 99 %, y en general se encuentra dada por el grado de certidumbre que se desee y por las disposiciones legales que instituciones reguladoras (Banco de México por ejemplo) soliciten en el análisis.

Si tomamos por ejemplo un nivel de confianza de 99 % el analista observará -tras el periodo elegido- que las pérdidas reales solamente sobrepasan en alrededor de un 1 % el VaR. Si por el contrario, se elige 90 % las pérdidas se sub estiman en alrededor de un 10 %.

En resumen, si decimos que el valor en riesgo al 99 % de nuestro portafolio para el próximo mes es de un millón, es equivalente a afirmar que para el próximo mes nuestra pérdida máxima será de un millón y esto sucedará una de cien veces.

Estadísticamente, el VaR con un nivel de confianza del  $(1 - \alpha)$  corresponde al cuantil  $(1 - \alpha)$  de la distribución de rendimientos, es decir, si  $Z$  es la variable aleatoria que denota los rendimientos del portafolio:

$$\text{VaR}_{(1-\alpha)} = z^* \quad \text{donde } z^* \text{ es tal que } P[Z \leq z^*] = 1 - \alpha$$

Ahora bien, la distribución de  $Z$  está determinada por la distribución conjunta de los instrumentos que compongan el portafolio. Para el caso bivariado, si asumimos que  $X$  y  $Y$  son las variables aleatorias continuas que denotan los retornos de cada uno de ellos para un horizonte  $T$  común, ponderamos nuestra inversión en ellos como  $\gamma X + (1 - \gamma)Y$  y asumimos que conocemos la función de distribución conjunta  $H$ , podemos encontrar a través de una transformación:

$$P[Z \leq z] = P[\gamma X + (1 - \gamma)Y \leq z]$$

o bien, si condicionamos sobre  $Y$  tenemos:

$$\begin{aligned} P[\gamma X + (1 - \gamma)Y \leq z] &= \int_{-\infty}^{\infty} P[X \leq (1/\gamma)z - (1 - \gamma)/\gamma y, Y = y] dy \quad (4.1) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{(1/\gamma)z - (1 - \gamma)/\gamma y} h(x, y) dx dy \end{aligned}$$

y nuevamente buscaremos  $z^*$  tal que la probabilidad anterior sea igual a  $1 - \alpha$ .

#### 4.1.2. Cálculo del VaR

Teóricamente el VaR tiene una expresión relativamente simple. Sin embargo, en la práctica, la distribución  $H$  es en general inaccesible y por ello comúnmente es necesario incorporar otras herramientas y supuestos sobre ella. A continuación se presentan los tres métodos más comunes para resolver este problema:

- *Varianza-covarianza*: parte el supuesto de que los dos primeros momentos caracterizan la distribución y se asume un modelo específico para ella, como por ejemplo: Normal, Gamma, Beta, entre otros. A partir de ello, se estima 4.1.
- *Simulación histórica (SH)*: este no hace supuestos sobre la distribución conjunta pero asume que la distribución de ganancias y pérdidas sigue un comportamiento descrito por la historia. Partiendo de distintos escenarios históricos obtiene las ganancias y pérdidas que se obtendrían para el horizonte determinado con el portafolio actual. Repitiendo esto muchas veces, aproxima numéricamente la distribución real de ganancias y toma el cuantil  $(1 - \alpha)$ .
- *Simulación Monte Carlo (SMC)*: Al igual que la herramienta anterior, aproxima numéricamente la distribución de ganancias a partir de la simulación de una gran cantidad de observaciones de una distribución 'apropiada' -tradicionalmente una distribución normal multivariada- para los riesgos dentro del portafolio. Una vez hecho esto, se calculan los retornos a partir de las observaciones anteriores y se ordenan quitando la fracción  $\alpha$  de los peores retornos. El mínimo de los retornos restantes corresponde al VaR del portafolio.

Un estudio de las características y la eficiencia de estos métodos puede encontrarse en Hull <sup>1</sup> Para fines de nuestra presentación, sólo haremos algunas observaciones sobre ellos.

Primeramente, como desventaja del método de Varianza-Covarianza está el supuesto sobre la distribución, que limita nuestras posibilidades para modelar los comportamientos de los rendimientos y su interdependencia. Para instrumentos no lineales -como algunos derivados- la precisión de éste es bastante cuestionable. Como puntos favorables destacan la rapidez con que se calcula y la flexibilidad para incorporar nuevas observaciones o restringir el análisis a determinado tiempo e instrumentos.

En contraste, las dos últimas alternativas requieren de simulación -aunque el último de ellos requiere de un número de ensayos significativamente mayor- lo que representa una inversión importante de tiempo. La simulación histórica por una parte no requiere que se asuma una distribución específica, lo que nos da cierta ventaja sobre el primero. Sin embargo es muy sensible a la información histórica y a la ponderación que se de a esta. Si por ejemplo el día  $d_0$  ocurre una pérdida importante y nuestra ventana de escenarios sólo incorpora los últimos cinco años de historia, exactamente cinco años después de  $d_0$  el evento no será considerado y habrá una variación importante en el VaR estimado el día anterior.

La Simulación Monte Carlo resulta ser el método más sofisticado entre los anteriores. Por una parte permite asumir una distribución para el cálculo estimada a partir de la historia -evitando el supuesto rígido de normalidad- y con ello la posibilidad de actualizar esta distribución de acuerdo a determinados escenarios. Además de esto, amortigua el peso de la historia incorporada evitando que ocurra algo similar a SH. Lo anterior tiene su costo en el tiempo a invertir pues en general este método es el más tardado para portafolios relativamente grandes, aunque existen ya algunas estrategias para hacerlo eficiente <sup>2</sup>.

Para resumir los pasos que deben seguirse cuando se calcula el VaR empleando SMC es necesario seguir el siguiente procedimiento.

### Algoritmo 9

Dadas  $N$  observaciones de un portafolio con  $n$  instrumentos con un horizonte  $t=1$  y confianza  $1 - \alpha$

1. Calcular las tasas relativas de cambio de acuerdo al horizonte de tiempo requerido. Por ejemplo, si se desea conocer el VaR de  $t=1$ , las tasas requeridas son  $r_{i,j} = \frac{x_{i,j} - x_{i,j-1}}{x_{i,j-1}}$  donde  $i = 1, 2, \dots, n$  y  $j = 1, 2, \dots, N$
2. Estimar la distribución de las  $r_{i,j}$
3. Simular  $m$  ensayos de las  $n$  tasas relativas a partir de la distribución del punto anterior, denotados por  $r_i^k$  con  $k = 1, 2, \dots, m$  y  $i = 1, 2, \dots, n$

<sup>1</sup>Hull, J. Options, Futures & other Derivatives. 4a Ed. Prentice-Hall. 2000.

<sup>2</sup>Algunas herramientas están destinadas a generar instrumentos artificiales que resuman la información del riesgo en varios grupos de instrumentos; otras más, como las técnicas de reducción de varianza, se ocupan de hacer más eficiente la simulación, acelerando la convergencia de ella y reduciendo así el número de ensayos requeridos y aumentando la precisión.

De acuerdo a lo anterior, los valores simulados de los instrumentos del portafolio corresponden a  $x_{i,N+1}^k = x_{i,N}(1 + r_i^k)$  con  $k = 1, 2, \dots, m$  y  $i = 1, 2, \dots, n$ ,

4. Ordenar los vectores  $x_i^k$  de acuerdo a la suma de sus pérdidas en orden ascendente y eliminar la proporción  $\alpha$  de las pérdidas mayores.
5. El VaR corresponde a la pérdida del último de los vectores restantes.

En el algoritmo anterior, particularmente los puntos 2 y 3 son los que representan el reto teórico y práctico más importantes. Por una parte, la estimación o elección de una distribución multivariada no es simple y una vez que esta se ha identificado, la simulación de observaciones provenientes de ella también puede complicarse.

## 4.2. Empleo de una cópula para el cálculo del VaR

Como alternativa natural podemos emplear a las cópulas. La ventaja fundamental de su uso es la separación de los problemas a los que nos enfrentamos. Por una parte, nos permiten extraer la estructura de dependencia existente en la muestra y con ello la simulación de observaciones provenientes de esta, y por otra, la incorporación distribuciones marginales adecuadas no necesariamente idénticas.

Para ejemplificar lo anterior, mostramos a continuación la obtención del VaR para dos portafolios constituidos por acciones de Cemex(cmx), Telmex(tmx) y Televisa(tvsa). Para ello, básicamente seguiremos el Algoritmo 9, dando un gran peso al punto 2. Como parte de él, revisaremos el perfil estocástico de la muestra, realizaremos la estimación de las distribuciones márginales y de los parámetros de la cópula y las pruebas de bondad de ajuste necesarias.

### 4.2.1. Perfil de la muestra y estimación de los parámetros

La muestra consta de un total de 475 observaciones de octubre de 2002 a octubre de 2004 para cada una de las acciones tomando el precio al cierre de las operaciones. Los portafolios a considerar son [Televisa, Telmex] y [Televisa, Cemex] con una ponderación idéntica sobre las acciones contenidas. Para iniciar nuestro análisis, observemos el comportamiento de los precios al cierre de cada una de las acciones.

Para las series de Cemex y Televisa se observa una tendencia a la alza mientras que para Telmex una tendencia decreciente del precio de la acción. Comparativamente, el precio de la acción de Cemex mostró mayor volatilidad.

Revisemos ahora los rendimientos para cada una de las acciones. De antemano, sabemos que en general es difícil que una distribución normal sea adecuada para los rendimientos de este tipo de activos y por ello incorporamos para el estudio la distribución t de Student no central.

En la figura 4.2 se observan los ajustes para los rendimientos de las acciones de Cemex (arriba a la izquierda), Telmex(arriba a la derecha) y Televisa (abajo).

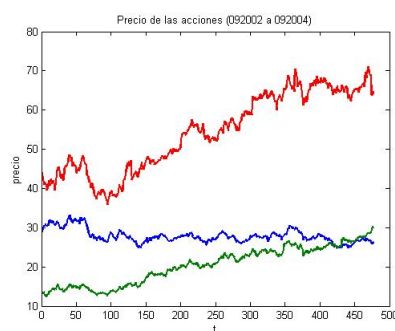


Figura 4.1: Precios de las acciones

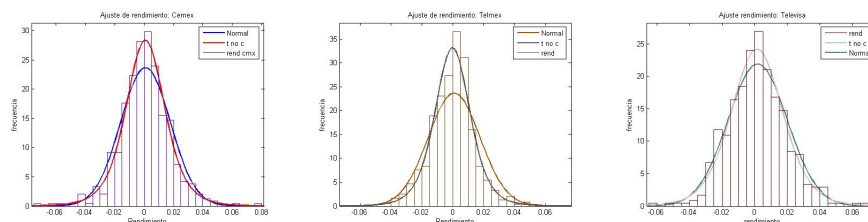


Figura 4.2: Ajustes para los rendimientos

En ellas, el histograma representa las frecuencias muestrales y las líneas continuas y punteadas los ajustes para distribución  $t$  no central y distribución Normal respectivamente. Notamos que los rendimientos de cada uno de los activos quedan descritos de una manera más adecuada por la distribución  $t$  no central que permite alcanzar frecuencias más altas tanto en la media de la distribución como en las colas. Lo anterior queda confirmado en el cuadro 4.1 al mostrarse que el valor de log-verosimilitud es mayor en todos los casos en comparación con el logrado por el ajuste con distribución Normal. Aquí se percibe ya una limitante de importancia de los supuestos de normalidad ya que en términos del comportamiento marginal de los activos, no se capturan las colas de la distribución, lo que es determinante para el cálculo del VaR pues implícitamente se asigna una menor frecuencia a pérdidas y ganancias grandes, lo que puede llevarnos a subestimar el valor real. En la figura 4.3 podemos observar una gran dispersión en el portafolio Televisa-Telmex lo que nos orilla a pensar que existe poca relación entre los rendimientos de los activos en este portafolio. En contraste, el de Televisa-Cemex muestra una tendencia un poco más clara al permanecer en una banda con pendiente uno y de hecho se observa que a rendimientos positivos (o negativos) en Televisa corresponden rendimientos positivos (o negativos) en Cemex. Lo anterior no es claro para el primer portafolio aunque también observamos una concentración de datos cerca de los rendimientos nulos -lo que es congruente con las estimaciones hechas en el cuadro 4.1 y representa una gran frecuencia de este suceso- y en conjunto implica que si una de las variables se anula, también la otra lo hará.

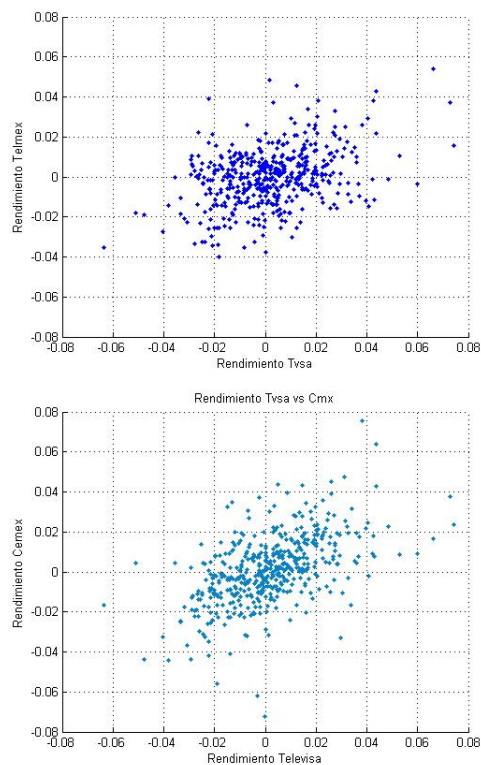
Realicemos ahora un análisis formal empleando las medidas de concordancia revisadas en la sección 2.3 y el coeficiente de correlación de Pearson sobre el



Distribución	Rendimiento	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\nu}$	Log-likelihood
Normal	Cemex	0.000929	0.016868	-	1268.26
	Telmex	-0.000116	0.014032	-	1355.89
	Televisa	0.001840	0.018241	-	1231.01
t no central	Cemex	0.000848	0.013423	5.3337	1283.65
	Telmex	-0.000144	0.011529	5.8468	1364.45
	Televisa	0.001415	0.016057	8.9825	1236.88

Cuadro 4.1: Parámetros estimados para distribución Normal y t no central. (Media, Varianza y Grados de libertad)

Figura 4.3: Arriba: Rendimientos de Televisa vs Telmex. Abajo: Rendimientos de Televisa vs Cemex



Cuadro 4.2: Estadísticos de correlación y concordancia en cada uno de los portafolios

Portafolio	$\rho$ Pearson	$\tau$ Kendall	$\rho$ Spearman
tmx - tvsa	0.3672	0.2046	0.2978
cmx - tvsa	0.5376	0.3944	0.5523

que haremos una observación importante una vez que se hallan transformado los datos.

Los resultados se muestran en el cuadro 4.2 y de acuerdo con esto puede afirmarse la existencia de dependencia de cuadrante positivo en ambos portafolios y una mayor probabilidad de concordancia en el segundo. El coeficiente de Pearson también muestra una relación positiva en ambos portafolios aunque en el segundo es mayor. Como mostramos en el segundo capítulo, ninguno de estos estadísticos nos permite aceptar la independencia en los portafolios, por el contrario, parecen indicar que existe dependencia.

Para refinar nuestro estudio y hacer uso de la teoría de cópulas, asumamos que la cópula de las variables del primer portafolio es  $C_1$  y la del segundo es  $C_2$ . Por el Corolario 1.3.1 es factible extraer las estructuras de dependencia anteriores que en términos prácticos se traduce en el empleo de la transformación mostrada en la igualdad 3.1 con la cual obtendremos las variables  $U_i$ ,  $V_i$  y  $W_i$  asociadas a los rendimientos de Cemex, Telmex y Televisa y que se distribuirán de manera casi uniforme.

Cuadro 4.3: Estadísticos de correlación, concordancia e independencia en cada uno de los portafolios bajo la transformación (3.1)

Portafolio	$\rho$ Pearson	$\tau$ Kendall	$\rho$ Spearman	$\ C_n - \Pi\ $
tmx - tvsa	0.2975	0.2046	0.2978	0.0587
cmx - tvsa	0.5523	0.3944	0.5523	0.1146

Hecha esta transformación es más fácil observar el comportamiento conjunto de las variables. En la Figura 4.4 a la izquierda se muestra la estructura de dependencia del portafolio Televisa-Telmex y a la derecha el de Televisa-Cemex. Como se anticipaba, los rendimientos en el segundo portafolio tienen una dependencia más fuerte que la del primero al tener una mayor concentración de observaciones hacia las esquinas y en la diagonal. El primer portafolio muestra también cierta concentración aunque no tan definida como para el anterior.

Los estadísticos para los datos transformados en el Cuadro 4.3 muestran para el primer portafolio un decremento del coeficiente de correlación en contraste con el aumento marginal para el segundo y -como se esperaba por lo expuesto en el segundo capítulo- las medidas de concordancia permanecieron sin cambios tras la transformación.

Con lo anterior queda expuesto de una manera más clara cómo el coeficiente de Pearson es susceptible a la variación en las marginales mientras que las medidas de concordancia permanecen sin cambios ante ello, revelando su relación única con la estructura de dependencia.

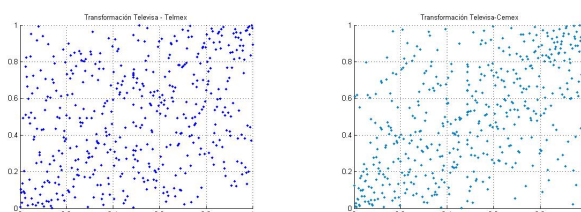


Figura 4.4: Dependencia en el rendimientos de los activos de cada portafolio bajo la transformación (3.1)

Haciendo empleo del estadístico de independencia presentado en 2.5.2 generado a partir de la cópula empírica, observamos que para el primer portafolio está más cercano a cero mientras que el segundo se aleja bastante, lo que corrobora nuestras afirmaciones anteriores. Ahora bien, si comparamos estos valores con los valores críticos de una muestra de tamaño 450, mostrados en el cuadro 2.2, podemos rechazar la existencia de independencia para ambos portafolios tanto al 90 % como al 95 % y al 99 % de confianza.

#### 4.2.2. Bondad de ajuste y elección de la cópula

A partir de lo revisado en la sección anterior podemos concluir dos cosas de interés que motivan el empleo de los conocimientos expuestos en el tercer capítulo: a) *no puede aceptarse la existencia de independencia* y b) *no puede aceptarse la existencia de independencia* al menos las marginales no siguen una distribución normal, lo que en términos conjuntos impiden que se emplee una distribución normal bivariada para modelar los retornos de ambos portafolios.

Dentro de nuestro estudio consideraremos cinco familias arquimideanas y dos no arquimideanas que pueden encontrarse en el apéndice. Todas ellas permiten modelar dependencia de cuadrante positivo y sólo dos de ellas con compressivas.

La estimación paramétrica se realizará de acuerdo a los métodos de máxima verosimilitud canónica y Genest Rivest sobre la medida de concordancia de Kendall.

Para la revisión gráfica en el primer portafolio optaremos por la alternativa expuesta en la sección 3.3.1 de la distribución condicional que nos permite comparar el ajuste para cópulas arquimideanas y no arquimideanas los estimadores obtenidos a partir de la tau de Kendall con excepción de la familia Gumbel-Barnett en el que se toma el máximo verosimil. Esta prueba no pudo realizarse para la familia Cuadras-Augé pues no tiene derivada en  $u = v$ ; para el resto se observa un buen ajuste y al realizarse las pruebas K-S no se puede rechazar uniformidad en ninguno de los casos, pues los cuantiles (p-values) de estas pruebas son altos: Ali-Makhil (0.8318), Clayton (0.8174), Gumbel-Hougaard (0.9766),

Gumbel Barnett (.9230), Frank (0.9923) y Ferlie-Morgensten (0.9934).

Con la información anterior es muy difícil optar abiertamente por alguna de las cópulas y de hecho, si basamos nuestra elección en aquella que obtuvo un estadístico mayor, optaríamos por la última de las alternativas que como se verá no coincide con nuestra conclusión final.

Adelante se muestran los estimadores obtenidos a partir del método de Genest Rivest sobre la tau de Kendall, los estimadores máximo verosímiles y los estadísticos de distancia (3.8) y (3.9).

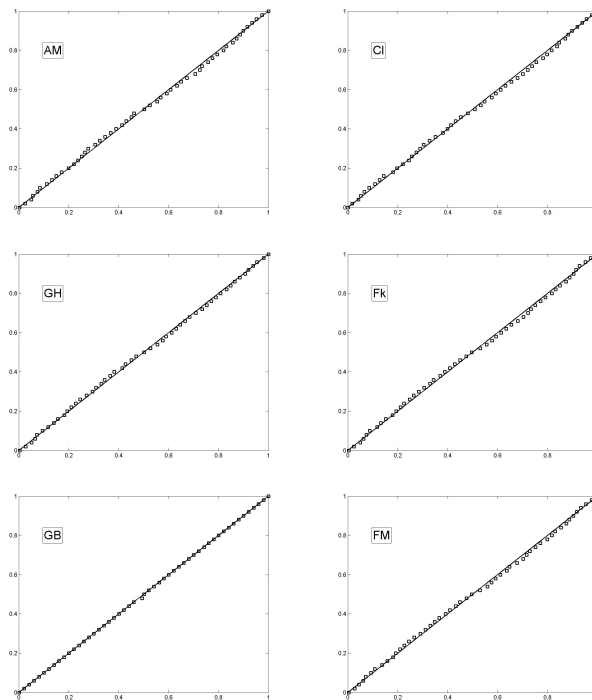


Figura 4.5: Revisión gráfica (3.3.1) para el primer portafolio

De acuerdo con estos, la familia Gumbel-Hougaard es la que logra el mejor ajuste al poseer los estadísticos  $Tn_C$  y  $Tn_K$  de distancia más pequeños en el grupo tanto para el estimador Genest-Rivest como para el máximo verosímil. De hecho, si observamos el renglón correspondiente a esta familia, notamos que los parámetros estimados son muy parecidos entre sí con una diferencia de menos de una centésima, lo cuál, como se mencionó en el capítulo anterior, era un criterio importante para la elección de la cópula.

Si obedecemos únicamente los valores de  $Tn_C$  la segunda mejor opción resulta ser la familia Frank mientras que si tomamos solo a  $Tn_K$  es la Ali-Mikhail-Haq, que de hecho es la segunda familia con estimadores más parecidos.

Observamos también que el ajuste de la cópula Frank no es de las mejores alternativas de acuerdo con  $Tn_K$ . En contraste AM ocupa un lugar intermedio en las alternativas de acuerdo con  $Tn_C$ . Para las familias no arquimedeanas, notamos

que la familia Ferlie-Gumbel-Morgensten ocupa el tercer sitio de acuerdo con  $Tn_C$  para ambos estimadores con diferencias de alrededor de 3 milésimas con la cópula Frank lo cual muestra la ineficacia del criterio a partir del cuantil obtenido en la prueba Kolmogorov-Smirnov antes mencionada.

Familia	Genest-Rivest			Maxima Verosimilitud C.		
	$\theta_\tau$	$Tn_C$	$Tn_K$	$\hat{\theta}$	$Tn_C$	$Tn_K$
AM	0.7259	0.0319	0.9189	0.714	0.0322	0.8823
Clayton	0.5145	0.0357	0.9409	0.4041	0.0383	0.9355
GH	1.2572	0.0228	0.8103	1.2609	0.0226	0.7999
Frank	1.9068	0.0258	0.9931	1.9599	0.0268	1.0415
GB	-	-	-	0.0001	0.0587	1.8793
CA	0.3397	0.0334	-	0.9998	0.2058	-
FM	0.9208	0.028	-	0.8310	0.0300	-
Gauss	0.3159	0.0600	-	0.2975	0.0337	-

Cuadro 4.4: Portafolio Televisa-Telmex: Estimadores a partir de tau de Kendall ( $\theta_\tau$ ) y máxima verosimilitud ( $\hat{\theta}$ ) y estadísticos de ajuste ( $Tn_C$  y  $Tn_K$ ) para cada uno de ellos.

Obteniendo los valores críticos para los estadísticos máximo verosímiles de las familias Gumbel-Houggard, Frank y Ali-Makhil, observamos que únicamente podemos rechazar bajo  $Tn_K$  a la familia Frank con  $\alpha = 10\%$ . El resto de los valores muestran que tanto GH como AM no pueden rechazarse con ninguno de los estadísticos bajo ningún nivel de significancia. Se concluye que el mejor ajuste corresponde a la cópula Gumbel-Houggard con el estimador máximo verosímil canónico  $\hat{\theta}=1.2609$ . Además, hacemos la observación de que la cópula de Gaussiana ocupa uno de los últimos sitios de acuerdo con el estadístico de distancia.

Familia ( $\hat{\theta}$ )	cuantil $Tn_C$		cuantil $Tn_K$	
	90 %	95 %	90 %	95 %
AM ( 0.7140)	0.0346	0.0375	1.1571	1.2817
GH (1.2609)	0.0359	0.0389	0.9814	1.1034
Frank (1.9599)	0.0380	0.0406	0.9807	1.0783

Cuadro 4.5: Valores críticos para los estimadores de máxima verosimilitud canónica de algunas familias (n=475 tamaño de la muestra y N=1000 ensayos Monte Carlo)

Estudiemos ahora la estructura de dependencia del segundo portafolio recordando que de acuerdo con el Cuadro 4.3 también se rechaza la independencia en él. Para ello, nuevamente optaremos al inicio por una revisión gráfica sólo que en esta ocasión basándonos en (3.3.3) para las cópulas arquimedeanas.

En la Figura 4.6 se observan  $\lambda$  y  $\lambda_n$  para las cinco familias arquimedeanas consideradas. Notamos a partir de ella que los mejores ajustes se logran para las familias Gumbel-Houggard y Frank mientras que las cópulas Ali-Makhil y Clayton se alejan bastante de  $\lambda_n$  al igual que la Gumbel-Barnett.

En el Cuadro 4.6 se muestran los parámetros y los estadísticos de distancia para el segundo portafolio para el cual las mejores alternativas son la cópula Gauss y la Gumbel Hougaard seguidas por la familia Frank.

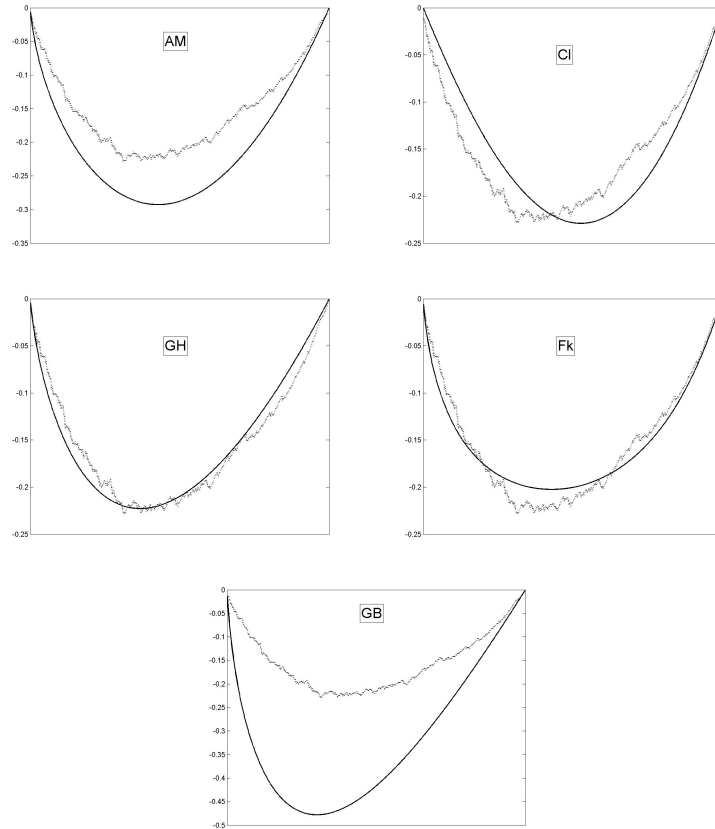


Figura 4.6: Revisión gráfica (3.3.3) para el portafolio Televisa-Cemex

Podemos ver que los estimadores de la cópula Gauss y los de la familia Frank son los que guardan mayor cercanía con una diferencia de menos de 3 centésimas mientras que para GH la diferencia es casi de una décima. Respecto a los estadísticos, vemos que  $Tn_C$  reporta un valor inferior con la cópula Gauss tanto para el estimador máximo verosímil como para el de Genest-Rivest con diferencias marginales respecto a las dos restantes.

De acuerdo con esto, para cualquiera de las tres familias mencionadas el mejor estimador será el obtenido a partir de la tau de Kendall y no el de máxima verosimilitud canónica.

Para terminar nuestro análisis, revisemos los valores críticos de los estadísticos de estas familias (Cuadro 4.7). Cabe mencionar que los valores críticos no se obtuvieron dada la excesiva tardanza que requiere este proceso influenciada principalmente por el cálculo de  $C(u, v)$  en (1.9) que requiere de una aproximación numérica en conjunto con el tamaño de la matriz a calcular (450 x 450). Comparativamente, la obtención de los valores críticos para alguna de las otras

Familia	Genest-Rivest			Maxima Verosimilitud C.		
	$\theta_\tau$	$Tn_C$	$Tn_K$	$\hat{\theta}$	$Tn_C$	$Tn_K$
AM	-	-	-	0.954	0.0509	1.3608
Clayton	1.3023	0.042	1.1729	-0.5049	0.8316	7.274
GH	1.6511	0.0226	0.7415	1.5587	0.0286	0.9328
Frank	4.0839	0.0273	0.6951	4.1188	0.0287	0.7168
GB	-	-	-	0.0001	0.1146	3.5377
CA	0.5657	0.0321	-	0.9998	0.1512	-
FM	1.7746	0.0344	-	1	0.0524	-
Gauss	0.5806	0.0207	-	0.5523	0.0224	-

Cuadro 4.6: Portafolio Televisa-Cemex: Estimadores a partir de tau de Kendall ( $\theta_\tau$ ) y máxima verosimilitud ( $\hat{\theta}$ ) y estadísticos de ajuste ( $Tn_C$  y  $Tn_K$ ) para cada uno de ellos.

Familia ( $\hat{\theta}$ )	cuantil $Tn_C$		cuantil $Tn_K$	
	90 %	95 %	90 %	95 %
GH (1.6511)	0.0425	0.0457	0.9310	1.0184
Frank (4.0839)	0.0358	0.0391	0.9196	1.0284

Cuadro 4.7: Valores críticos para los estimadores Genest-Rivest de algunas familias (muestra  $n=475$ ;  $N=1000$  ensayos Monte Carlo)

familias toma alrededor de 10 minutos para los 1,000 ensayos MonteCarlo mientras que un sólo ensayo para la cópula gaussiana toma un par de minutos.

De acuerdo con el cuadro, ninguna de las cópulas GH y Frank puede rechazarse bajo los niveles de significancia presentados por lo cual se puede elegir libremente alguna de estas dos estructuras.

Aun sin contar con los valores críticos para la cópula Gaussiana, podemos afirmar que esta no ha de rechazarse como estructura de dependencia de este portafolio dado que presenta un ajuste superior al de las dos familias mencionadas y no rechazadas. Por lo anterior, tomaremos a la cópula Gaussiana con el estimador Genest-Rivest como mejor alternativa para el ejercicio del cálculo del VaR, recordando que la cópula Frank y la GH son también buenas alternativas.

### 4.2.3. Obtención del VaR

Retomando el Algoritmo 9 presentado al principio de este capítulo, notamos que hemos cubierto los dos primeros puntos de este mediante la información contenida en los cuadros 4.1, 4.4 y 4.6 recordando que las distribuciones marginales elegidas son t no centrales, las cópulas elegidas en los portafolios son la Gumbel-Hougaard y Gauss con parámetros (1.2609) y (0.5806) respectivamente.

Realizaremos ahora la simulación de las ganancias en los portafolios, para lo cual emplearemos los Algoritmos 1 y 2 del primer capítulo de la siguiente manera:

primero simularemos la estructura de dependencia con marginales uniformes y posteriormente las transformaremos por separado en las marginales  $t$  no centrales.

A partir de lo anterior obtendremos el vector bidimensional  $r^*$  de  $N$  tasas relativas de cambio para el periodo  $n + 1 = 476$  y posteriormente los  $N$  posibles valores para cada uno de los activos en los portafolios multiplicando el valor de la última observación por  $(1 + r_{ij}^*)$  es decir:  $x_{476}^{ij} = x_{475} \times (1 + r_{ij}^*)$   $i = 1, 2$  activos  $j = 1, 2, \dots, N$  ensayos.

Ahora, con  $x_{476}^{ij}$  y  $x_{475}^j$  obtenemos las pérdidas y ganancias posibles de cada activo  $v_{ij} = (x_{476}^{ij} - x_{475}^j)$  y luego las pérdidas y ganancias posibles para el portafolio  $v_j = v_{1j} + v_{2j}$ . Abajo presentamos el gráfico de  $v_j$  para cada uno de los portafolios con  $N = 5000$  ensayos Monte Carlo.

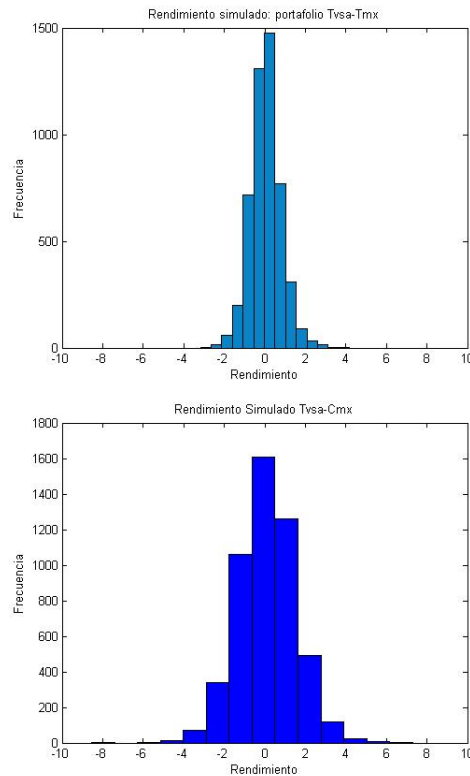


Figura 4.7: Ganancias simuladas para los portafolios.  $N = 5000$  ensayos Monte Carlo. Arriba: portafolio Televisa-Telmex. Abajo: portafolio Televisa-Cemex

Notamos a primera vista que la distribución de ganancias del segundo portafolio posee colas más largas que el primero, lo que repercutirá en un VaR mayor. Complementario a esto, en el primer portafolio se observa una concentración mayor de observaciones cerca del rendimiento nulo en comparación con el portafolio Televisa-Cemex.

Para la obtención del VaR únicamente resta ordenar de forma descendente las  $v_j$  y extraer la fracción  $\alpha$  de peores pérdidas; el último de los  $v_j$  restante es el



VaR al  $(1 - \alpha)\%$  para cada uno de los portafolios. Además del VaR obtenido a partir de las cópulas estimadas también es posible la obtención de las cotas para este a través del empleo de 1.4.1 y del VaR bajo supuesto de independencia empleando a  $\Pi$ . Lo anterior se muestra en el cuadro siguiente.

Portafolio Cópula	Valor en riesgo	
	95 %	99 %
<b>Televisa-Telmex</b>		
GH (1.2609)	1.138	1.6465
$\Pi$ (independencia)	1.017	1.6037
W (cota inferior)	0.28235	0.57817
M (cota superior)	1.4293	2.2902
<b>Televisa-Cemex</b>		
Gauss (0.5806)	2.2498	3.5447
$\Pi$ (independencia)	1.8756	2.8988
W (cota inferior)	0.74262	1.655
M (cota superior)	2.5004	4.1591

Cuadro 4.8: Valor en Riesgo para los portafolios considerados en 5000 ensayos Monte Carlo bajo diferentes estructuras de dependencia.

Observamos que el segundo de los portafolios tiene un VaR mayor para todas las estructuras de dependencia presentadas. Ponderando de acuerdo al valor medio de cada portafolio, considerando únicamente las cópulas GH y Gauss, llegamos a un VaR al 95 % en términos porcentuales de 2.36 % y 3.01 % respectivamente y de 3.42 % y 4.75 % para VaR al 99 %, manteniéndose la primera observación.

El cuadro anterior nos permite identificar la influencia que tiene la especificación de la estructura de dependencia y las marginales en ambos portafolios. El efecto de de estas últimas se observa si se comparan por ejemplo los valores en riesgo obtenidos a partir de la cópula independencia. De acuerdo con esto, el VaR del segundo portafolio portafolio es superior en más de cuatro quintos de unidad debido precisamente por el cambio de la marginales de los rendimientos de Telmex por la de Cemex, recordando que éste era el activo más volátil de acuerdo con lo revisado en la sección 4.1.

En cuanto al efecto de las estructuras de dependencia, es oportuno recordar los resultados del Cuadro 4.3, donde el estadístico de independencia para el segundo de los portafolios, al igual que los coeficientes de correlación de Pearson y Spearman y la tau de Kendall, eran significativamente mayores a los obtenidos en el primer portafolio, lo que en términos prácticos representaba una mayor cercanía del comportamiento conjunto del primer portafolio en comparación al segundo con respecto a la independencia.

En el cálculo del VaR, partiendo nuevamente del valor obtenido bajo el supuesto de independencia, el portafolio Televisa-Cemex tiene una mayor variación respecto al primero una vez que se comparan estos valores con los obtenidos bajo las estructuras correctas para ellos. De esta manera, notamos que el VaR del primer portafolio se incrementa en a penas un décimo de unidad al incorporar la cópula G-H mientras que para el segundo hay un incremento de casi media unidad.

### 4.3. Contraste de los resultados

Hasta este momento se han mostrado el procedimiento y algunos de los resultados obtenidos a partir del empleo de una cópula para la estimación del VaR y también se ha hecho notar que tanto las distribuciones marginales como la estructura de dependencia juegan un papel importante en este cálculo.

El objetivo principal en esta sección, será comparar estos cálculos con los logrados a partir de los supuestos de normalidad y ponderar sobre la importancia de la correcta especificación de las componentes anteriores, con lo que estaremos en posición de emitir un juicio respecto a la utilidad de las cópulas en nuestro ejercicio.

Recordamos aquí que gracias a los resultados de la sección 4.2.2, tanto las estructuras como las marginales aquí propuestas son adecuadas para modelar nuestros portafolios y de hecho pueden considerarse como las 'correctas' en adelante. En todo caso, de existir un mejor ajuste para la cópula de cada portafolio, los estadísticos garantizan que no habrá diferencias significativas con las que elegimos.

En función de lo anterior realizaremos el contraste para ambos portafolios anticipando que el segundo permitirá una visualización más clara de la influencia de las marginales y el primero sobre la cópula. Para lo anterior nuevamente emplearemos simulación ilustrando el comportamiento de las estimaciones del VaR para diferentes números de ensayos MonteCarlo. Finalizamos con algunas observaciones sobre la infraestructura e inversión de tiempo requerida en ambos casos, así como algunas observaciones interesantes sobre esta aplicación.

#### 4.3.1. Cópulas G-H y Gauss con marginales $t$ no centrales vs distribución Normal Bivariada

Dado que los cálculos a partir del supuesto de normalidad bivariada bajo simulación MonteCarlo siguen el Algoritmo 9 y la estimación de los parámetros de esta distribución es accesible al igual que la simulación de sus observaciones, entraremos directamente con la revisión numérica para identificar las diferencias en los cálculos logrados a partir de ella y los hechos a partir de las cópulas ajustadas para ambos portafolios. Cabe mencionar que emplearemos la información de los cuadros 4.1 y 4.2 que es la que se obtendría sin la teoría de cópulas. Se presenta adelante el cuadro comparativo para el portafolio Televisa-Telmex. Para este portafolio observamos que el VaR 99% se subestima en promedio en 0.17 unidades bajo supuestos de normalidad. En contraste, los cálculos del VaR 95% muestran diferencias más pequeñas que en promedio llegan a 0.08 unidades, es decir, aproximadamente la mitad de lo que se subestima para el otro cuantil. Si bien esto puede parecer extraño, es un resultado lógico si se considera que algunas de las diferencias más importantes entre las marginales ocurren precisamente en las colas de la distribución y también las cópulas tienen diferencias en estos extremos. Comparativamente el VaR 99% es mayor que el de 95%, lo que es congruente a nivel teórico puesto que al tomar un cuantil más grande implícitamente estamos considerando una pérdida mayor.

Para redondear lo anterior en la Figura 4.8 se presentan los gráficos correspondientes a este portafolio. Donde se muestra el VaR para la cópula GH con

Miles de ensayos MC	VaR 95 %			VaR 99 %		
	C. G-H	Normal B.	Dif.	C. G-H	Normal B.	Dif.
1	1.1802	1.0083	0.1719	1.6571	1.4485	0.2086
2	1.2065	1.0344	0.1721	1.7719	1.5167	0.2552
3	1.1979	1.0576	0.1403	1.7244	1.5222	0.2022
4	1.1828	1.0764	0.1064	1.7461	1.5074	0.2387
5	1.1380	1.0859	0.0521	1.6465	1.5609	0.0856
6	1.1701	1.0483	0.1218	1.6506	1.4996	0.151
7	1.1393	1.1032	0.0361	1.6669	1.6122	0.0547
8	1.1637	1.0853	0.0784	1.7330	1.4936	0.2394
9	1.1393	1.0755	0.0638	1.7376	1.5639	0.1737
10	1.1582	1.0724	0.0858	1.7366	1.5916	0.1450
11	1.1417	1.1228	0.0189	1.6904	1.5962	0.0942
12	1.1763	1.1097	0.0666	1.7459	1.5525	0.1934
13	1.1802	1.0933	0.0869	1.7173	1.5497	0.1676
14	1.1331	1.0984	0.0347	1.6900	1.5109	0.1791
15	1.1524	1.0941	0.0583	1.7017	1.5916	0.1101
Promedio	1.1640	1.0777	0.0863	1.7077	1.5412	0.1666

Cuadro 4.9: Portafolio Televisa-Telmex. VaR 99 % y diferencias entre cálculos con GH y marginales t no centrales y distribución Normal Bivariada

marginales t no central ( línea continua) y el obtenido con la Normal Bivariada (línea punteada). Se incluyen también las cotas superior e inferior (líneas punteadas en gris con triángulos arriba y abajo respectivamente) así como la cópula producto (cuadro gris con línea punteada en gris).

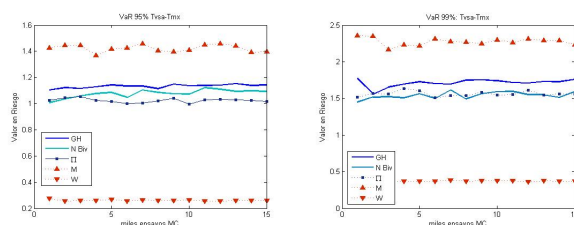


Figura 4.8: VaR para portafolio Televisa-Telmex de 1,000 a 15,000 ensayos Monte Carlo. Izquierda VaR 95%; Derecha: VaR 99%. Distribución Normal Bivariada línea punteada; cópula GH y marginales t no centrales línea continua. Líneas grises punteadas con triángulo Cotas de Frechet-Hoeffding y con cuadro cópula producto

Las líneas con triángulos representan los escenarios extremos del VaR dadas las distribuciones marginales t no centrales estimadas. La cópula  $W$  corresponde al mejor escenario, mientras que  $M$  equivale al peor escenario posible. Para comprender el significado de estas cotas, retomamos los resultados de la sección 1.4.1. En palabras, la cota  $W$  corresponde al caso en que ambos activos no pueden sufrir una pérdida simultáneamente y de hacerlo esto sucede con probabilidad casi nula, mientras que para  $M$ , los activos están propensos a presentar pérdidas o ganancias simultáneamente.

Finalmente, notamos que la serie del VaR de la Normal Bivariada cruza en

varios puntos la de la cópula independencia con una frecuencia mayor para el cuantil 99 %. Esto se debe a que el coeficiente de correlación estimado no es muy grande (ver cuadro 4.2) y como producto de ello el VaR estimado no será muy distinto al obtenido si se asume independencia, lo que nos habla de un error de especificación de la cópula dado que la independencia se rechaza para ambos portafolios de acuerdo con el cuadro 4.3.

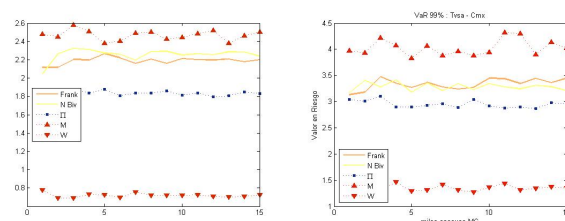


Figura 4.9: VaR para portafolio Televisa-Cemex de 1,000 a 15,000 ensayos Monte Carlo. Izquierda VaR 95 %; Derecha: VaR 99 %. Distribución Normal Bivariada línea punteada; cópula Gauss y marginales t no centrales línea continua. Líneas grises punteadas con triángulo Cotas de Frechet-Hoeffding y con cuadro cópula producto

Realizando una revisión similar para el portafolio Televisa-Cemex, en la Figura 4.9 notamos en primera instancia que las cotas y el cálculo bajo independencia se incrementan a casi el doble de lo observado en el primer portafolio. Como las estructuras de dependencia ( $M$ ,  $W$  y  $II$ ) se mantienen, podemos asegurar que estas variaciones se deben al cambio en una de las marginales, a saber, el cambio del rendimiento de Telmex por el de Cemex, que como se recordará era más volátil que la anterior, lo que en general tendrá una influencia importante en las estimaciones logradas.

Notamos que para el VaR 95 % las diferencias no son significativas entre el VaR bajo supuestos de Normalidad y el que emplea la teoría de cópulas; sin embargo, para el VaR 99 % se hace evidente que el primero subestima el cálculo hecho con la estructura gaussiana y marginales t-no centrales. Lo anterior tiene justificación en los cuantiles de las distribuciones marginales, ya que compartivamente, los de la distribución t empuja más el VaR cuanto más se reduce el error considerado. Lo anterior puede observarse en el cuadro 4.10.

Percentil	Dist. Marginal	Cemex	Telmex	Televisa
95 %	t no central	0.027562	0.022398	0.030887
	Normal	0.028677	0.022965	0.031846
	Dif	-0.0011142	-0.00056675	-0.00095862
99 %	t no central	0.044967	0.03651	0.046825
	Normal	0.040171	0.032526	0.044276
	Dif	0.0047963	0.0039839	0.0025497

Cuadro 4.10: Cuantiles 95 % y 99 % para las distribuciones marginales estimadas

Por último, hacemos un examen sobre las ventajas y las desventajas del cálculo del VaR con la teoría de cópulas. Sabemos por una parte que el cálculo

a partir del supuesto de normalidad será mucho más parsimonioso que el que emplea la cópula; no requiere de una revisión tan extensa dentro de la parte de la estimación y ésta, al igual que la simulación, puede realizarse con cualquier paquete estadístico sin necesidad de desarrollar la infraestructura que requiere el manejo de cópulas.

En contraste, el cálculo logrado a partir del empleo de cópulas es más preciso al permitir la incorporación de marginales y estructura de dependencia más adecuadas. Si bien la cópula gaussiana puede ser adecuada, esto puede ser mostrado con suficiente claridad a partir de la teoría de cópulas como se hizo para el segundo portafolio y dotando al modelo de las marginales convenientes. En cuanto a éstas últimas, difícilmente seguirán un comportamiento normal lo cual limita bastante el alcance del cálculo logrado a partir del supuesto de normalidad.

En conclusión, el empleo de la teoría de cópulas supone un mayor costo en cuanto implementación y análisis, que es premiado por una mayor exactitud en la estimación del VaR.

## 4.4. Otras aplicaciones de interés actuarial

En esta última sección presentamos otras aplicaciones que tienen por objeto ilustrar la utilidad de las cópulas en otros campos de interés actuarial. Las cuatro primeras tienen relación con las áreas de Seguros y Finanzas mientras que las últimas dos tienen un sentido teórico más general orientado al área Estadística.

### 4.4.1. Indicadores de riesgo

Realizamos a continuación algunas importantes observaciones que pretenden mostrar la utilidad de la teoría de cópulas en la cuantificación del riesgo. En primer término, a partir del VaR<sup>3</sup> y sus cotas en términos de la estructura de dependencia existente, veíamos que  $W$  representa la estructura bajo la cual el VaR alcanza sus valores más pequeños y de hecho nos permite establecer una regla para minimizar el valor en riesgo: se preferirá un portafolio cuya estructura de dependencia esté más cercana a la cota inferior  $M$  que a la  $W$  -asumiendo que posean activos similares en cuanto a volatilidad se refiere- puesto que esta última corresponde al caso en que a pérdidas grandes de un activo corresponden pérdidas grandes en el otro mientras que  $M$  representa el caso en que no pueden suceder pérdidas grandes simultáneamente.

Para extrapolar la observación anterior -al menos de forma intuitiva- consideremos a  $\xi$  como una medida del riesgo de un portafolio integrado por dos activos  $(X, Y)$  y que considera para su conformación tanto el comportamiento conjunto como el individual de las pérdidas, tenemos que:

- $\xi$  varía de acuerdo al comportamiento conjunto descrito por  $H_{X,Y}$
- $\xi$  tenderá a incrementar si el comportamiento conjunto indica que habrá pérdidas simultáneas

---

<sup>3</sup>Cuyo empleo debe ser cuidadoso de acuerdo con los requerimientos de esta herramienta y los resultados esperados a partir del empleo de cópulas.

- $\xi$  tenderá a disminuir si el comportamiento conjunto indica que no es muy factible que halla pérdidas simultáneas.

Para dos funciones  $H_1$  y  $H_2$  conjuntas asociadas al portafolio de activos  $(X, Y)$  tales que  $H_1(x, y) \geq H_2(x, y)$  a partir de valores  $x \geq x_0$  y  $y \geq y_0$ <sup>4</sup>, entonces  $\xi$  guarda la siguiente relación:

$$\xi_{H_1}(X, Y) \leq \xi_{H_2}(X, Y)$$

Además, para una medida con estas características y una función de distribución  $H$  tendremos :

$$\xi_W(X, Y) \leq \xi_H(X, Y) \leq \xi_M(X, Y)$$

Resumiendo, la siguiente afirmación es factible:

*Bajo una medida de riesgo que considere la distribución conjunta del portafolio, éste será menos riesgoso si tiene una estructura de dependencia más cercana a  $W$  que si se aproxima a  $M$ .*

partir de ello podemos inducir un indicador de riesgo mediante el estadístico  $Cn$  presentado en el capítulo 2, que como veíamos puede determinar la distancia de la cópula asociada a nuestro portafolio y la cota superior o bien la cota inferior de la manera siguiente:

$$Tn_W = \sup|Cn - W| \quad (4.2)$$

o bien,

$$Tn_M = \sup|Cn - M| \quad (4.3)$$

donde estaremos interesados en obtener la distancia mayor para asumir menos riesgo en el primer caso y una distancia que se aproxime a cero para el segundo.

Notamos que los indicadores tendrán la ventaja de ser acotados y adimensionales al no incorporar el comportamiento marginal de los activos. Aunque estos no cumplen con los axiomas de coherencia de una medida de riesgo, son bastante informativos al brindar un punto de comparación para el riesgo atribuible a la estructura de dependencia en varios portafolios. En conclusión, los indicadores anteriores son un resumen del riesgo conjunto en un portafolio en términos de la probabilidad de obtener pérdidas simultáneas. Para el caso bivariado los valores extremos de estos quedan como:

Notamos que si empleamos los indicadores anteriores en un portafolio que incluya el rendimiento de un activo  $X$  y el rendimiento  $R_M$  del mercado también estaremos obteniendo un indicador de pérdidas conjuntas, sólo que en este caso representará la relación existente con el mercado.

<sup>4</sup>Esta condición nos dice que la función  $H_1$  acumula su probabilidad antes de  $x_0$  y  $y_0$ , lo que representa menor probabilidad de sucesos extremos y en las colas de la distribución. Por el contrario  $H_2$  al ser menor que la otra función a partir de estos puntos, indica que tarda un poco más en acumular su probabilidad y por ende, que los sucesos extremos son más probables.

Probabilidad de pérdidas simultáneas	Indicador	
	$Tn_W$	$Tn_M$
c.s. = 0	2	0
c.s. = 1	0	2

Cuadro 4.11: Indicadores de pérdidas conjuntas

En estos términos, existe una relación estrecha con la  $\beta$  del modelo CAPM <sup>5</sup> en la ecuación:

$$R = R_f + \beta[E(R_m - R_f)]$$

donde  $\beta = \frac{cov(R, R_m)}{\sigma_m^2}$  representa el riesgo sistemático o no diversificable y es en términos generales la cantidad 'lineal' de riesgo existente entre el activo o portafolio en cuestión y el mercado.

Como observamos, los indicadores propuestos en 4.3 y 4.4 no tienen una interpretación en términos del 'premio al riesgo' obtenido como en el caso del CAPM, sin embargo tienen la ventaja de no limitar su alcance a la búsqueda de una relación lineal.

Finalmente, solo mencionaremos que un estudio mediante simulación puede revelar valores críticos para los indicadores anteriores en términos del grado de 'pérdida conjunta' que se esté dispuesto a tomar una vez que se seleccionen las distribuciones marginales de los rendimientos de cada activo en el portafolio.

#### 4.4.2. Reaseguro

Revisemos ahora otra aplicación de interés general en el ámbito actuarial. Dada su importancia, damos primero un breve esbozo de sus características. En términos simples, el reaseguro es un mecanismo de transferencia de riesgo mediante el cual una compañía reaseguradora acepta indemnizar a una aseguradora por una parte, o inclusive, por la totalidad de pérdidas que esta última afronte. Por lo anterior, el reaseguro es una de las herramientas más efectivas para la administración del riesgo en una aseguradora. Algunos de los beneficios obtenidos son:

- Protección de la solvencia de la aseguradora al reducir la probabilidad de ruina.
- Incremento de la capacidad para asumir otros riesgos por parte de la aseguradora e incluso de la entrada a nuevos mercados.
- Mayor estabilidad en las ganancias al transferir los elementos más riesgosos
- Obtención de asesoría para la administración del riesgo por parte de la reaseguradora que en general posee una experiencia mayor.
- Bajo costo del contrato al tener las reaseguradoras bajos gastos administrativos.

<sup>5</sup>Por sus siglas en inglés, Capital Asset Pricing Model, se puede consultar [12] para una descripción de este.

Dicho esto, podemos identificar distintos tipos de reaseguro, a continuación presentamos algunos de ellos:

- *Stop-loss*: creado para controlar la magnitud de la pérdida. Una vez fijados los límites de la responsabilidad y el nivel de retención para un portafolio, la reaseguradora acuerda cubrir las reclamaciones agregadas que excedan el nivel de retención hasta el límite de la póliza haciéndose cargo la aseguradora de la parte restante.
- *Quota-share*: Como su nombre lo indica, la reaseguradora se encargará de cubrir un porcentaje de cada una de las pérdidas. En términos generales, este porcentaje puede no ser fijo y estar en función del monto reclamación.
- *Excess-of-loss*: similar al stop-loss con la diferencia de que este se aplica sobre cada una de las reclamaciones y no sobre las reclamaciones agregadas.

Además de los anteriores, existe otro tipo de reaseguro que contempla también los gastos extraordinarios a los que se enfrentan las aseguradoras una vez hecha una reclamación, por ejemplo: honorarios por disputas legales, gastos de investigación, etcétera, que en general no son independientes de la reclamación.

Los reaseguros que incluyen cobertura por los gastos extraordinarios también toman formas similares a las coberturas expuestas. A continuación estudiamos el caso en que la cobertura sobre las reclamaciones es de tipo excess-of-loss mientras que la cobertura sobre gastos es similar a la del quota-share; a este se le denomina 'pro-rata sharing of expenses'.

Como mencionamos, la mayoría de las veces no podremos afirmar la existencia de independencia entre las reclamaciones y los gastos, lo que origina un problema para la estimación de la prima de reaseguro exacta al requerirse la función de distribución conjunta de estas.

Revisemos primero cómo se obtienen los gastos en los que incurriría una reaseguradora bajo el esquema anterior, para lo cual descomponemos a este en la parte correspondiente a la reclamación y la de los gastos. Para ello, asumamos que  $X$  y  $Y$  son las variables aleatorias que denotan el monto de la reclamación y los gastos respectivamente, además, que el límite de retención de la aseguradora es  $L$  mientras que el límite de responsabilidad es  $R$ . Dicho esto, bajo el esquema de excess-of-loss el pago de la reaseguradora corresponde a:

$$\begin{cases} 0 & X < R \\ X - R & R \leq X < L \\ L - R & L \leq X \end{cases}$$

Ahora, para los gastos, asumiendo que la reaseguradora se hace cargo de una proporción variable de acuerdo al tamaño de la reclamación y en general los gastos serán menor a esta ( $X > Y$ ) tenemos:

$$\begin{cases} (X - R)Y/X & R \leq X < L \\ (L - R)Y/L & L \leq X \end{cases}$$



De los razonamientos anteriores, sumando los pagos en los que incurriría la aseguradora obtenemos la función:

$$g(X, Y) = \begin{cases} 0 & X < R \\ X - R(1 + Y/X) & R < L \\ L - R(1 + Y/L) & L \leq X \end{cases}$$

Para obtener el pago esperado obtenemos la esperanza de esta función bajo la distribución conjunta de las variables aleatorias  $H$ ; es decir:

$$E[g(X, Y)] = \int_0^{\infty} g(x, y)h(x, y) dx dy$$

Sabemos que las reclamaciones y los gastos en general pueden modelarse a partir de las distribuciones Pareto y lognormal entre otras. Ahora, para conocer la distribución conjunta, podemos ajustar una cópula a  $X$  y  $Y$  y posteriormente -siguiendo un procedimiento similar para la simulación de los retornos en las secciones anteriores- obtener el valor esperado.

Para ilustrar de una manera más clara esto, debemos seguir los siguientes pasos:

1. Ajustar una cópula adecuada a  $X$  y  $Y$  y ajustar las distribuciones marginales.
2. Simular un número grande  $N$  de observaciones de  $X$  y  $Y$  bajo la distribución conjunta obtenida en el punto anterior.
3. Aplicar la función  $g$  a estas observaciones y almacenar estos valores en  $p_i$ ,  $i = 1, 2, \dots, N$ .
4. Obtener la media muestral de las  $p_i$  que de hecho corresponde al valor esperado buscado.

#### 4.4.3. Seguro de vida conjunta

En Matemáticas Actariales, un supuesto usual cuando se aborda el tema de status de vida múltiple, es la independencia dentro de los tiempos de vida restantes de cada individuo que integra el status.

Intuitivamente sabemos que esto no es necesariamente cierto. Suponiendo un status conformado por un matrimonio, sabemos que la mortalidad de uno no es completamente ajena a la del otro. Por una parte, ambos se encuentran expuestos a riesgos individuales e independientes como son el lugar de trabajo y las enfermedades hereditarias entre otros. Por otra, también se enfrentan a riesgos en común como son el lugar en que residen y su forma de vida entre otros.

En términos estadísticos, Frees y Valdez <sup>6</sup> muestra la existencia de dependencia positiva en un status como el anterior para experiencia norteamericana y Youn

<sup>6</sup>Frees, W., Carriere, J., Valdez, E. (1995) *Annuity Valuation with Dependent Mortality*. Actuarial Research Clearing House.

y Shemyakin <sup>7</sup> para experiencia canadiense. Ambos textos muestran la necesidad de incorporar modelos multivariados para el cálculo de probabilidades de supervivencia y particularmente optan por las cópulas como herramienta fundamental para ello.

Retomando las expresiones básicas para los modelos de vida conjunta y de último superviviente de dos individuos con edades  $x$  y  $y$ :

1.  ${}_tP_x$  denota la probabilidad de que el individuo con edad  $x$  llegue con vida al tiempo  $t$ .
2.  ${}_tq_x$  denota la probabilidad de que el individuo con edad  $x$  fallezca antes de que transcurran  $t$  años.
3.  ${}_tP_{xy}$  denota la probabilidad de que el status de vida conjunta llegue con vida al tiempo  $t$ , es decir, que ninguno de los integrantes fallezca antes de  $t$  años.
4.  $\overline{{}_sP_{xy}}$  denota la probabilidad de que el status de último superviviente llegue con vida a  $s$ , es decir, que al menos uno de los integrantes llegue con vida a  $t$ .

Cuando se asume independencia, es posible reexpresar las probabilidades de los status en términos de las probabilidades individuales como:

$${}_tP_{xy} = {}_tP_x \cdot {}_tP_y$$

y

$$\overline{{}_sP_{xy}} = 1 - {}_s q_x \cdot {}_s q_y$$

y además se da la siguiente relación:

$$\overline{{}_tP_{xy}} + {}_tP_{xy} = {}_tP_x + {}_tP_y$$

lo que en general es bastante práctico para la realización de cálculos al ser en general inaccesible una expresión directa para la probabilidad conjunta.

Si consideramos a dos individuos recién nacidos y representamos las edades en las que fallecen por las variables aleatorias  $X$  y  $Y$ , podemos escribir las funciones de supervivencia conjunta e individuales como:

$$S(x, y) = 1 - P[X > x, Y > y]$$

y

$$S_x(x) = 1 - P[X > x] \text{ y } S_y(y) = 1 - P[Y > y]$$

y siguiendo con los supuesto de independencia obtenemos:

$${}_tP_{xy} = \frac{S(x+t, y+t)}{S(x, y)} = \frac{S_x(x+t)S_y(y+t)}{S_x(x)S_y(y)}$$

y también

$$\overline{{}_tP_{xy}} = \frac{S(x, y+t) + S(x+t, y) - S(x+t, y+t)}{S(x, y)}$$

---

<sup>7</sup>Youn, H., Shemyakin, A. (1999) *Statistical aspects of joint life insurance pricing*. Proceedings of American Statistical Association. pp34-38.

$$= 1 - \frac{[S_x(x) - S_x(x+t)][S_y(y) - S_y(y+t)]}{S_x(x)S_y(y)}$$

que nuevamente está en términos de las distribuciones marginales de probabilidad.

Si en lugar de lo anterior asumimos que  $X$  y  $Y$  están asociadas por una cópula  $C$ , por el teorema de Sklar podemos escribir a la función de supervivencia conjunta como:

$$S(x, y) = C[S_x(x), S_y(y)]$$

y con ello podemos obtener:

$${}_tP_{xy} = \frac{C[S_x(x+t), S_y(y+t)]}{C[S_x(x), S_y(y)]}$$

y

$$\overline{{}_tP_{xy}} = \frac{C[S_x(x), S_y(y+t)] + C[S_x(x+t), S_y(y)] - C[S_x(x+t), S_y(y+t)]}{C[S_x(x), S_y(y)]}$$

que además de estar en términos de las funciones de supervivencia individuales, incorpora la estructura de dependencia adecuada.

Como aplicación directa de lo anterior, empleando las cotas de Frechet-Hoeffding, podemos afirmar que:

$$\max[S_x(x) + S_y(y) - 1, 0] \leq {}_tP_{xy} \leq \frac{\min[S_x(x+t), S_y(y+t)]}{\min[S_x(x), S_y(y)]}$$

para un status de vida conjunta y una expresión un tanto más extensa para el status de último superviviente.

Con lo anterior es posible obtener también las cotas para anualidades y seguros relacionadas con los status anteriores de una manera bastante rápida.

Ahora, si se realiza el ajuste sobre una tabla de mortalidad divariada, podemos extraer la cópula correspondiente y a partir de ello obtener un cálculo más preciso y mejor fundamentado del que supone independencia. Bajo experiencia canadiense y norteamericana, los ajustes para un seguro de vida conjunta de un matrimonio corresponden a las familias Gumbel-Hougaard y Frank respectivamente y en ambos casos se presenta dependencia de cuadrante positivo.

Para cerrar con esta aplicación, cabe mencionar que aún no existe un estudio para México debido principalmente a la relativamente nueva implementación de estas herramientas y a la falta de estadísticas de mortalidad conjunta que permitan su realización.

#### 4.4.4. Valuación de Opciones

El empleo de los derivados financieros ha cobrado en los últimos tiempos una gran importancia. Sus características los hacen instrumentos muy efectivos para la administración de riesgos financieros y su flexibilidad, en conjunto con las crecientes necesidades empresariales, ha impulsado la creación de nuevos instrumentos de este tipo. Desafortunadamente, dada la flexibilidad de estos, su complejidad en términos de valuación, es bastante grande y su precio es en

general un problema. En México, si bien el mercado de derivados está aún en desarrollo, se aprecia ya un interés por su empleo abriendo campos de investigación para nuevas técnicas de valuación. Si bien las técnicas existentes (modelo de Black & Scholes por ejemplo) resuelven estos problemas, nuevamente reposan sobre supuestos fuertes sobre la estructura de dependencia existente entre las variables de interés.

En general, el precio de uno de estos instrumentos se verá influenciado principalmente por el o los activos contemplados en el contrato, la fecha de vencimiento y el tipo de interés. Por ejemplo, para las opciones sobre activos múltiples ('rainbow options') que incorporan la alternativa de comprar (o vender) el mejor (o peor) de los activos en un portafolio, dependerá directamente de cada uno de los activos subyacentes y desde luego si se trata de una opción de compra o de venta, además del vencimiento del contrato y el interés vigente. Como vemos, las características de la relación de dependencia existente entre los activos contemplados jugará un papel importante para la valuación.<sup>8</sup>

En términos generales, el precio de una opción Europea con vencimiento en  $T$  y activos subyacentes  $X_1$  y  $X_2$  es:

$$c = B(T - t)E_t[P(X_1, X_2, t)]$$

donde  $B(T - t)$  es el valor en  $t$  de un cupón cero con vencimiento en  $T$  y  $E_t(\cdot)$  es el valor esperado en  $t$  calculado bajo una medida neutral al riesgo y  $P(X_1, X_2, t)$  es la función de pago para estos subyacentes.

De acuerdo con los resultados de Black-Scholes-Merton, bajo supuestos de log-normalidad en los activos y de distribución normal conjunta es posible llegar a una solución satisfactoria una vez que se da a conocer la función de pago y se realiza la estimación sobre los parámetros marginales. Una desventaja de lo anterior es precisamente el supuesto de normalidad y de acuerdo con ello, de que la correlación es una medida de dependencia suficiente para el comportamiento de los activos.

A partir del empleo de las cópulas es posible ir más allá de los supuestos mencionados para este tipo de instrumentos. Por una parte, se evita el supuesto sobre la distribución conjunta y por otra, se da una mayor holgura sobre la gama de distribuciones marginales a considerar.

Para ejemplificar esto, consideremos el call de una opción del mínimo de dos activos. Por una parte, asumamos que los precios de los activos  $X_1$  y  $X_2$  al tiempo  $t$  están dados por  $S_1(t)$  y  $S_2(t)$  respectivamente y  $K$  al que se comprará el activo. La función de pago para esta opción, queda dada por:

$$P(X_1, X_2, T) = \max[\min(S_1(T), S_2(T)) - K, 0]$$

y asumiendo que las tasa de interés es fija, bajo la primer identidad de esta sección, llegamos a:

$$c = B(T - t) \left\{ \int_K^\infty qg(q) dq - K(1 - G(K)) \right\}$$

<sup>8</sup>Para una revisión de los conceptos relacionados con valuación de opciones se recomienda la referencia [12]

donde  $g(\cdot)$  es la densidad neutral al riesgo del mínimo y  $G$  su distribución acumulada. Además para la distribución neutral al riesgo y del mínimo observamos:

$$\begin{aligned} G(q) &= P[S_1 \leq q \wedge S_2 \leq q] \\ &= P[S_1 \leq q] + P[S_2 \leq q] - P[S_1 \leq q \wedge S_2 \leq q] \\ &= F_{S_1}(q) + F_{S_2}(q) + H(q, q) \end{aligned}$$

asumiendo que  $F_{S_1}$ ,  $F_{S_2}$  y  $H$  son las distribuciones marginales y conjunta respectivamente.

Con lo anterior, si  $C$  es la cópula de  $S_1$  y  $S_2$  tenemos:

$$G(q) = F_{S_1}(q) + F_{S_2}(q) + C[F_{S_1}(q), F_{S_2}(q)]$$

y la densidad  $g$  puede obtenerse derivando la expresión anterior. A partir del ajuste de una cópula y de las marginales adecuadas sobre las observaciones de los activos es posible proveer un cálculo sin supuestos de lognormalidad. Además, este mismo esquema puede ser empleado para valorar opciones sobre el máximo y sobre 'spread' entre otras, y también para la evaluación bajo la posición de venta (put).

La incorporación de las cópulas como herramientas para esta valuación es reciente y hay aún mucho camino por recorrer. En Cherubini & Luciano <sup>9</sup> se examinan otras características de las cópulas relacionadas con la valuación de opciones y se justifica el hecho de que  $C$  es una medida neutral al riesgo y cumple con las restricciones de arbitraje. En van den Goorberg <sup>10</sup> se estudia el empleo de cópulas dinámicas (con parámetro dependiente del tiempo) para la valuación.

#### 4.4.5. Regresión

Como se mencionó muy al principio de esta presentación, la regresión es una de las herramientas más importantes para la descripción de la dependencia entre dos o más variables y que además se caracteriza por ser bastante simple en su comprensión. En esta sección mostramos cómo pueden aprovecharse las cópulas para construir regresiones no lineales.

Si poseemos una muestra divariada con  $n$  componentes, para obtener una regresión lineal basta estimar los coeficientes por mínimos cuadrados o máxima verosimilitud y con ello describiremos el *comportamiento medio* de la muestra asumiendo que una de las variables es la descriptiva y la otra, dependiente.

Bajo este planteamiento, si asumimos ahora que las variables tienen funciones de distribución marginales  $F$  y  $G$  respectivamente y distribución conjunta  $H$  en términos conceptuales estaremos obteniendo la curva:

$$y = E(Y | X) = \int_0^\infty [1 - H(y | x)] dy + \int_0^\infty H(y | x) dy$$

<sup>9</sup>Cherubini, U. , Luciano, E. *Multivariate Option Pricing* International Centre for Economic Research. Applied Mathematics. Working Paper 5- 2002

<sup>10</sup>van den Goorbergh, R. Genest, C., Werker, B. *Bivariate option pricing using dynamic copula models* Insurance: Mathematics and Economics 37 (2005) 101-114

Asumiendo que  $X$  y  $Y$  tienen cópula  $C$  podremos encontrar una variedad de regresiones para ellas en términos de la distribución condicional ilustrada en la sección 1.4.4.

Por ejemplo, si las marginales son uniformes y su cópula es de la familia Frank con parámetro  $\theta$ , la curva de regresión de  $x$  sobre  $y$  es

$$y = E(Y | X) = \frac{(1 - e^{-\theta})xe^{-\theta x} + e^{-\theta}(e^{-\theta x} - 1)}{(e^{-\theta x} - 1)(e^{-\theta} - e^{-\theta x})}$$

La utilidad del planteamiento anterior desde luego se centra en la obtención de una regresión no necesariamente lineal íntimamente ligada con la estructura de dependencia existente entre las variables y para la cuál el único parámetro a estimar es  $\theta$ , el parámetro de la cópula.

Hagamos ahora un replanteamiento para la curva de regresión para los cuantiles de la distribución en vez del valor esperado. Para ello observemos que:

$$P[Y \leq y | X = x] = P[V \leq G(y) | U = F(x)] = \frac{\partial}{\partial u} C(u, v)$$

A partir de lo anterior, si tomamos el cuantil  $\xi$  y resolvemos para  $v$  la siguiente expresión:

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

obtendremos la regresión de  $u$  sobre  $v$ .

Para ilustrar el significado de esta última regresión, si consideramos que  $X$  y  $Y$  determinan las pérdidas y los gastos en una compañía, podremos encontrar para una pérdida específica  $x$  el percentil correspondiente de las pérdidas.<sup>11</sup>

Para finalizar esta sección, realizamos una observación importante sobre la relación de las curvas de regresión presentadas con la estimación de parámetros y ajuste de una cópula. Retomando el método de máxima verosimilitud canónica en la sección 3.1.1, la estimación del parámetro  $\theta$  se realiza de acuerdo a la maximización de

$$\sum_{i=1}^n \log C_{12}(U_i, V_i)$$

con  $U$  y  $V$  presentados en (3.3); para la obtención del parámetro en la regresión mediante máxima verosimilitud consideraríamos a  $x_i$  distribuídas como  $\frac{\partial}{\partial u} C(u, v)$  y al obtener la función de verosimilitud llegaríamos a algo similar.

No obstante, queda la interrogante de lo que sucedería si se realiza una estimación por mínimos cuadrados -que como es bien sabido queda en los mismos términos para la regresión lineal- y comparar:

$$\theta_{MC} = \operatorname{argmin} \sum_i^n [v_i - \frac{\partial}{\partial u} C(u_i, v_i)]^2$$

con  $\hat{\theta}_{MVC}$  expresado en la ecuación (3.2).

<sup>11</sup>Notamos que la regresión lineal toma la media como punto de partida.

#### 4.4.6. Procesos de Markov

A continuación mostramos una interesante aplicación de las cópulas para la generación de nuevos procesos de Markov. Realizaremos primero una revisión básica de este tipo de procesos para después vincular a las cópulas en su estudio.

La estructura de dependencia de un proceso de Markov es muy particular en comparación a la existente entre dos variables aleatorias cualesquiera. Decimos que un proceso de Markov es una secuencia de variables aleatorias  $\{X_t \mid t \in T\}$  -donde  $X_t$  puede verse como el estado de la variable  $X$  en el tiempo  $t$ - que cumple la siguiente característica para  $t_1 < t_2 < \dots < t_n < t_s$ :

$$P[X_t \leq x \mid X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_n} = x_n, X_{t_s} = x_s] = P[X_t \leq x \mid X_{t_s} = x_s]$$

es decir, en el futuro del proceso solo interesa su pasado inmediato. En adelante usaremos la siguiente notación para la probabilidad condicional:

$$P(x, s; y, t) = [X_t \leq y \mid X_s = x]$$

Si tomamos a  $F_s$  y  $F_t$  como las distribuciones de  $X_s$  y  $X_t$ , y además hacemos a  $H_{st}$  su distribución conjunta, es claro que por el teorema de Sklar podemos obtener

$$H_{st}(x, y) = C_{st}[F_s(x), F_t(y)]$$

y a partir de ello extraer la cópula del proceso.

Un ejemplo interesante de esto, es la obtención de la cópula del proceso Browniano. Sabemos que para este proceso, la probabilidad de transición está dada por  $\Phi[(t-s)^{-2}(y-x)]$ . Si asumimos que  $X_0 = 0$  y retomamos la distribución condicional para cópulas, podemos observar que para un proceso Browniano:

$$\begin{aligned} P(x, s; y, t) &= \Phi[(t-s)^{-2}(y-x)] \\ &= \frac{\partial}{\partial F_s} C_{st}[F_s(x), F_t(y)] \end{aligned}$$

de acuerdo con esto,

$$C_{st}[F_s(x), F_t(y)] = \int_{-\infty}^x \Phi[(t-s)^{-2}(y-z)] dF_s(w)$$

Por nuestra condición de valor inicial sabemos que  $F_s(z) = \Phi(z/\sqrt{t})$ . Ahora, haciendo un primer cambio de variable  $F_s(z) = w$ , observando que  $z = -\infty \Rightarrow w = 0$ , que  $z = x \Rightarrow w = F_s(x)$ ,  $dF_s(z) = dw$  obtenemos:

$$C_{st}[F_s(x), F_t(y)] = \int_0^{F_s(x)} \Phi[(t-s)^{-2}(y - \sqrt{s}\Phi^{-1}(w))] dw$$

Haciendo otro cambio de variables  $F_s(x) = u$  y  $F_t(y) = v$  llegamos a la cópula del proceso:

$$C_{st}(u, v) = \int_0^u \Phi\{(t-s)^{-2}[\sqrt{t}\Phi^{-1}(v) - (-\sqrt{s}\Phi^{-1}(w))]\} dw$$

A partir de este resultado podemos construir nuevos procesos de Markov con marginales no normales, bastando para ello evaluar las marginales deseadas en la cópula anterior y la obtención de la distribución condicional.

Ahora, ¿es posible construir un proceso de Markov a partir de una cópula cualquiera?. Para resolver esta interrogante definiremos a continuación una operación binaria entre cópulas que resulta ser útil para discriminar aquellas que satisfacen la ecuación de Chapman-Kolmogorov, condición necesaria para la construcción de este tipo de procesos.

Se define el producto  $*$  de las cópulas  $C_1$  y  $C_2$  como:

$$(C_1 * C_2)(u, v) = \int_0^\infty \frac{\partial}{\partial u} C_1(u, t) \frac{\partial}{\partial v} C_2(t, v) dt$$

Puede probarse que  $C_1 * C_2$  es una cópula, que  $\Pi$  y  $M$  son los elementos nulo y neutro respectivamente y además que el operador anterior es asociativo y no conmutativo.

Independientemente de ello, podemos afirmar que si

$$(C_{st} * C_{su}) = C_{ut}$$

entonces  $C_{st}$  satisface las ecuación de Chapman-Kolmogorov.

En Muriel<sup>12</sup> se expone una condición necesaria y suficiente a partir de otro operador binario entre cópulas, lo que provee de una potente herramienta para la conformación de distintos procesos de Markov con probabilidades de transición y distribuciones marginales arbitrarias.

---

<sup>12</sup>Muriel, Op Cit pp78



# Conclusiones

En este trabajo se han presentado tanto elementos teóricos como prácticos que exhiben las características de las cópulas.

Primeramente, las cópulas bivariadas son funciones bien comportadas: están acotadas y definidas en el cubo unitario, son continuas, no decrecientes en cada uno de sus argumentos e invariantes a transformaciones no decrecientes. El Teorema de Sklar, resultado fundamental para el uso de las cópulas, garantiza la existencia de una cópula que relaciona de manera funcional dos distribuciones marginales y una distribución conjunta.

En términos probabilísticos, las cópulas representan la estructura de dependencia que vincula los comportamientos marginales en una distribución conjunta; por ello, los modos de dependencia, las medidas de asociación y dependencia tienen una caracterización a partir de ellas. En contraste, el coeficiente de correlación de Pearson no puede expresarse en términos de cópulas al estar ligado no solo a la estructura de dependencia sino también al comportamiento de las marginales de la distribución.

La cópula empírica tiene una construcción muy simple y es el equivalente a la distribución empíricas para una variable. Gracias a ella, es factible la construcción de estadísticos que emulan el funcionamiento de la prueba de Kolmogorov-Smirnov en el caso de una variable, para la elaboración de pruebas de independencia y de bondad de ajuste.

Dentro de la parte del ajuste de cópulas, se mostró que tanto las pruebas de bondad de ajuste como la estimación de parámetros es factible para cópulas que tienen segundas derivadas o cuyos parámetros pudieran expresarse en términos de las medidas de asociación de Kendall o de Spearman, además de los métodos ya conocidos para cópulas arquimedeanas.

Como ejemplo del empleo de las cópulas, se hizo el estudio del valor en riesgo de dos portafolios compuestos cada uno por un par de acciones que cotizan en la Bolsa Mexicana de Valores, con un instrumento en común (Televisa CPO) y las acciones de Cemex y Telmex. A partir de los estadísticos de prueba de independencia del segundo capítulo, se confirmó la inexistencia de esta para niveles de confianza de 95 y 99 %. Para el modelado de las marginales, se optó por el empleo de distribuciones  $t$  no centrales en todos los casos y como estructuras de dependencia a las cópulas Gumbel-Hougaard y Gauss respectivamente. Una vez hecha la simulación sobre las distribuciones estimadas, se observó que en el primer portafolio, el VaR se subestimaba en poco más de 5 % para el cuantil del 95 % y de casi 6.5 % para el de 99 % con 15 mil ensayos MonteCarlo cuando se hacía supuesto de normalidad bivariada. Para el segundo caso, el error era menor, dada la correcta especificación de la cópula a emplear. Se bosquejaron además otras aplicaciones, a partir de lo cual se concluye que es posible emplear

a las cópulas en una variedad de temas de interés actuarial.

Como conclusión general, las cópulas representan un gran avance para el modelado del comportamiento conjunto al permitir la incorporación de distribuciones marginales adecuadas y también -como en el caso del primer portafolio- estructuras de dependencia mejor ajustadas a la muestra. En todo caso, un estudio mediante cópulas permite la validación de los supuestos empleados sobre la forma de dependencia.

En cuanto a la medición de la dependencia, vimos que los estadísticos mostrados en la segunda sección son en buena medida un avance para ello y permiten elaborar pruebas sobre la muestra sin importar el comportamiento marginal de esta.

Como aspecto negativo, se hizo notar que el empleo de las cópulas es más costoso en cuanto a la infraestructura que se requiere. En este respecto, es notable la necesidad de generar algoritmos o lograr implementaciones más eficientes para la estimación de parámetros, simulación de observaciones y generación de valores críticos, ya que para la aplicación a gran escala de esta teoría se vuelve indispensable una justa relación sobre el tiempo invertido.

Las últimas observaciones realizadas en la parte teórica tienen que ver con el problema que representa la especificación de las cópulas a emplear, ya que, si bien los estadísticos proveen una herramienta de contraste, no son garantía de que la cópula elegida sea la mejor alternativa existente a pesar de ser lo suficientemente fuertes para discriminar con cierto nivel de confianza aquellos modelos que definitivamente no son útiles.

Para finalizar, como futuros proyectos, se deja el desarrollo de las aplicaciones aquí bosquejadas para probar la efectividad de las cópulas en estos casos, el tratamiento del caso multivariado desde un enfoque actuarial y el perfeccionamiento de los estadísticos  $T_n$  -ahora talvez en términos de distancias promedio y no distancias máximas- así como el ya mencionado perfeccionamiento de las aplicaciones computacionales presentadas en el apéndice. Finalmente, se concluye que las cópulas, aún en este acercamiento muy particular, tienen propiedades que pueden y deben explotarse por los actuarios ya que como se mostró, poseen características que las hacen superiores a otras técnicas de modelado de dependencia.

## Apéndice A

# Modos de dependencia y $\tau$ de Kendall

Familia	$C_\theta(u, v)$	$\theta \in$	Límites	Modo de dependencia
Cuadras-Augé	$[\min(u, v)]^\theta [uv]^{1-\theta}$	$[0, 1]$	$C_0 = \Pi, C_1 = M$	DCP
Farlie-Gumbel-Morgensten	$uv + \theta uv(1-u)(1-v)$	$[-1, 1]$	$C_0 = \Pi$	$\theta > 0$ , DCP $\theta < 0$ DCN
Fréchet	$(1 - \theta)\Pi + \theta W$	$[0, 1]$	$C_0 = \Pi, C_1 = W$	

Cuadro A.1: Cópulas no arquimedeanas

Familia	$\theta \in$	Límites	Modo de dependencia
Ali-Mikhail-Haq	$(0, \infty)$	$C_0 = \Pi$	DCP
Clayton	$[-1, \infty) \setminus \{0\}$	$C_{-1} = W, C_0 = \Pi,$ $C_\infty = M$	$\theta > 0$ DCP $\theta < 0$ DCN
Gumbel-Hougaard	$[1, \infty]$	$C_0 = \Pi, C_1 = M$	DCP
Frank	$(-\infty, \infty) \setminus \{0\}$	$C_{-\infty} = W, C_0 = \Pi$ $C_\infty = M$	DCP DCN
Gumbel-Barnett	$(0, 1]$	$C_0 = \Pi$	DC

Cuadro A.2: Cópulas arquimedeanas

Familia	$\tau$	$\theta$	$\tau \in$	$\lambda(t)$
Cuadras-Augé	$\frac{\theta}{2-\theta}$	$\frac{2\tau}{1-\tau}$	$[0,1]$	-
Farlie-Gumbel -Morgensten	$\frac{2\theta}{9}$	$\frac{9\tau}{2}$	$[-\frac{2}{9}, \frac{2}{9}]$	-
Ali-Milkhail-Haq	$-\frac{1}{3\theta^2}[-3\theta^2 + \log(1-\theta)]$ $\times [\theta(2\theta - 4) + 2]$	-	$[-0.1817, \frac{1}{3}]$	$\frac{\log((1-\theta)(1-t))}{t(\theta-1)}$ $\times \frac{1-\theta(1-t)}{t(\theta-1)}$
Clayton	$\frac{\theta}{\theta+2}$	$\frac{2\tau}{1-\tau}$	$[-1,1]$	$\frac{t}{\theta}(1-t^\theta)$
Gumbel-Hougaard	$\frac{\theta-1}{\theta}$	$\frac{1}{1-\tau}$	$[0,1]$	$\frac{\theta}{t \log t}$
Frank	$1 + 4 \frac{(D(\theta)-1)}{\theta}$	-	$[-1,1]$	$\frac{-\log((e^{-\theta t}-1))}{(e^{-\theta}-1)}$ $\times \frac{e^{-\theta t}-1}{\theta e^{-\theta t}}$
Gumbel-Barnett	$1 + 4 \int_0^1 \log(1-\theta \log(t))$ $\times \frac{t(\log(t)-\theta)}{\theta \log(t)} dt$	-	$(0,1)$	$\frac{\log(1-\theta \log(t))}{-\theta/t/(1-\theta \log(t))}$

Donde la función de Debye está dada por  $D(-\theta) = \frac{1}{\theta} \int_0^\theta \frac{t}{e^t-1} dt$

Cuadro A.3:  $\tau$  de Kendall para algunas familias paramétricas

## Apéndice B

# Algoritmos en Matlab

```
%SimulaCopula.m
%Simula observaciones de Cópulas (fam,a,nu,n)
function [u,v] = SimulaCopula(varargin)
    fam = varargin{1};
    a = varargin{2};
    if nargin == 4
        nu = varargin{3};
        n = varargin{4};
    else
        n = varargin{3};
    end
    flag = validParam(fam,a);
    if flag ~= 1
        disp('Parametro incorrecto');
        u = 0;
        v = 0;
    else
        switch fam
            %Ali-Mikhail-Haq  $C(u,v)=u*v/(1-a*(1-u)*(1-v))$ 
            case 'AM'
                % a in (0,1)
                u1 = rand(n,1);
                t = rand(n,1);
                u2 = condCDFInv(@amCond,u1,t,a);
            %Clayton  $(u^{-a}+v^{-a}-1)^{-1/a}$ ,
            %du= $(u^{-a}+v^{-a}-1)^{-(a+1)/a}*u^{-a-1}$ 
            case 'Cl'
                u1 = rand(n,1);
                t = rand(n,1);
                u2 = u1.*(t.^(-a/(1+a)) - 1 + u1.^a).^(-1/a);

            %Gumbel-Hougaard
            % logC =  $-((-\log(u)).^a + (-\log(v)).^a).^(-1/a)$ ;
            case 'GH'
                u1 = rand(n,1);
```

```

        t = rand(n,1);
        u2 = condCDFinv(@ghCond,u1,t,a);
%Frank  $-(1/a)*(\log(1+(\exp(-a*u)-1)*(\exp(-a*v)-1)/$ 
%  $(\exp(-a)-1)))$  du=  $(\exp(-a*v)-1)*\exp(-a*u)/$ 
%  $(\exp(-a)+\exp(-a*u-a*v)-\exp(-a*u)-\exp(-a*v))$ 
        case 'Fk'
            u1 = rand(n,1);
            t = rand(n,1);
            u2 = -log((exp(-a.*u1).*(1-t)./t + exp(-a))./
(1 + exp(-a.*u1).*(1-t)./t))./a;
%Gumbel-Barnett  $u*v*\exp(-a*\log(u)*\log(v))$ 
%du=  $-v*u^{(-a*\log(v))*(-1+a*\log(v))}$ 
        case 'GB'
            u1 = rand(n,1);
            t = rand(n,1);
            u2 = condCDFinv(@gbCond,u1,t,a);
% Familias no arquimedeanas<<<<%
% Cuadras-Auge (adaptación para CA
% generalizada Nelsen(1999)pp47
        case 'CA'
            r = rand(n,1);
            s = rand(n,1);
            t = rand(n,1);
            u1 = max(r.^(1/(1-a)),t.^(1/a));
            u2 = max(s.^(1/(1-a)),t.^(1/a));
%FMG
        case 'FM'
            u1 = rand(n,1);
            t = rand(n,1);
            u2 = condCDFinv(@fmCond,u1,t,a);
%(algoritmo Johnson(1986) en Nelsen(1999)pp77
%xa = a.*(1-2.*u1)-1;
%xb = sqrt((xa.^2)-4.*(xa+1).*t);
%u2 = 2.*t./(xb-xa);

        case 'Nd'
            NormD = normcdf(mvnrnd([0 0],[1 a; a 1],n));
            u1 = NormD(:,1);
            u2 = NormD(:,2);

        case 'Td'
            stdtD = tcdf(mvtrnd([1 a; a 1],nu,n),nu);
            u1 = stdtD(:,1);
            u2 = stdtD(:,2);

        otherwise
            u1 = rand(n,1);
            u2 = rand(n,1);
    end
    u = u1;

```

```

v = u2;
end
end

%Función inversa por método de bisección
function u2 = condCDFinv(condCDF,u1,t,a)
    infi = zeros(size(t));
    supr = ones(size(t));
    dist = 1;
    tol = 1e-12;
    while dist > tol
        u2 = .5 .* (infi + supr);
        lo = feval(condCDF,u1,u2,a) < t;
        infi(lo) = u2(lo);
        supr(~lo) = u2(~lo);
        dist= .5 .* dist;
    end
end
function t = ghCond(u1,u2,a)
% -((-log(u))^a + (-log(v))^a)^
% (-(-1+a)/a)*(-log(u))^a/u
%/log(u)*exp(-((-log(u))^a+(-log(v))^a)^(1/a))
    l1 = -log(u1);
    l2 = -log(u2);
    t = -exp(-(l1.^a + l2.^a).^(1/a)).*
        ((l1.^a + l2.^a).^(1-a/a)).*
        (l1.^a)./u1./log(u1);
end

function t = gbCond(u1,u2,a)
% du = -v*u^(-a*log(v))*(-1+a*log(v))
    l2 = log(u2);
    t = -u2.*u1.^(-a.*l2).*(-1 + a*l2);
end

function t = amCond(u1,u2,a)
    t = (1-a+a.*u2).*u2./
(a- a*u1 + a*u1.*u2 - a*u2-1).^2;
end

function t = fmCond(u1,u2,a)
    t = u2.*(1 + a*(1-u2).*(1-2*u1));
end

%. . . . .
%tauParam.m
%retorna parámetro estimado a partir de tau de Kendall
%y valor de tau de Kendall
function [a,t] = tauParam(fam,u,v)
tau=corr(u,v,'type','Kendall');

```

```

%tau = tauKendall(u,v);
t=tau;
switch fam
    %<<<Ali-Mikahil-Haq C(u,v)=u*v/(1-a*(1-u)*(1-v))
    case 'AM'
        a = fzero(@(alpha) RootFun(alpha,tau,fam),tau);

    %<<<<Clayton (u^(-a)+v^(-a)-1)^(-1/a)
    case 'Cl'
        a = (2*tau)/(1-tau);

    %<<<<Gumbel-Hougaard exp((-(-log(u))^a-(-log(v))^a)/a)
    case 'GH'
        a = 1/(1-tau);

    %<<<<Frank -(1/a)*(log(1+(exp(-a*u)-1)*(exp(-a*v)-1)
    % / (exp(-a)-1)))
    case 'Fk'
        a = fzero(@(alpha) RootFun(alpha,tau,fam),tau);

    %<<<<Gumbel-Barnett
    case 'GB'
        a = NaN;

    %Familias no arquimedeanas
    %<<<<Farlie Morgensten
    case 'FM'
        a = 9*tau /2;

    case 'CA'
        a = 2*tau/(1+tau);

    case 'Nd'
        a = sin(tau*pi/2);

    otherwise
        a = NaN;
end
end

%funcion para encontrar la inversa del parámetro
function err = RootFun(alpha,ttau,fam)

switch fam
    case 'Fk'
        tau = 1 + 4 .* (debye1(alpha)-1) ./ alpha;
    case 'AM'
        tau = -1/3* (-3*alpha^2 + 2*log(1-alpha)
* alpha^2
+ 2*alpha - 4*log(1-alpha)*alpha

```



```

+ 2*log(1-alpha)) / alpha^2;
    case 'GB'
        t = sym('t');
        tau = 1 + 4*int(log(1-alpha*log(t))*t
*(log(t)-alpha)/alpha*log(t),0,1);
    end
    err = tau - ttau;
end

%.....
%MaxV.m
%Funcion MaxV(familia, datos) minimiza negativo de
%log verosimilitud. Retorna parametro y
%valor de log verosimilitud
function [mlest,mlval]= MaxV(fam,u,v)
    % if (fam~='CA')
        [inf,sup] = validParam(fam);
        [mlest,mlval] = fminbnd(@(a
copulaNegL(fam,a,u,v),inf,sup);
    %else
        % mlest = NaN;
        % mlval = NaN;
    % end
end

%copulaNegL.m
%Función copulaNegL( familia, datos, parámetro)
%evalua negativo f verosimilitud retorna matriz
%de evaluaciones y suma de ellas
function [val,c12] = copulaNegL(fam,a,u,v)
tol= 1.0000e-16;
c12=0;
    switch fam
        %<<<Ali-Mikahil-Haq C(u,v)=u*v/(1-a*(1-u)*(1-v))
%dc(u,v)=-(-a^2*v+a^2*u*v+a*v-2*a+a*u+a*u*v+1+a^2-a^2*u)/
%(a-a*u+a*u*v-a*v-1)^3
        case 'AM'
            c12 = -(-a^2*v + a^2*u.*v + a*v- 2*a +
a*u + a*u.*v + 1 + a^2 - a^2*u)./
(a - a*u + a*u.*v - a*v-1).^3;
            c12= -log(c12);
            %<<<<Clayton (u^(-a)+v^(-a)-1)^(-1/a)
        case 'Cl'
            % fact = 1;
            fact= -1 + u.^(-a)+ v.^(-a);
            c12= (1+a)*(u.^(-1-a)).*(v.^(-1-a)).*
fact.^(-2-1/a);
            c12= -log(c12);
            %<<<<Gumbel-Hougaard exp(-((-log(u))^a +
%(-log(v))^a).^(-1/a))

```

```

        %dC(u,v) = %(-1+a+((-log(u))^a+(-log(v))^a)
%^(1/a))/u/log(u)*(-log(v))^a/
%v*(-log(u))^a*((-log(u))^a+(-log(v))^a)^(-2+1/a)/
%log(v)*exp(-((-log(u))^a+(-log(v))^a)^(1/a))
        case 'GH'
            lua = (-log(u)).^a;
            lva = (-log(v)).^a;
            lm = lua + lva;
            c12 = (-1 + a + lm.^(1/a))./u./log(u).*lva./
v.*lua.*lm.^(-2 + 1/a)./log(v).*
exp(-lm.^(1/a));
            %c12 = ((-log(u)).^(a-1)).*((-log(v)).^(a-1)).*
%(lm.^(-2 + 1/a)).*
%(-1 + a + lm.^(1/a))./((exp(lm).^(1/a)).*u.*v);
            c12 = -log(c12);

        %<<<<<Frank -(1/a)*(log(1+(exp(-a*u)-1)*
%(exp(-a*v)-1)/(exp(-a)-1)))
        case 'Fk'
            c12 = -(-a*(exp(-a*(u+v+1))-exp(-a*(u+v)))./
(-exp(-a)-exp(-a*(u+v)))+
exp(-a*u)+exp(-a*v)).^2);
            c12= -log(c12);

        %<<<<<Gumbel-Barnett
        %u*v*exp(-a*log(u)*log(v))
        %dGB = u^(-a*log(v))*(1-a*log(u)-a*log(v)
%-a+a^2*log(v)*log(u))
        case 'GB'
            lu = log(u);
            lv = log(v);
            c12 = u.^(-a*lv).*(1 - a*lu -a*lv -a + a^2 *lu.*lv);
            c12= -log(c12);
        %>>>Familias no arquimedeanas<<<<<%
        %<<<<<Cuadras-Auge
        case 'CA'
            mmo1 = max(u.^(1/(1-a)),v.^(1/(1-a)));
            mmo2 = max(u.^(1/a),v.^(1/a));
            mmo = min(mmo1, mmo2);
            c12 = mmo.^-a;
            c12 = -log(c12);
        %<<<<<FMG
        case 'FM'
            c12 = 1+a*(1-u).*(1-v)-a*v.*(1-u)-a*u.*(1-v)+a*u.*v;
            c12 = -log(c12);
        otherwise
            disp('Familia no identificada');
    end
    val = nansum(c12);
end

```

# Bibliografía

- [1] Nelsen, R. *An introduction to copulas*, Springer-Verlag. 1999
- [2] Joe, H. *Multivariate Models and Dependence Concepts*. Chapman & Hall. 1997
- [3] Erdely, A. *Construcción de medidas de dependencia por medio de cópulas*. Tesis UNAM. 2003
- [4] Muriel, N. *Sobre cópulas y dependencia de variables aleatorias*. Tesis UNAM. 2004
- [5] Bhat. B. *Modern probability theory*. 2a Ed. John Wiley & Sons. 1985
- [6] Mood, A., Graybill, F. , Boes, D. *Introduction to the Theory of Statistics*. McGraw-Hill. 1974
- [7] Embrechts, P., Lindskog, F., McNeil, A. *Modelling dependence with copulas and applications to risk management*. Department of Mathematics, ETHZ, Zürich, 2001
- [8] Genest, C., Goudi, K. , Rivest, L-P. 'A semiparametric estimation procedure for parameters in multivariate families of distribution.' *Biometrika*, Vol.82, Issue 3, 1995
- [9] Genest, C. ,Quessy, J-F. , Remillard, B. 'Goodness-of-fit Procedures for Copula Models Based on the Probability Integral Transformation.' *Board of the Foundation of the Scandinavian Journal of Statistics*, Blackwell, Vol 32. 2005
- [10] Davison. A., Hinkley. D. *Bootstrap Methods and their Applicattions*. Cambridge University. 1997
- [11] De Matteis, R. *Fitting copulas to data*. Diploma thesis. Institute of Mathematics of the University of Zurich. 2001
- [12] Hull, J. *Options, Futures & other Derivatives*. 4a Ed. Prentice-Hall. 2000